

Министерство науки и образования Украины
Донбасская государственная машиностроительная академия

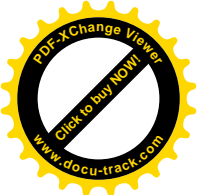
Е.Ю. Ивченко

МОДЕЛИРОВАНИЕ СЛОЖНЫХ СИСТЕМ

конспект лекций

Утверждено
на заседании кафедры ИСПР
Протокол № 2 от 09.09.2014 г.

Краматорск 2014

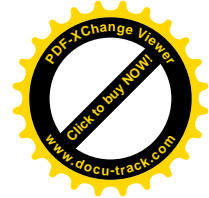


Конспект лекций по дисциплине «Моделирование сложных систем» / для студентов направления 6.040303 «Системный анализ» очной и заочной формы обучения / Составители: Е.Ю.Ивченкова. – Краматорск: ДГМА, 2014.

Конспект по дисциплине «Моделирование сложных систем» предназначены для студентов направления 6.040303 «Системный анализ» всех форм обучения, и всех специалистов имеющих дело с обработкой экономической информации при помощи компьютерной техники.

Составитель: Ивченкова Е.Ю., к.э.н., ст. преподаватель

Отв. за выпуск Мельников А.Ю., к.т.н., доц., и.о. зав. кафедрой



Модуль 1 Основные понятия моделирования и управления системами

Тема 1.1. Основные положения теории систем

ЛЕКЦИЯ 1. СУЩНОСТЬ И ПРИНЦИПЫ СИСТЕМНОГО ПОДХОДА К ИССЛЕДОВАНИЮ СЛОЖНЫХ ДИНАМИЧЕСКИХ ОБЪЕКТОВ.

1. Понятие системы. Принцип «черного ящика»

Определение любой конкретной системы является произвольным. По существу, вся вселенная состоит из множества систем, каждая из которых содержится в более крупной системе подобно множеству пустотелых кубиков, вложенных друг в друга. Так же, как всегда можно представить себе более обширную систему, куда входит данная, всегда можно выделить из данной системы более ограниченную.

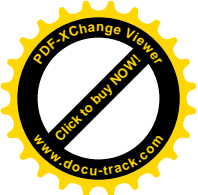
Следовательно, при стремлении исследовать все воздействия, влияющие на какой-либо единичный материальный объект, необходимо рассматривать его как часть некоторой системы. Эта система является системой потому, что она состоит из взаимосвязанных частей и в определенном смысле представляет собой замкнутое целое.

Система – целостный комплекс взаимосвязанных элементов, имеющий определенную *структуру* и взаимодействующий с *внешней средой*.

Внешняя среда – все то, что не вошло в систему.

Выделение системы носит субъективный характер и требует наличия:

- *объекта*, состоящего из множества элементов, связанных в некоторую совокупность;
- *субъекта*, исследующего поведение объекта;



Министерство науки и образования Украины
Донбасская государственная машиностроительная академия

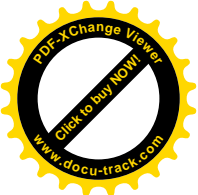
Е.Ю. Ивченко

МОДЕЛИРОВАНИЕ СЛОЖНЫХ СИСТЕМ

конспект лекций

Утверждено
на заседании кафедры ИСПР
Протокол № 2 от 09.09.2014 г.

Краматорск 2014

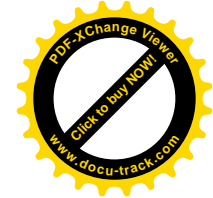


Конспект лекций по дисциплине «Моделирование сложных систем» / для студентов направления 6.040303 «Системный анализ» очной и заочной формы обучения / Составители: Е.Ю.Ивченкова. – Краматорск: ДГМА, 2014.

Конспект по дисциплине «Моделирование сложных систем» предназначены для студентов направления 6.040303 «Системный анализ» всех форм обучения, и всех специалистов имеющих дело с обработкой экономической информации при помощи компьютерной техники.

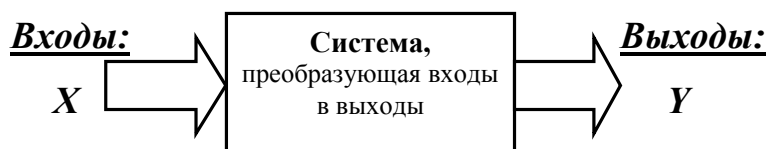
Составитель: Ивченкова Е.Ю., к.э.н., ст. преподаватель

Отв. за выпуск Мельников А.Ю., к.т.н., доц., и.о. зав. кафедрой



- *цели исследования*, характеризующей отношение наблюдателя к объекту и определяющей набор рассматриваемых элементов и их существенных свойств.

В наиболее общем случае система может быть представлена как «черный ящик». Данный термин ввел английский ученый-кибернетик У. Эшби.



«**Черным ящиком**» именуют такой объект, внутренняя структура которого не известна, поведение его описывается только зависимостью входов от выходов.

Входы – каналы, с помощью которых внешняя среда влияет на систему; через входы из внешней среды в систему поступают вещество, энергия или информация.

Выходы – каналы влияния системы на внешнюю среду; через *выходы* Y результаты процессов преобразования входов поступают во внешнюю среду.

Введем обозначения:

- множество входов системы: $X = \{X_i\} = \{X_1, X_2, \dots, X_n\}$;
- множество выходов системы: $Y = \{Y_j\} = \{Y_1, Y_2, \dots, Y_m\}$

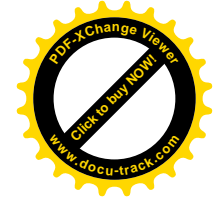
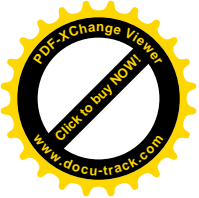
Отсюда система типа «черного ящика» может быть определена как:

$$Y = \{R\}X \quad (1)$$

где $\{R\}$ - преобразователь входов в выходы.

Преимущества концепции «черного ящика»: дает возможности изучения поведения систем, внутреннее устройство которых неизвестно или не является предметом изучения (например, станок с ЧПУ, доменная печь).

Недостатки концепции: возможности исследования «черного ящика» достаточно ограничены, т.к. более детальное исследование предполагает знание структуры системы.



2. Элемент и структура системы

Первичным элементом системы является объект, выполняющий определенные функции и не подлежащий дальнейшему разбиению в границах данного исследования.

Концепция первичного элемента позволяет производить структурный анализ системы, причем элементы являются модулями структуры, «черными ящиками», внутренняя структура которых не является предметом исследования. Целостность системы обеспечивается связями между ее элементами.

Состав элементов и способ их объединения определяют *структуру системы*. Формализуем понятие структуры системы.

Система представляет собой совокупность конечного числа подсистем:

$$S = \{S_i\}, i = \overline{1, I} \quad (2)$$

где $\forall S_i \in S$;

$$\bigcap_i S_i = \emptyset;$$

$$\bigcup_i S_i = S.$$

Все подсистемы взаимосвязаны между собой:

$$\Delta = \langle S; (S_i, S_j) \rangle, i \neq j; i, j \in I \quad (3)$$

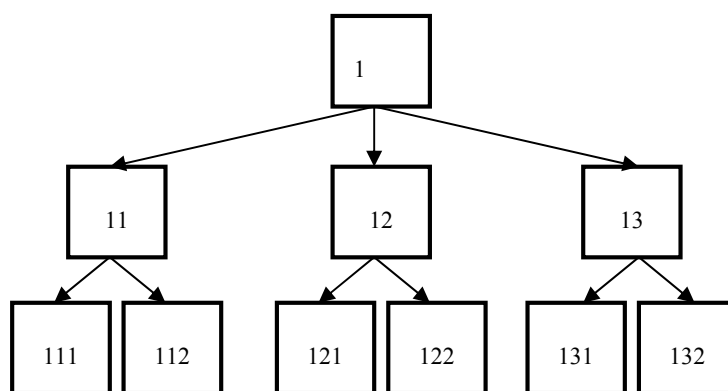
где Δ - отношение связи.

Отсюда в наиболее общем виде структура системы определяется следующим образом:

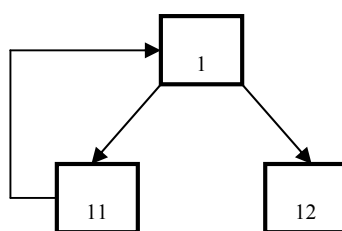
$$G = \langle \{S_i\}, \Delta \rangle \quad (4)$$

Связи, обеспечивающие взаимодействия элементов системы между собой и с внешней средой, разнообразны так же, как свойства объекта и среды. При анализе и синтезе системы принимают во внимание лишь *существенные* связи, а прочими пренебрегают, рассматривая их как помехи или «шум».

Структура системы является наиболее устойчивой ее характеристикой.



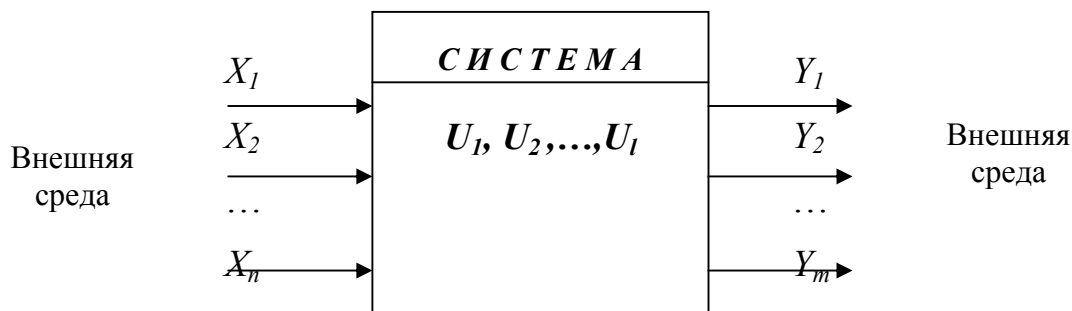
Структура системы без обратной связи



Структура системы с обратной связью

Внутренняя структура системы определяет ее состояния.

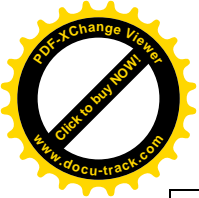
Состояние системы – внутренняя характеристика системы, определяемая совокупностью параметров.



$U = \{U_i\} = \{U_1, U_2, \dots, U_l\}$ - состояния системы.

Классификация систем

Классификационный признак	Виды систем
1. Способ преобразования входных параметров в выходные	<ul style="list-style-type: none"> • дискретные; • непрерывные



2. Наличие связи с внешней средой	<ul style="list-style-type: none">• закрытые;• открытые
3. Изменение состояния системы во времени	<ul style="list-style-type: none">• статические;• динамические
4. Природа поведения системы и ее элементов	<ul style="list-style-type: none">• детерминированные;• стохастические
5. Неоднородность, разнокачественность элементов и связей	<ul style="list-style-type: none">• простые;• сложные;• очень сложные (большие)

В соответствии с представленной классификацией введем ряд определений.

1. **Функциональная система.**

Если S является функцией:

$$S: X \rightarrow Y \quad (5)$$

где $X = \{X_{ij}\}$ – входы системы;

$Y = \{Y_{ij}\}$ – выходы системы,

то соответствующая система называется *функциональной*.

Примечание. Иными словами, такая система представляет собой «черный ящик», где преобразование входов в выходы задается некоторой функцией:

$$Y = f(X) \quad (6)$$

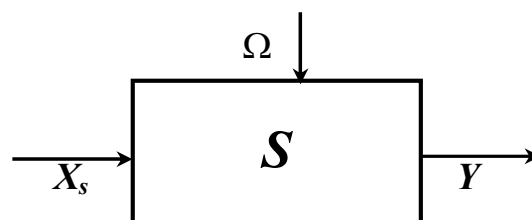
Функциональная система может быть как дискретной, так и непрерывной.

2. **Дискретная система** – система, в которой преобразование входов в выходы рассматривается в фиксированные моменты времени.

3. **Непрерывная система** – преобразование входов в выходы рассматривается как непрерывный процесс.

В наиболее общем случае функциональная система как объект кибернетики является открытой, динамической, стохастической.

4. **Открытая система** – система, состояние которой зависит от факторов внешней среды.



Множество входных воздействий X включает в себя два подмножества:

- управляющих воздействий X_s - набор факторов внутренней среды, сформированных в ответ на внешние воздействия;
- возмущающих воздействий Ω – набор факторов внешней среды, наиболее значимых для системы в данное время;

Отсюда функциональная модель системы принимает вид:

$$S : X \times \Omega \rightarrow Y \quad (7)$$

или модель открытой системы определяется как функция:

$$y = f(x_s, \omega) \quad (8)$$

где x_s – управляющие воздействия;

ω – возмущающие воздействия.

Если учесть, что на выходы системы влияют не только входы, но и внутреннее состояние системы, имеем:

$$y = f(x_s, \omega, u) \quad (9)$$

5. **Динамическая система** – система, состояние которой изменяется во времени. В любой момент времени состояние такой системы суть функция:

$$u(t) = \varphi[t, \tau, u(\tau)]; \tau < t; t, \tau \in T \quad (10)$$

где t - произвольный момент времени;

τ - момент времени, предшествующий рассматриваемому;

$u(\tau)$ - состояние системы в предшествующий момент времени.

Отсюда моделью динамической открытой системы является функция:

$$y(t) = f[x_s(t), \omega(t), u(t)]; t \in T \quad (11)$$



где $x_s(t)$ – управляющие воздействия в момент времени t ;

$\omega(t)$ – возмущающие воздействия в момент времени t ;

$u(t)$ – состояние системы в момент времени t (определяется формулой (10)).

6. Стохастическая (вероятностная) система – система, для которой переход из состояния в состояние осуществляется случайным образом. Моделью стохастической открытой динамической системы является случайная функция:

$$y(t) = \xi[x_s(t), \omega(t), u(t)]; t \in T \quad (12)$$

Рассмотренные свойства (5) – (12) лежат в основе определения сложной системы.



ЛЕКЦИЯ 2. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ СИСТЕМЫ

Важнейшей характеристикой системы является *разнообразие*, определяющее число различных состояний системы.

Если разнообразие системы превосходит возможности анализа и синтеза всех ее элементов и связей между ними, то такая система называется **большой системой**. Как правило, большие системы являются системами *сложными*. Формализуем понятие сложной системы.

Сложная система представляет собой динамическую, открытую, вероятностную систему, определяемую следующими параметрами:

$$S = \langle T, \Phi, X_s, \Omega, U, Y, G, R \rangle \quad (13)$$

где T - множество моментов времени;

Φ - макрофункция системы;

X_s – множество управляющих воздействий;

Ω – множество возмущающих воздействий;

U – множество состояний системы;

Y - множество значений выходных величин;

G - структура системы;

R - отношение эмерджентности.

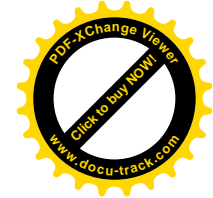
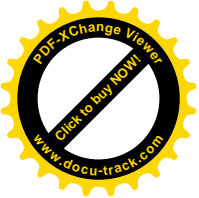
1. **T - множество моментов времени**, таких, что между ними задано отношение порядка (множество упорядоченных моментов времени)

$$\varepsilon = [T, (t_k, t_p)] \Big| t_k \neq t_p; t_k, t_p \in T \quad (14)$$

2. **Φ - макрофункция системы** является количественным выражением цели системы и зависит от управляющего воздействия X_s . Выбор макрофункции обеспечивает достижения требуемого значения Y_0 . Таким образом, макрофункция связана с решением глобальной задачи, стоящей перед системой:

$$\Phi : X_s \rightarrow Y_0, X_s \subset X, Y_0 \subset Y \quad (15)$$

Обоснование выбора определенного вида макрофункции производится в соответствии с некоторым эвристическим критерием ψ :



пусть $\{\Phi_1, \Phi_2, \dots, \Phi_J\}$ – некоторый конечный набор функций, связанных с целью системы S :

$$\Phi = \underset{\Psi}{extr} \{ \Phi_j \}, j = \overline{1, J} \quad (16)$$

3. **Множество возмущающих воздействий Ω** непусто:

$$\Omega \neq 0 \quad (17)$$

1. Существует **переходная функция U** , значениями которой служат состояния (см. формулу (10)) системы в момент времени $t \in T$, если в начальный момент времени $\tau < t$ она находилась в состоянии $u(\tau) \in U$, и в течение отрезка $[\tau, t)$ на нее действовали входные воздействия $x \in X$:

$$u(t) = \varphi(t, \tau, u, x), u(t) \in U \quad (18)$$

Примечание. В наиболее общем случае переходная функция является случайной (вероятностной).

2. **Структура системы G** в силу многообразия связей и элементов системы определяет наиболее существенные связи Δ между ее элементами (см. формулы (2) – (4)):

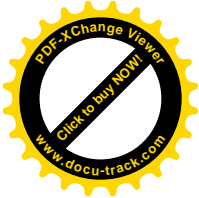
$$G = \langle \{S_i\}, \Delta \rangle \quad (19)$$

3. **Отношение эмерджентности R** задает соответствие между макрофункцией системы и реализующей ее структурой и изменяется всякий раз, когда это соответствие нарушается:

$$R : \Phi \rightarrow G \quad (20)$$

Эмерджентность – свойство системы, состоящее в том, что системе как целостности, присущи такие характеристики, которые не имеют отдельные ее элементы.

Основные характеристики сложной системы: эмерджентность, многоуровневость, динамизм, стохастичность, инерционность, открытость и адаптивность.



**Основные
характеристи
ки сложной
системы**

эмерджентность

свойство системы, состоящее в том, что системе как целостности, присущи такие характеристики, которые не имеют отдельные ее элементы

многоуровневость

наличие различных уровней исследования системы, каждый из которых имеет свои особенности

динамизм

быстрое изменение параметров системы во времени

стохастичность

вероятностный характер поведения системы

инерционность

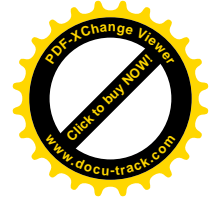
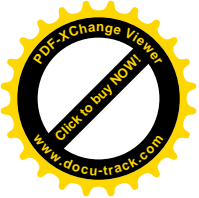
запаздывание реакции системы на управляющие воздействия

открытость

наличие взаимосвязи с «внешней средой»

адаптивность

способность генерировать сигналы, компенсирующие возмущающие воздействия, направленные на изменение их текущих свойств



Тема 1.2. Моделирование систем

ЛЕКЦИЯ 3. ПОНЯТИЕ МОДЕЛИ СИСТЕМЫ, КЛАССИФИКАЦИЯ МОДЕЛЕЙ

1 Концепция модели. Изоморфизм и гомоморфизм

В процессе познания окружающего мира у человека вырабатываются представления о тех или иных свойствах окружающих объектов и их взаимосвязях. Они формируются в виде описаний этих объектов на обычном языке, фиксируются в виде рисунков, графиков, формул или реализуются как макеты или другие устройства. Подобные способы описания обобщаются в едином понятии – модель.

Модель – представление объекта в некоторой форме, отличной от формы их реального существования.

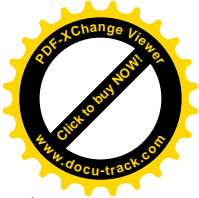
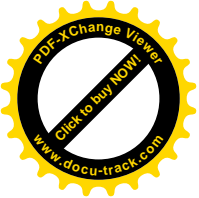
Моделирование – построение и изучение моделей объектов.

Предметом моделирования могут быть конкретные и абстрактные объекты, действующие и проектируемые системы.

В кибернетическом моделировании основную роль играет сходство поведения и/или структуры оригинала и модели, различия в содержании не имеют определяющего значения, поскольку сходные зависимости между входами и выходами могут быть, по определению, реализованы объектами различной природы. При анализе поведения системы обычно стремятся к следующему:

- изучить поведение объекта, как целостной системы, используя модель «черного ящика»;
- определить причины поведения объекта с помощью моделей, раскрывающих его внутреннюю структуру.

Моделирование приобретает особое значение при изучении объектов, недоступных в полной мере прямому наблюдению или экспериментированию (в частности, экономических объектов).



Если между двумя объектами может быть установлено сходство хотя бы в каком-либо одном определенном смысле, то один из этих объектов может рассматриваться как оригинал, а другой – как его модель. Отношение «оригинал-модель» могут иметь место не только между двумя, но и между любым числом объектов. Оценка адекватности пары «оригинал-модель» может быть осуществлена с использованием понятий *изоморфизма* и *гомоморфизма*.

Рассмотрим оригинал объекта как «черный ящик» с вектором входов $X(t)$ и выходов $Y(t)$. При длительном наблюдении за поведением объекта можно достигнуть такого уровня знаний его свойств, чтобы предсказывать его выходы при различных заданных изменениях его входов. Однако даже при самом длительном изучении «черного ящика» невозможно получить однозначное представление о его внутреннем устройстве, поскольку одни и тем же поведением могут обладать системы различной структуры и физической природы.

Рассмотрим две системы A и B , описываемые следующими векторами:

- входы $X_A(x_{1A}, x_{2A}, \dots, x_{mA})$ и $X_B(x_{1B}, x_{2B}, \dots, x_{mB})$;
- выходы $Y_A(y_{1A}, y_{2A}, \dots, y_{nA})$ и $Y_B(y_{1B}, y_{2B}, \dots, y_{nB})$ соответственно.

Говорят, что системы A и B **изоморфны**, если для любого момента времени $t \in T$ соблюдается условие:

$$\begin{cases} X_{1A}(t) = X_{1B}(t) \\ X_{2A}(t) = X_{2B}(t) \\ \dots\dots\dots \\ X_{nA}(t) = X_{nB}(t) \end{cases} \Rightarrow \begin{cases} Y_{1A}(t) = Y_{1B}(t) \\ Y_{2A}(t) = Y_{2B}(t) \\ \dots\dots\dots \\ Y_{mA}(t) = Y_{mB}(t) \end{cases} \quad (1)$$

Изоморфными называются системы, одинаковые для наблюдателя, следящего только за их входами и выходами. Здесь изоморфизм трактуется как идентичность поведения рассматриваемых систем в некоторой среде, с которой они контактируют через свои входы и выходы. При этом системы A и B могут иметь разную физическую природу (примеры: местность и географическая карта, объект съемки и фотография, настоящий самолет и его макет).



Однако изоморфизм не является необходимым условием соответствия модели оригиналу. Частным случаем отношения «оригинал-модель» является *гомоморфизм*.

Гомоморфизм характеризует однозначное соответствие между состоянием систем А и В и неоднозначно-обратное соответствие (рис.1).

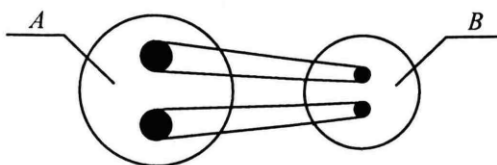


Рис. 1. Гомоморфные системы

Пусть система А - оригинал. Система В - модель системы А, полученная из системы А путем ее упрощения (уменьшения числа рассматриваемых переменных за счет их объединения). Система А имеет размерность n , а В - размерность $m < n$.

Если при этом каждому состоянию А отвечает одно определенное состояние В, но не каждому состоянию В отвечает одно состояние А, то систему В называют *гомоморфной моделью* системы А.

Итак, *модель* — это гомоморфный образ оригинала.

2 Экономико-математические модели. Их классификация

Моделирование — метод, повышающий эффективность управления системой как кибернетическим объектом. Математическое моделирование позволяет изучать и оптимизировать поведение исследуемой системы путем формального представления ее свойств, целевой функции управления, накладываемых ограничений.

Экономико-математическое моделирование рассматривает построение моделей, используемых для управления экономическими объектами.

Экономико-математическая модель - это формализованное описание различных экономических явлений, исследование которых позволяет получить необходимую информацию для реализации целей управления моделируемой системой.



Требования, предъявляемые к экономико-математическим моделям:

- построение на основе экономической теории и отражение ее объективных законов;
- правильное отображение функции и/или структуры изучаемой реальной экономической системы;
- удовлетворение определенным математическим условиям (разрешимость, согласованность размерностей и т.д.).

Как и всякая модель, экономико-математическая модель является упрощенным отражением реальности, выделяя наиболее существенные факторы (переменные). Переменные в экономико-математических моделях делятся на 2 группы:

1. *Экзогенные (внешние, независимые)* - их величина определяется вне модели.
2. *Эндогенные (внутренние, зависимые)* - их значения устанавливаются в результате решения модели.

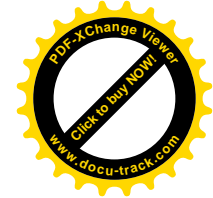
Классификация экономико-математических моделей

Классификационный признак	Типы моделей
1. Учет фактора времени	<ul style="list-style-type: none">• статические;• динамические
2. Наличие запаздывания между входом и выходом	<ul style="list-style-type: none">• кинематические (безынерционные);• инерционные
3. Определение состояния объекта	<ul style="list-style-type: none">• дескриптивные;• нормативные

Статическая экономико-математическая модель в общем виде может быть представлена так:

$$Y = F[x_s, \omega, \alpha] \quad (2)$$

где x_s – экзогенные управляемые переменные (учтенные факторы);



ω – экзогенные неуправляемые переменные (неучтенные факторы);

α - параметры модели;

Y - эндогенные переменные (выходы, отклики);

F - определяет вид функциональной зависимости эндогенных переменных от экзогенных.

При изучении динамики экономической системы в формулу (2) добавляется фактор времени:

$$Y(t) = F[x_s(t), \omega(t), \alpha] \quad (3)$$

В свою очередь, динамические модели делятся на *кинематические* и *инерционные*.

Кинематическая (безынерционная) модель – модель, основанная на предположении, что реакция системы на входные воздействия является мгновенной, то есть фактически существующие в любой реальной системе запаздывания (лаги) между входом и выходом не учитываются.

Инерционная модель учитывает запаздывание реакции выходов на изменение входов в системе.

Формально отличия между кинематической и инерционной моделью можно представить следующим образом:

- инерционная модель

$$u(t) = \varphi[t, \tau, x(t), u(\tau)]; \tau \prec t; t, \tau \in T \quad (4)$$

где t - произвольный момент времени;

τ - момент времени, предшествующий рассматриваемому;

$x(t)$ – набор входных воздействий в момент времени t ;

$u(\tau)$ - состояние системы в предшествующий момент времени.

- кинематическая модель

$$u(t) = \varphi[x(t)]; t \in T \quad (5)$$

где t - произвольный момент времени;

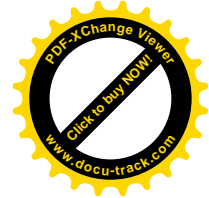
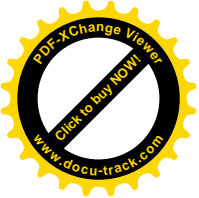
$x(t)$ – набор входных воздействий в момент времени t .

Дескриптивная модель – модель, предназначенная для объяснения



наблюдаемых факторов или прогноза поведения объекта (например, модель однофакторной (тренд) и многофакторной регрессии, имитационная модель).
Дескриптивные модели отвечают на вопросы «как это происходит, как будет развиваться?»

Нормативная модель – модель, предназначенная для определения наилучшего или допустимого с точки зрения исследователя состояния объекта (например, оптимизационная модель). Такие модели отвечают на вопросы «как должно быть?»



ЛЕКЦИЯ 4. МЕТОДИКА МОДЕЛИРОВАНИЯ СИСТЕМ

Моделирование представляет собой сложный трудоемкий процесс, состоящий из следующих *этапов*:

1. Постановка задачи.
2. Построение математической модели.
3. Подготовка исходной информации.
4. Численное решение.
5. Анализ полученного решения и уточнение модели.
6. Практическая реализация модели.

1. Постановка задачи.

Постановка экономической проблемы и ее качественный анализ предполагает:

- изучение объекта моделирования путем выделения его основных свойств, структуры (предприятие или его подразделения, регион, отрасль экономики и пр.);
- определение цели моделирования, которая состоит в исследовании (и, возможно, оптимизации) одного или нескольких процессов, протекающих в моделируемом объекте;
- изучение различных подходов к реализации поставленной цели, которые уже применялись или могут быть применены.

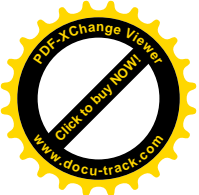
2. Построение математической модели.

Формализация экономической проблемы, т.е. выражение ее в виде конкретных математических зависимостей:

- определение типа экономико-математической модели;
- изучение возможностей ее применения в данной задаче;
- уточнение перечня переменных и формы связей.

3. Подготовка исходной информации.

Наиболее трудоемкий этап. Математическое моделирование предъявляет жесткие требования к информации. В этой связи при подготовке информации



используются методы теории вероятностей и математической статистики для организации выборочных обследований, оценки достоверности данных.

4. Численное решение.

Получение числовых характеристик модели и расчеты исследуемых экономических показателей:

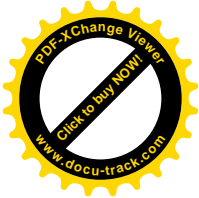
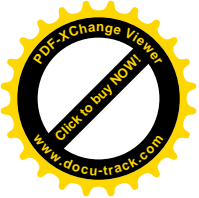
- разработка алгоритма численного решения задачи;
- компьютерная реализация модели.

5. Анализ полученного решения и уточнение модели.

Решение вопроса о правильности результатов моделирования и применимости их для управленческих решений:

- *верификация* – ее проведение убеждает в том, что модель ведет себя так, как было задумано;
- *оценка адекватности* – проверка соответствия поведения модели поведению реальной системы;
- *проблемный анализ* – формулировка значимых выводов на основе результатов, полученных в ходе моделирования.

Процесс моделирования носит циклический характер, свидетельствующий о необходимости возврата к предшествующим этапам. Наиболее часто это происходит на этапе подготовки исходной информации.

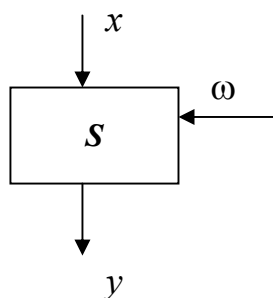


Тема 1.3. Управление системой

ЛЕКЦИЯ 5. УПРАВЛЕНИЕ СИСТЕМОЙ

1. Понятие управления системой. Объект и субъект управления

Управление системой – целенаправленное воздействие на ее поведение (состояние) при изменяющихся условиях внешней среды.

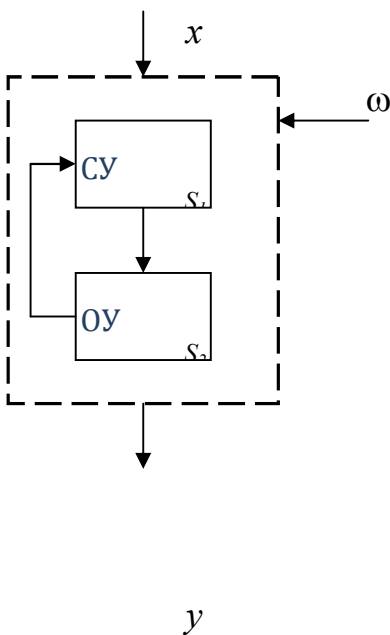


Любую систему, являющуюся объектом кибернетического исследования, можно представить как *систему управления*.

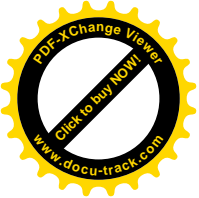
Система управления включает две подсистемы:

S_1 – субъект управления (управляющая подсистема);

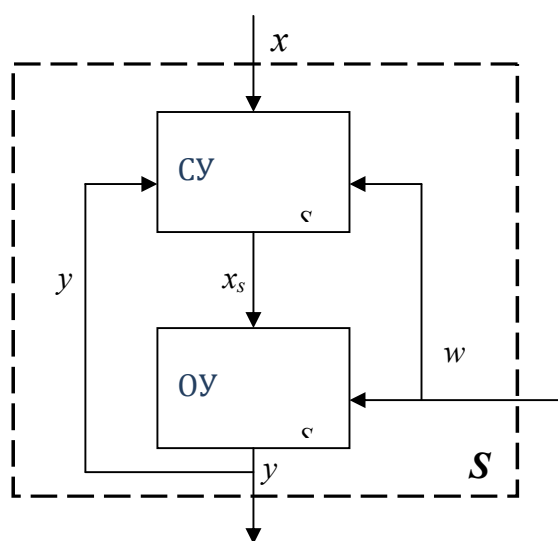
S_2 – объект управления (управляемая подсистема)



Системой управления называется организованная динамическая система с обратной связью. Управление в такой системе состоит в целенаправленном



воздействии СУ на ОУ с целью достижения последним желаемого состояния и поддержания этого состояния на заданном уровне.



Характеристика процесса управления:

- *фаза планирования (определение плановых параметров функционирования ОУ):* СУ на основе цели управления, задающей значения входов x , и априорной информации о законах функционирования системы во внешней среде A генерирует управляющие воздействия x_s :

$$S_1 : X \times A \rightarrow X_s \quad (1)$$

- *фаза учета (установление фактических параметров функционирования ОУ):* ОУ реагирует на управляющие воздействия x_s с учетом возмущающих воздействий w :

$$S_2 : X_s \times \Omega \rightarrow Y \quad (2)$$

- *фаза анализа (оценка отклонение фактических параметров от плановых):* результаты работы объекта Y по каналу обратной связи поступают на вход СУ и анализируются;



- фаза регулирования (корректировка плановых параметров функционирования ОУ): СУ, используя результаты анализа фактического состояния объекта Y и прогнозные значения возмущающих воздействий w , вырабатывает новые управляющие воздействия x_s :

$$S_l : X \times Y \times \Omega \rightarrow X_s \quad (3)$$

Итак, в общем случае состояние входов системы $x(t)$ зависит от:

- $y(t)$ - текущего состояния ОУ;
- $w(t)$ - возмущающих воздействий.

СУ генерирует управляющие воздействия с учетом как $y(t)$, так и $w(t)$.

2. Основные принципы управления

Для обеспечения эффективности процесса управления необходимо соблюдение следующих принципов.

1. Принцип необходимого разнообразия

Основная задача управления – уменьшение разнообразия в управляемой системе (ОУ).

Оценим соответствие разнообразия СУ разнообразию ОУ.

Пусть в дискретные моменты времени $t = 1, 2, \dots$ происходит изменение входов $x(t)$ (под влиянием $y(t)$ и $w(t)$), а СУ вырабатывает вектор управляющих воздействий $x_s(t)$, в результате которых состояние ОУ определяется как:

$$u(t) = \varphi[x(t), x_s(t)]; t \in T \quad (4)$$

Для того чтобы в соответствии с целью управления перевести ОУ из состояния $u(t)$ в состояние $u(t+1)$, СУ должен «прогнозировать» $x(t)$ (путем прогнозирования $y(t)$ с учетом $w(t)$) и вычислить значение $x_s(t)$:

$$u(t+1) = \varphi[x(t+1), x_s(t+1), u(t)]; t \in T \quad (5)$$

Если разнообразие задачи управления, измеряемое количеством информации, определить как V , а информационную мощность СУ как W , то для осуществления перехода $u(t) \rightarrow u(t+1)$, необходимо, чтобы в каждый момент времени выполнялось условие:



$$W(t) \geq V(t) \text{ для } \forall t \in T \quad (6)$$

Отсюда **принцип необходимого разнообразия** применительно к управлению таков: *необходимо, чтобы разнообразие СУ было не меньше разнообразия ОУ.*

2. Принцип управления воздействием на «главный» фактор

В реальных системах управления сложными объектами (в частности, экономическими) «полное» разнообразие ОУ и внешней среды столь велико, что соотношение (6) не выполняется. Поэтому СУ формирует гомоморфную модель, прибегая к агрегированию, превращению переменных в константы, линеаризации связей и другим методам упрощения модели. Это производится с целью выделения основных параметров («главных» факторов) функционирования системы, на которые направлены управляющие воздействия.

3. Принцип обратной связи

Обратная связь является одним из основных понятий кибернетики, позволяющих понять многие явления, происходящие в системах управления.

Вид соединения элементов, при котором выходное воздействие одного элемента передается на вход другого, называется **прямой связью**.

Вид соединения элементов, при котором выходное воздействие одного элемента передается на вход того же самого элемента, называется **обратной связью**. Различают два вида обратной связи:

- *отрицательная* – если под действием обратной связи первоначальное отклонение выхода системы y , вызванное возмущающими воздействиями w , уменьшается (управление экономическими и биологическими системами);
- *положительная* - если под действием обратной связи первоначальное отклонение выхода системы y , возрастает (системы обучения).

Положительная ОС усиливает действие входного сигнала, а отрицательная – уменьшает. Следовательно, положительная ОС способствует увеличению возникшего в системе отклонения, что приводит к ее неустойчивости. Напротив,



отрицательная ОС способствует восстановлению равновесия в системе, то есть обеспечивает ее устойчивость.

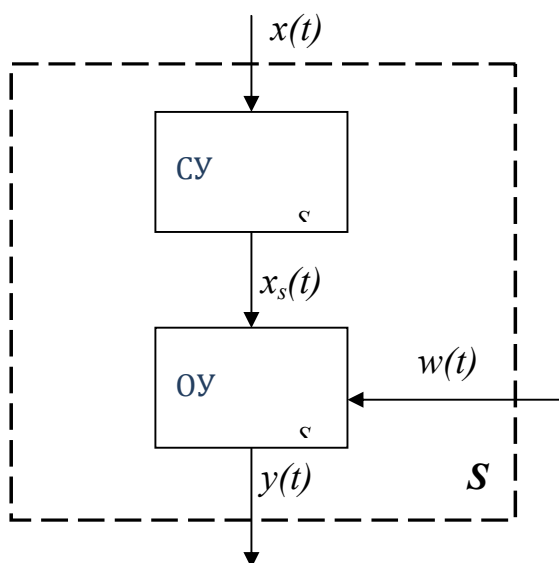
ОС создает возможность эффективного управления в изменяющихся условиях функционирования ОУ, позволяя оценить качество управления в соответствии с выбранным критерием.

3. Виды управления

Свойства систем управления существенно зависят от способа формирования управляющих воздействий, определяющего виды управления.

1. Жесткое управление

Под **жестким управлением** понимается такое воздействие на ОУ, при котором достижение результата не сообщается СУ (обратная связь отсутствует). Для жесткого управления характерно наличие разомкнутого контура управления.



Назначение СУ состоит в следующем:

- в соответствии с некоторой программой на вход СУ поступают задающие воздействия $x(t)$, требующие, чтобы управляемый параметр $y(t)$ принял значение y_0 ;
- СУ генерирует управляющие воздействия $x_s(t)$, по возможности учитывая влияние на ОУ возмущающих воздействий $w(t)$;
- ОУ формирует результаты в виде выходных значений $y(t)$.



В связи с невозможностью предвидения всех возмущений СУ, не имея информации о влиянии их на ОУ, добиться выполнения равенства $y(t)=y_0$ крайне трудно.

Преимущества жесткого управления: относительно простая алгоритмическая и программная реализация.

Недостатки жесткого управления: ограниченность области применения на практике.

Применение: простейшие автоматические устройства, жесткое администрирование.

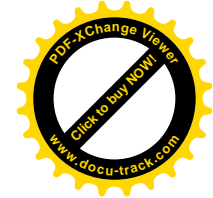
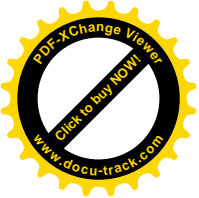
1 Регулирование

Регулирование представляет собой такое воздействие на ОУ, при котором информация о состоянии ОУ поступает в СУ, что позволяет обеспечивать поддержание соответствия значений выходных параметров ОУ заданным (за счет наличия обратной связи). Для регулирования характерно наличие замкнутого контура.

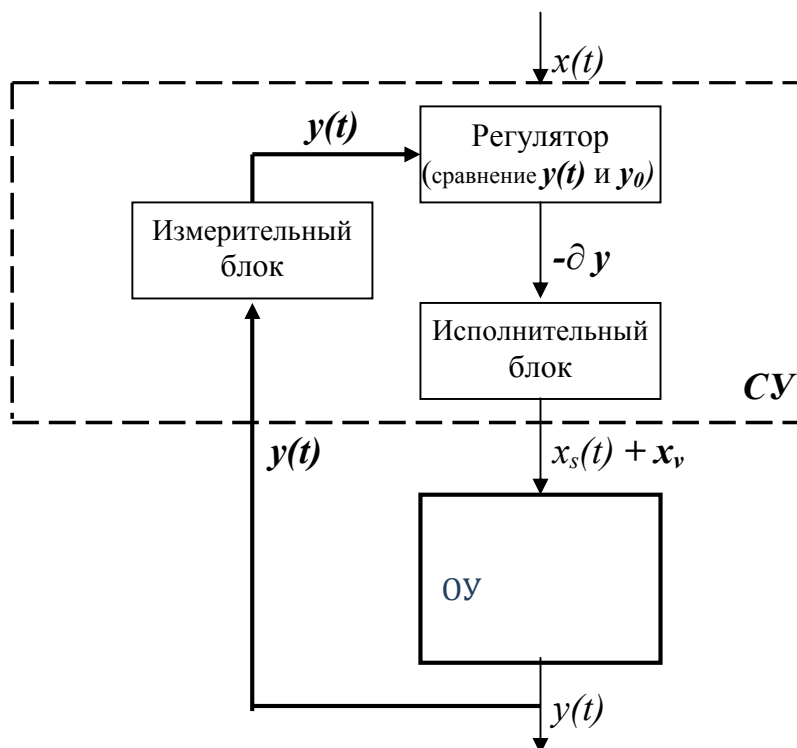
Назначение СУ таково:

- в соответствии с заданными значениями входов и соответствующими им управляющими воздействиями ОУ выдает вектор выходных результатов $y(t)$;
- измерительный блок служит для измерения фактического значения выходной переменной $y(t)$ и ее преобразования в сигнал, удобный для передачи в регулятор СУ;
- регулятор осуществляет сравнение $y(t)$ с заданным значением выхода y_0 и вырабатывает сигнал $\partial y = k[y(t)-y_0]$, который с обратным знаком поступает в исполнительный блок;
- исполнительный блок формирует новые управляющие воздействия, увеличивая или уменьшая их интенсивность $x_s(t) + x_v$.

Если регулирование обеспечивает изменение выхода ОУ в соответствии с некоторой программой, оно именуется *программным регулированием*.



2.1. Регулирование по отклонению



Преимущества регулирования по отклонению: повышение эффективности управления за счет учета состояния ОУ при выработке управляющих воздействий.

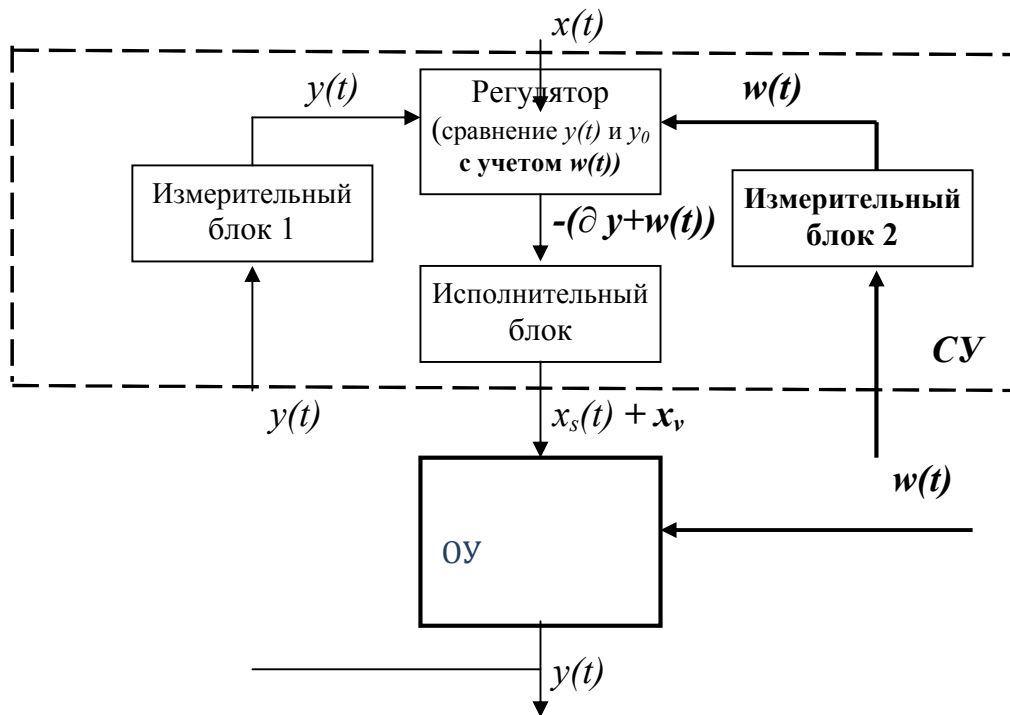
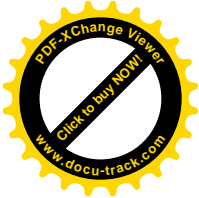
Недостатки регулирования по отклонению:

- более сложная алгоритмическая и программная реализация;
- неполнота учета возмущающих воздействий на ОУ.

Применение: системы автоматической стабилизации в технике (температура, влажность, курс самолета), станки с программным управлением, управление жизнедеятельностью живых организмов (поддержание температуры тела, кровяного давления), управление перевозками грузов.

2.2. Регулирование по возмущению

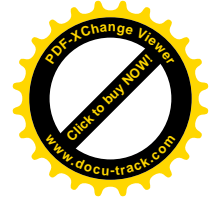
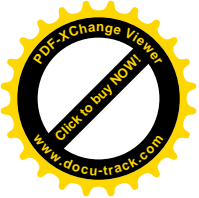
В этом случае программа управления не рассчитывается заранее, а определяется поведением наблюдаемого объекта. СУ учитывает, измеряет *возмущающие воздействия* и компенсирует их регулятором.



Преимущества регулирования по возмущению: повышение эффективности управления не только за счет учета внутреннего состояния ОУ, но и оценки влияния внешней среды.

Недостатки регулирования по возмущению:

еще более сложная алгоритмическая и программная реализация.



Модуль 2 Типовые математические схемы моделирования систем

Тема 2.1. Стохастические модели

ЛЕКЦИЯ 6. КОМПОНЕНТЫ И КЛАССИФИКАЦИЯ МОДЕЛЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ

Система массового обслуживания (СМО) – система, предназначенная для обслуживания потока заявок (требований), поступающих в случайные моменты времени (например, станок, обрабатывающий детали; магазин, обслуживающий покупателей; компьютер, решающий задачи).

В любой СМО предусматривается наличие:

- заявок на обслуживание (деталей, покупателей, задач, заказов), образующих поток;
- каналов обслуживания (приборов), обслуживающих заявки.

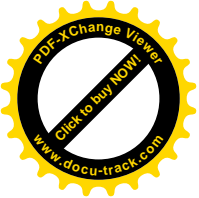
Обслуживание каждой поступившей в случайный момент времени заявки, осуществляется в течение некоторого случайного периода времени, после чего канал освобождается и готов к приему следующей заявки.

Случайный характер потока заявок приводит к тому, что в некоторые промежутки времени на входе СМО скапливается избыточное число заявок (они образуют очередь либо покидают систему необслуженными); в другие же периоды СМО работает с недогрузкой или простаивает.

Целью моделирования СМО является установление зависимости между характером потока заявок, числом каналов, их производительностью, правилами работы СМО и эффективностью обслуживания.

Оптимизация и оценка эффективности СМО состоит в нахождении средних суммарных издержек на сервис каждой заявки и нахождение средних суммарных затрат от заявок не обслуженных.

СМО состоит из определенного числа обслуживающих каналов и предназначена для выполнения заявок с различным характером распределения момента времени на сервис.



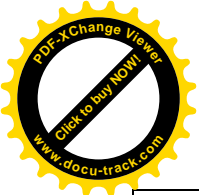
Моделирование СМО предполагает:

- 1) построение ЭММ, связывающих характеристики СМО (число каналов, их производительность и т.П.) с показателями эффективности;
- 2) оптимизацию данных характеристик с целью получения наибольшей эффективности.

Для описания СМО может использоваться обозначение $A/B/m - d$, где A – обозначение закона распределения интервалов времени между заявками; B – обозначение закона распределения времени обслуживания заявок; m – количество каналов; d – обозначение дисциплины обслуживания. В качестве A и B обычно используются следующие обозначения: M – экспоненциальное распределение, G – любое другое. Для некоторых распределений используются специальные обозначения, например, D – детерминированная величина, E_k – распределение Эрланга k -го порядка, и т.д. Имеется несколько вариантов классификации СМО (табл.1).

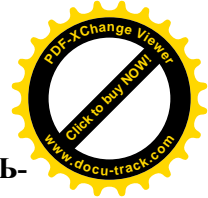
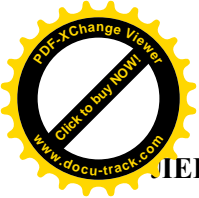
Таблица 1 – Классификация СМО

Признак классификации	Типы СМО	Описание
1	2	3
Законы распределения интервалов времени	Марковские	Интервалы времени между моментами поступления заявок и времена обслуживания заявок распределены по экспоненциальному закону
	Немарковские	Закон распределения интервалов времени между моментами поступления заявок и/или времен обслуживания заявок отличается от экспоненциального
Количество каналов	Одноканальные	В любой момент времени в СМО может обслуживаться не более одной заявки
	Многоканальные	В СМО одновременно может обслуживаться несколько заявок
Ограничения на очередь	Без ограничений на очередь	Если заявка поступает на обслуживание в момент, когда все каналы заняты, то она становится в очередь. Ограничений на количество заявок в очереди и на время пребывания в ней нет.
	С ограничениями на очередь	Имеются ограничения на количество заявок в очереди и/или время их пребывания в очереди. Если заявка поступает на обслуживание в момент, когда в очереди уже имеется предельное количество заявок, или время пребывания заявки в очереди превышает допустимое, то она покидает СМО (не обслуживается)



	Без очереди	Если заявка поступает на обслуживание в момент, когда все каналы заняты, то она покидает СМО (не обслуживается). Очередь в такой СМО не образуется
Количество заявок	Замкнутые	Имеется фиксированное количество заявок, периодически требующих обслуживания.
	Разомкнутые	Количество заявок, которые могут поступать в СМО, не ограничено
	Смешанные	Имеется фиксированное количество заявок, периодически требующих обслуживания. При этом возможно поступление дополнительных заявок
Количество узлов СМО и связь между ними	Однофазные (одиночные)	Один типовой узел
	Многофазные	Последовательность типовых узлов. Все заявки, обслуженные в одном узле, направляются в следующий узел
	Сеть СМО	СМО, состоящая из нескольких типовых узлов. Количество узлов, в которых требуется обслуживание, и порядок их прохождения могут быть различными для разных заявок
Дисциплина обслуживания (порядок обслуживания заявок из очереди)	FIFO	«Первым пришел – первым обслужен» (обслуживание в порядке поступления)
	LIFO	«Первым пришел – последним обслужен»
	С относительными приоритетами	Первыми из очереди выбираются заявки с более высоким приоритетом. Если обслуживание заявки началось, то оно всегда доводится до конца, даже если в это время поступает заявка с более высоким приоритетом
	С абсолютными приоритетами	Первыми из очереди выбираются заявки с более высоким приоритетом. Обслуживание заявки прерывается, если поступает заявка с более высоким приоритетом
	Квантованное обслуживание	На обслуживание каждой заявки выделяется определенное время. Если за это время обслуживание не завершается, то заявка возвращается в очередь, и обслуживается следующая заявка
	По необходимому времени обслуживания	Первыми обслуживаются заявки, для которых требуется меньше времени

Тема 2.2. Детерминированные модели



ЛЕКЦИЯ 7. СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ЭКСПЕРИМЕНТА

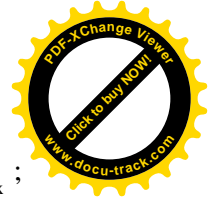
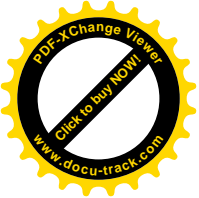
Эффективность машинных экспериментов с имитационными моделями системы S , формализуемых в виде схем массового обслуживания, существенно зависит от выбора плана эксперимента, так как именно план определяет объем и порядок проведения вычислений на ПЭВМ, приемы накопления и статистической обработки результатов моделирования системы и в целом влияет на эффективность использования ресурсов ПЭВМ при моделировании.

Математические методы планирования экспериментов основаны на кибернетическом представлении процесса проведения эксперимента, наиболее подходящей моделью которого является абстрактная схема типа «черного ящика» вида $Y = \varphi(X)$, где $Y = (y_1, y_2, \dots, y_k)$ – множество векторов зависимых выходных переменных, называемых реакциями или откликами (для машинного эксперимента эти переменные являются эндогенными); $X = (x_1, x_2, \dots, x_m)$ – множество векторов входных независимых переменных, называемых факторами (для машинного эксперимента они являются экзогенными).

Функция φ , связывающая отклик с факторами, называется функцией реакции.

При проведении машинного эксперимента с моделью для оценки характеристик процесса функционирования исследуемой системы необходимо создать такие условия, которые способствовали бы выявлению влияния факторов, находящихся в функциональной связи с искомой характеристикой. Для этого необходимо:

- 1) отобрать факторы $x_i, i = \overline{1, m}$, влияющие на искомую характеристику, и описать функциональную зависимость;



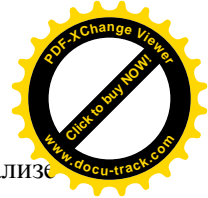
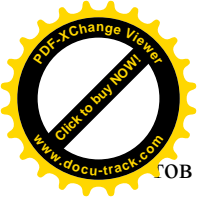
- 2) установить диапазон изменения факторов $x_{i_{\min}} \dots x_{i_{\max}}$;
- 3) определить координаты точек факторного пространства $\{x_1, x_2, \dots, x_m\}$, в котором следует проводить эксперимент;
- 4) определить необходимое число реализаций и их порядок в эксперименте.

Свойства объекта исследования, т.е. процесса машинного моделирования системы S на многофакторное возмущение, – одна из задач математического планирования эксперимента. Наиболее распространенными и полно отвечающими задачам статистического моделирования являются полиномиальные модели. Задача нахождения полиномиальной модели, описывающей систему или отдельные ее характеристики, состоит в оценке вида и параметров некоторой функции от переменных (x_1, x_2, \dots, x_m) .

Рассмотрим влияние m количественных факторов $x_i, i = \overline{1, m}$ на некоторую реакцию (отклик) η в отведенной для экспериментирования локальной области $[x_{i_{\min}} \dots x_{i_{\max}}, i = \overline{1, m}]$ факторного пространства G . Функцию реакции $\varphi(x_1, x_2, \dots, x_m)$ представим в виде полинома степени d от m переменных, который содержит C_{m+d}^d коэффициентов:

$$\begin{aligned} \eta \equiv \varphi = & \tilde{b}_0 + \sum_{1 \leq i \leq m} \tilde{b}_i x_i + \sum_{i, j} \tilde{b}_{ij} x_i x_j + \\ & + \sum_{i_1, i_2, \dots, i_m} \tilde{b}_{i_1, i_2, \dots, i_m} x_1^{i_1} x_2^{i_2} \dots x_m^{i_m}, \end{aligned} \quad (5)$$
$$\sum ij = d$$

Для нахождения значений коэффициентов данного полинома используют регрессионный анализ (метод наименьших квадратов). Для него необходимо получить $N \geq C_{m+d}^d$ результа-

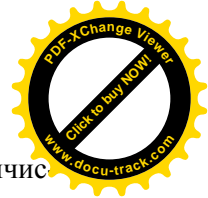
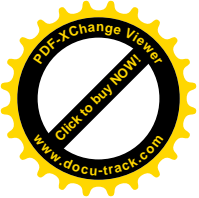


гов наблюдения величины отклика. В регрессионном анализе полагается, что выполняется ряд предпосылок.

1) Результаты наблюдений y_1, y_2, \dots, y_N — независимые, нормально распределенные случайные величины. Речь идет о распределении y относительно некоторой фиксированной точки (x_1, x_2, \dots, x_m) , так как на значение y влияют и другие неконтролируемые параметры. Если эта предпосылка не удовлетворяется, то коэффициенты регрессии найти можно, однако ничего нельзя будет сказать об эффективности метода, т.е. нельзя оценить точность уравнения регрессии. Если y не подчиняется нормальному распределению, то стараются подобрать такую функцию преобразования, чтобы перейти от y новой случайной величине $q = f(y)$, приближенной распределенной по нормальному закону. Например, для многих ассиметричных распределений делается замена $q = \ln(y)$.

2) Дисперсии $\sigma_y^2 = \sigma^2\{y_u\}$, $u = \overline{1, N}$ равны друг другу. Это означает, что если производить многократные и повторные наблюдения над величиной y_u при некотором определенном наборе значений $(x_{1u}, x_{2u}, \dots, x_{ku})$, то получим дисперсию σ_y^2 , которая не будет зависеть от математического ожидания Mu , т.е. не будет отличаться от σ_y^2 , полученной при повторных наблюдениях для любого другого набора независимых переменных. Это требование также не всегда выполняется для реального эксперимента.

3) Независимые переменные x_1, x_2, \dots, x_m измеряются с пренебрежимо малой ошибкой по сравнению с ошибкой в определении y .



При таких предпосылках оказывается возможным вычислить коэффициенты полинома, а также оценить их точность и точность уравнения регрессии в целом.

Проведение эксперимента над моделью любой сложной системы завершается получением больших массивов данных. В состав данных массивов входят наборы исходных данных, значения промежуточных и результирующих параметров. Практически для любой системы требуется определить степени влияния исходных параметров на результат, меру линейной зависимости, и получить иные статистические показатели функционирования системы.

В данном разделе приводятся методы получения статистической информации работы системы путем использования соответствующей имитационной модели на основе статистических исследований. Для обработки результатов эксперимента используются линейные многофакторные модели.

С помощью грамотно построенной многофакторной модели можно определить, как себя ведет отклик рассматриваемой сложной системы в зависимости от изменения факторов. Построение линейной многофакторной модели осуществляется с помощью методов корреляционно-регрессионного анализа.

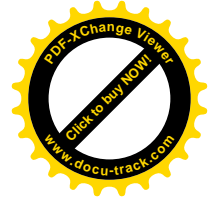
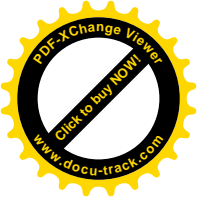
7.1. Обозначения

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \text{значения отклика (зависимой переменной) в пе-}$$

риодах времени, где n – число периодов наблюдения.

$X = (x^1, x^2, \dots, x^m)$ – ряды m факторов (независимых или объясняющих переменных).

$\overline{x^i}$ – среднее значение i - й переменной.



7.2. Предварительные понятия и определения

Дисперсия случайной величины x определяется по формуле:

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}.$$

Ее называют еще среднеквадратическим отклонением.

Определение 7.1. Величину $\sigma_x = \sqrt{\sigma_x^2}$ называют стандартным отклонением величины x .

Коэффициент корреляции между случайными величинами x и y :

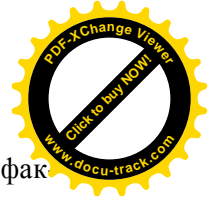
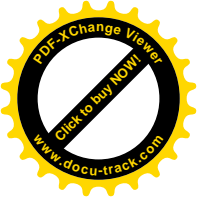
$$r_{xy} = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Коэффициент корреляции характеризует тесноту или силу связи между переменными y и x . Значения принимаемые r_{xy} находятся на отрезке $[-1, 1]$. Если $r_{xy} > 0$, то корреляция положительна, т.е. с увеличением (уменьшением) одной переменной (x) увеличивается (уменьшается) другая переменная (y). Если корреляция отрицательна, то наблюдается обратная ситуация. При нулевой корреляции связи между переменными нет.

Если речь идет о многофакторной модели, то вычисляется

коэффициент множественной корреляции $R = \sqrt{1 - \frac{\sigma_e^2}{\sigma_a^2}}$, где

σ_e^2 – остаточная дисперсия отклика, $\sigma_a^2 = \sigma_y^2$ – общая дисперсия отклика.



Общая дисперсия характеризует разброс наблюдений фактических значений от среднего \bar{y} .

Остаточная дисперсия:

$$\sigma_e^2 = \frac{\sum_{i=1}^n (y_i - y_{it})^2}{n - m - 1},$$

где y_{it} – вычисленные значения переменной y посредством уравнения регрессии при подстановке в него наблюдаемых фактических значений x^i . Остаточная дисперсия характеризует ту часть рассеяния переменной y , которая возникает из-за всякого рода случайностей и влияния неучтенных факторов.

Коэффициент детерминации

$$D = 1 - \frac{\sigma_e^2}{\sigma_a^2}.$$

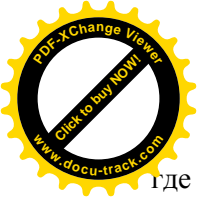
Данный коэффициент служит для оценки точности регрессии, т.е. соответствия полученного уравнения регрессии имеющимся эмпирическим данным. Значение коэффициента изменяется в пределах $[0, 1]$. Чем D ближе к единице, тем точнее регрессионная модель. Стандартная ошибка оценки равна $\sqrt{\sigma_e^2}$.

Можно показать, что при $D = 0$ корреляционно-регрессионная связь между переменными x и y отсутствует. Если $D = 1$, то связь функциональная.

7.3. Оценка параметров линейной многофакторной модели

Общий вид линейной многофакторной модели следующий:

$$\hat{y} = \sum_{j=0}^m a^j x^j, \quad (6)$$



где $x^0 = 1$ – дополнительная константа для учета в уравнение регрессии свободного коэффициента (в этом случае в матрицу X добавляется единичный вектор-столбец и единичная строка):

$$\hat{y} = (y_{1t}, \dots, y_{nt}).$$

В матричной форме уравнение (6) примет вид

$$\hat{y} = a^T x.$$

Для оценки параметров регрессионной модели необходимо минимизировать следующую функцию:

$$S(a) = (y - \hat{y})^2 = (y - Xa)^T (y - Xa) \rightarrow \min,$$

т.е. минимизировать сумму квадратов остатков. Такой метод, называемый методом наименьших квадратов Вам известен.

Дифференцируя $S(a)$ по компонентам вектора a , получаем следующее соотношение:

$$\frac{\partial S(a)}{\partial a} = -2X^T y + 2X^T Xb = 0$$

или

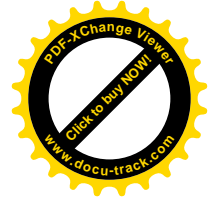
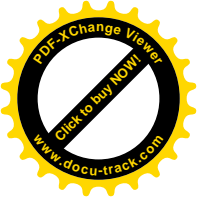
$$X^T Xa = X^T y. \quad (7)$$

Соотношение (7) является системой линейных уравнений, решением которой при условии обратимости матрицы $X^T X$ является

$$a = (X^T X)^{-1} X^T y.$$

Решение системы (7) можно осуществить соответствующими методами, например, методом Жордана-Гаусса. Здесь

$$X^T X = \begin{pmatrix} n & \sum x_i^1 & \dots & \sum x_i^m \\ \sum x_i^1 & \sum (x_i^1)^2 & \dots & \sum x_i^1 x_i^m \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^m & \sum x_i^m x_i^1 & \dots & \sum (x_i^m)^2 \end{pmatrix};$$



$$X^T y = \begin{pmatrix} \sum y_i \\ \sum x_i^1 y_i \\ \vdots \\ \sum x_i^m y_i \end{pmatrix}.$$

Когда нам известна методика оценки параметров линейной многофакторной модели, мы можем приступить к ее построению.

7.4. Построение линейной многофакторной регрессионной модели

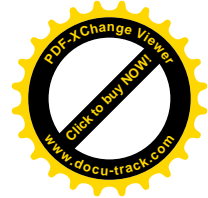
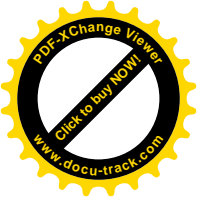
Построение многофакторной модели содержит не только этапы вычислений регрессионных коэффициентов, но и оценку правомерности ее использования по различным критериям. Также необходима проверка обсужденных ранее предпосылок.

Построение модели выполняется по следующим этапам:

- 1) априорное исследование системы;
- 2) формирование перечня факторов и их логический анализ;
- 3) сбор исходных данных и их первичная обработка;
- 4) спецификация функции регрессии;
- 5) оценка функции регрессии;
- 6) отбор главных факторов;
- 7) проверка адекватности модели;
- 8) интерпретация модели;
- 9) прогнозирование неизвестных значений зависимой переменной.

Последний этап может не выполняться, если только требуется оценить влияние выбранных факторов на отклик системы.

Рассмотрим подробнее содержание этапов.



7.4.1. Априорное исследование системы

В соответствии с целью работы на основе знаний в области рассматриваемой системы конкретизируются явления, процессы, зависимость между которыми подлежит оценке. При этом подразумевается, прежде всего, четкое определение явлений в системе, установление объектов и периода исследования.

На этом этапе должны быть сформулированы осмысленные и приемлемые гипотезы о зависимости явлений, например, зависимость производительности труда от автоматизации производственного процесса, уровня квалификации рабочих и размеров вознаграждений за труд.

7.4.2. Формирование перечня факторов и их логический анализ

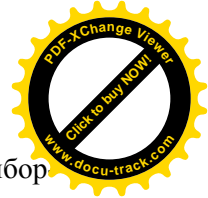
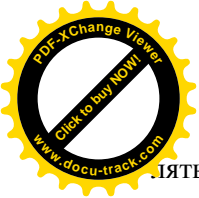
Для определения разумного числа переменных в регрессионной модели, необходимо, прежде всего, ориентироваться на соображения профессионально-теоретического характера. Исходя из физического смысла явления, производят классификацию переменных на зависимые и объясняющие.

7.4.3. Сбор исходных данных и их первичная обработка

Исходная информация может быть представлена в трех формах:

- 1) динамические (временные) ряды;
- 2) пространственная информация – информация о работе нескольких объектов в одном разрезе времени;
- 3) сменная (табличная форма) – информация о работе нескольких объектов за разные периоды.

Объем выборки (число значений переменных) не может быть произвольным и зависит от количества факторов, входящих в состав модели с учетом свободного члена. Рекомендуется для статистически значимой модели на один фактор осуществ-



для выборки в объеме 5 – 8 наблюдений. Общий объем выборки рассчитывается исходя из количества факторов, а также количества свободных членов уравнения регрессии.

7.4.4. Спецификация функции регрессии

На данном этапе исследования дается конкретная формулировка гипотезы о форме связи (линейная или нелинейная, простая или множественная и т.д.). Для этого используются различные критерии для проверки состоятельности гипотетического вида зависимости. Также проверяются предпосылки корреляционно-регрессионного анализа.

7.4.5. Оценка функции регрессии

Здесь определяются числовые значения параметров регрессии, и вычисляется ряд показателей, характеризующих точность модели.

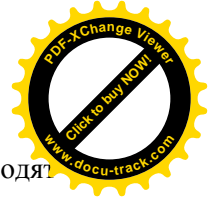
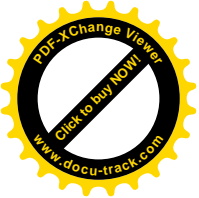
7.4.6. Отбор главных факторов

Одним из важных этапов составления многофакторной модели является отбор объясняющих переменных.

В первую очередь на данном этапе формируют перечень тех факторов, которые могут влиять на отклик системы. Обычно их число достигает нескольких десятков. Поэтому построенная с таким количеством факторов модель может быть неустойчивой. Она проявляется, когда при изменении некоторых факторов значение зависимой переменной будет расти вместо того, чтобы уменьшаться.

При небольшом количестве факторов полученная модель может приводить к ошибкам при принятии решений в ходе ее анализа.

В связи с этим определяют рациональный перечень факторов.



Отбор факторов из предварительного списка производят с помощью их анализ на мультиколлинеарность.

Определение 7.2. Мультиколлинеарность – попарная корреляционная зависимость между факторами.

Считается, что мультиколлинеарная зависимость между факторами x^i и x^j присутствует, если коэффициент парной корреляции $r_{ij} \geq 0,7 \div 0,8$.

Отрицательное воздействие мультиколлинеарности состоит в следующем:

- 1) усложняется процедура выбора главных факторов;
- 2) искажается смысл коэффициента множественной корреляции (он предполагает независимость факторов);
- 3) усложняются вычисления при построении самой модели;
- 4) снижается точность оценки параметров регрессии, искажается оценка дисперсии.

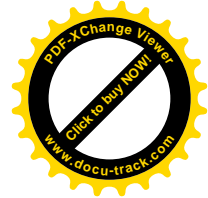
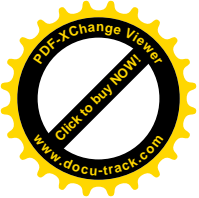
Следствием снижения точности является ненадежность регрессионных коэффициентов и отчасти неприемлемость их использования для интерпретации меры воздействия соответствующей объясняющей переменной на зависимую переменную.

Оценки коэффициента становятся очень чувствительными к выборочным наблюдениям. Здесь становится ненадежным применение критериев значимости. Таким образом, становится ясным, что при построении рассматриваемой модели необходимо установить стохастическую мультиколлинеарность, которую по возможности следует устранить.

Для измерения мультиколлинеарности можно использовать квадрат коэффициента множественной корреляции:

$$D = R^2.$$

Значение D является коэффициентом детерминации. Введем коэффициент парной детерминации $d_{yj} = r_{yj}^2$. Тогда, при отсутствии мультиколлинеарности



$$D = \sum_{j=1}^m d_{yj} . \quad (8)$$

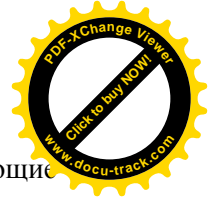
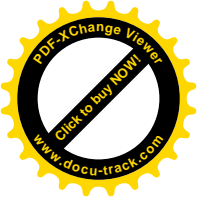
Соотношение (8) показывает, что в многофакторной модели присутствуют связи только между зависимой переменной y и факторами $x^j, j = \overline{1, m}$. Связи между факторами отсутствуют. Равенство (8) может выполняться в редких случаях. Поэтому вводят меру мультиколлинеарности:

$$M = D - \sum_{j=1}^m d_{yj} .$$

Чем меньше эта разность, тем меньше мультиколлинеарность. Для ее снижения используется метод исключения переменных. Суть метода определяется исключением тех факторов, которые хорошо коррелированы. При этом порог модуля коэффициента корреляции $r_{ij}, i = \overline{1, m}; j = \overline{1, m}$ задается значением больше 0,6. Как правило, если $|r_{ij}| \geq 0,7, i, j = \overline{1, m}$, то один из факторов исключают. Здесь сразу возникает вопрос: «Какой из двух хорошо коррелированных факторов исключать?» Решение об исключении того или иного фактора принимают, исходя из управляемости факторов на уровне системы.

Если фактор хорошо управляем и позволяет улучшить функционирование системы, то его оставляют в модели. Но часто возникают ситуации, когда оба фактора положительно влияют на систему, и на них можно воздействовать. Поэтому вопрос исключения факторов решают путем процедуры отбора главных факторов.

Принятие решения об отборе главных факторов производится с применением анализа значений специальных статистических характеристик.



Процедуры отбора главных факторов включает следующие этапы:

- 1) анализ факторов на мультиколлинеарность и ее снижение (исключение);
- 2) анализ тесноты взаимосвязи факторов $x^j, j = \overline{1, m}$ с зависимой переменной y ;
- 3) анализ степени влияния факторов на зависимую переменную;
- 4) проверка коэффициента регрессии $a_j, j = \overline{0, m}$ на статистическую значимость;
- 5) анализ факторов на управляемость;
- 6) построение новой регрессионной модели без исключенных факторов.
- 7) исследование целесообразности исключения факторов из модели с помощью коэффициента детерминации.

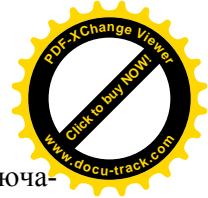
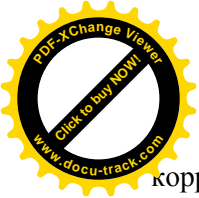
Для выполнения первых двух этапов составляют корреляционную матрицу следующего вида (таблица 2).

Таблица 2. Корреляционная матрица

№ переменной	x^1	x^2	...	x^m	y
x^1	1	r_{12}	...	r_{1m}	r_{1y}
x^2	r_{21}	1	...	r_{2m}	r_{2y}
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
x^m	r_{m2}	r_{m2}	...	1	r_{my}

Для анализа факторов на мультиколлинеарность используют коэффициенты парной корреляции $r_{ij}, i, j = \overline{1, m}$.

Анализ тесноты взаимосвязи $x^j, j = \overline{1, m}$ и y проводится с использованием соответствующих коэффициентов парной



корреляции r_{jy} , $j = \overline{1, m}$. Если $r_{jy} = 0$, то j -й фактор исключается автоматически. Далее факторы сортируются в порядке возрастания модулей коэффициентов парной корреляции. Принятие решения об их исключении начинается с первых элементов полученной последовательности. Окончательное решение об исключении принимается в ходе анализа других статистических характеристик.

Анализ степени влияния факторов на зависимую переменную осуществляется путем вычисления коэффициента β_k .

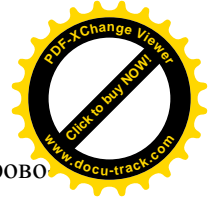
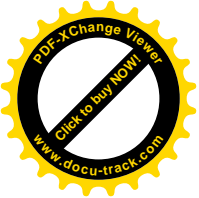
$$\beta_k = a_k \frac{\sigma_{x^k}}{\sigma_y}, \text{ где } k = \overline{1, m}.$$

Коэффициент показывает, на сколько σ_y изменится зависимая переменная y , если изменить на σ_{x^k} k -й фактор при фиксированных значениях остальных факторов. Он учитывает влияние анализируемого фактора на зависимую переменную «с учетом различий уровня его колеблемости». Фактор x^j с меньшим коэффициентом β_j по отношению к фактору x^i (коэффициент β_i) может быть исключен из модели.

Принятие решения об исключении фактора из модели путем вычисления коэффициента β_k имеет меньший вес по отношению к принятию решения на базе оценки тесноты связи фактора с зависимой переменной. Т.е. если для факторов x^i и x^j установлено, что $r_{iy} > r_{jy}$, но $\beta_i < \beta_j$, то исключению подлежит фактор x^j .

Проверка коэффициентов на статистическую значимость проводится по одному из двух критериев:

- 1) критерий Стьюдента;
- 2) критерий Фишера.



Проверка коэффициентов по критерию Стьюдента проводится по следующей формуле:

$$t_k = \frac{a_k}{S_{a_k}}, k = \overline{1, m},$$

где a_k – коэффициент регрессии при k -м факторе,

S_{a_k} – стандартное отклонение оценки параметра a_k , вычисляемое по формуле:

$$S_{a_k} = \frac{\sigma_e^2}{\sigma_k(1 - d_{yy})(n-1)}.$$

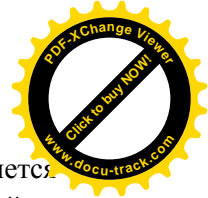
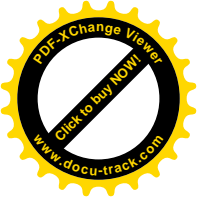
Здесь число степеней свободы статистики Стьюдента равно $f = n - m - 1$.

Полученные значения $t_k, k = \overline{1, m}$ сравниваются с критическим значением $t_{f, \alpha}$, где α – уровень значимости. Если $t_k \geq t_{f, \alpha}$, то очевидно a_k существенно больше нуля, а фактор x^k оказывает большее влияние на зависимую переменную y . В противном случае фактор исключается из модели.

Проверка статистической значимости по критерию Фишера проводится с помощью формулы:

$$F_k = \left(\frac{a_k}{S_{a_k}} \right)^2, k = \overline{1, m}.$$

Здесь число степеней свободы статистики F_k следующие $f_1 = 1, f_2 = n - m - 1$. Как в предыдущей методике, значения F_k сравниваем с критическим значением $F_{f_1 f_2, \alpha}$. Если $F_k > F_{f_1 f_2, \alpha}$, то фактор x^k оказывает существенное влияние на y . В противном случае фактор x^k исключается.



При анализе факторов на управляемость осуществляется логический анализ факторов на предмет возможности воздействия на них на уровне системы. Те факторы, которые не поддаются управлению, исключаются из модели.

Например, мы исследуем зависимость объема выпускаемой продукции от следующих факторов: размер вознаграждения рабочим, скорость резки деталей, которую выполняет станок, количество технических перерывов за одну смену. Очевидно, что мы не можем воздействовать на второй фактор, т.к. он определяется техническими характеристиками оборудования. Поэтому, при близких значениях коэффициентов корреляции и β_k , мы его исключаем.

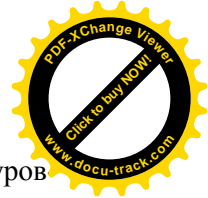
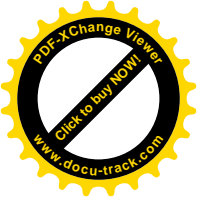
Построение новой регрессионной модели.

Пусть мы исключили из полученной ранее модели с коэффициентом детерминации D_m m_2 факторов. Тогда строим новую регрессионную модель с оставшимися $m_1 = m - m_2$ факторами. Для нее также вычисляем коэффициент детерминации D_{m_1} .

Исследование целесообразности исключения m_2 факторов из модели с помощью коэффициента детерминации осуществляется путем расчета следующей величины:

$$F_1 = \frac{(D_m - D_{m_1})(n - m - 1)}{(m - m_1)(1 - D_m)}. \quad (9)$$

С помощью данной величины, которая имеет F - распределение со степенями свободы $f_1 = m - m_1 = m_2$ и $f_2 = n - m - 1$, устанавливается, вносят ли совместно исключенные факторы существенную долю в вариацию объясняющей переменной y . Разность $D_m - D_{m_1}$ в (9) является мерой дополнительного объяснения вариации y за счет обратного включения выведенных из модели факторов.



Значение F_1 сравнивают с критическим F_{α, f_1, f_2} для уровня значимости α и степеней свободы f_1, f_2 . Если $F_1 > F_{\alpha, f_1, f_2}$, то m_2 объясняющих переменных совместно оказывают существенное влияние на вариацию переменной y , и, следовательно, все m_2 переменные из модели нельзя исключать. В противном случае, m_2 переменных совместно не оказывают существенного влияния на вариацию переменной y . Здесь эти переменные окончательно исключаются из модели.

7.4.7. Проверка адекватности модели

Данный этап анализа включает:

- 1) оценку значимости коэффициента детерминации;
- 2) проверку качества подбора теоретического уравнения;
- 3) вычисление специальных показателей.

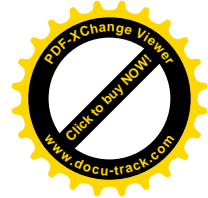
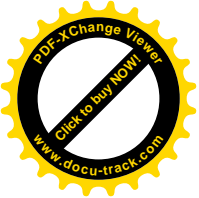
Оценка значимости коэффициента детерминации необходима для решения вопроса: оказывают ли выбранные факторы влияние на зависимую переменную? Оценка позволяет выявить такие ситуации, когда величина D целиком обусловлена случайными колебаниями в выборке, на основании которой он вычислен.

Для оценки значимости D используется F -статистика со степенями свободы $f_1 = m$, $f_2 = n - m - 1$:

$$F = \frac{D(n - m - 1)}{m(1 - D)}.$$

Если значение F превышает критическое, то включенные в уравнение регрессии переменные достаточно объясняют зависимую переменную y .

Проверка качества подбора теоретического уравнения осуществляется с использованием средней ошибки аппроксимации.



$$E = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - y_{it}}{y_{it}} \right) 100\% .$$

Специальные показатели применяются для характеристики воздействия отдельных факторов на итоговый показатель.

Коэффициент эластичности:

$$\beta_k^e = a_k \frac{\overline{x_k}}{\overline{y}} .$$

Он показывает, насколько процентов в среднем изменяется функция с изменением аргумента на 1% при фиксированных значениях других аргументов.

Мера вариации результативного признака за счет изолированного влияния фактора x^k :

$$g_k = \beta_k^2 .$$

Показатель системного эффекта факторов показывает общее влияние факторов без учета внутренних связей:

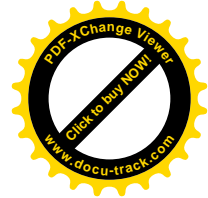
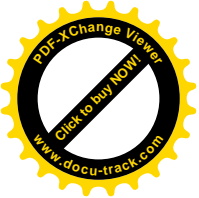
$$\eta_s = R^2 - \sum_{j=1}^m \beta_j^2 .$$

7.4.8. Интерпретация системы

Результаты регрессионного анализа сравниваются с гипотезами, сформулированными на первом этапе исследования, и оценивается их правдоподобие с точки зрения поведения системы.

7.4.9. Прогнозирование неизвестных значений зависимой переменной

Полученное уравнение регрессии находит практическое применение в прогностическом анализе. Прогноз получают путем подстановки в регрессию численно оцененных значений факторов.



ЛЕКЦИЯ 8. ПРОГНОЗИРОВАНИЕ ПАРАМЕТРОВ СИСТЕМЫ

Прогноз осуществляется для оценки неизвестного параметра в будущем периоде времени или в неисследованной части пространства.

При прогнозировании используют два понятия:

1) **Предсказание** – суждение о будущем состоянии процесса (или состоянии процесса в неисследованной части пространства), основанное на субъективных оценках качественных и количественных факторов.

2) **Прогнозирование** – оценка параметра в будущем периоде или в неисследованной области пространства на базе эвристических и математических методов.

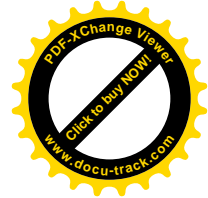
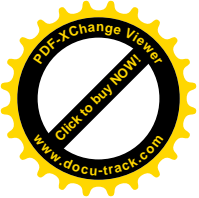
Эвристические методы базируются на использовании трудно формализуемых или не формализуемых явлений или процессов. В математических методах осуществляется подбор и обоснование соответствующей модели исследуемого процесса, а также способов определения неизвестных параметров. Здесь задача прогнозирования сводится к решению уравнений, описывающих данную модель для заданного момента времени.

Для экономических систем наиболее популярны методы экстраполяции, которые отличаются простотой, наглядностью и легко реализуются на ЭВМ. Экстраполяционные методы предполагают, что существует связь между прошлым, настоящим и будущим. Когда речь идет о временных рядах, то существенен порядок, в котором осуществляются наблюдения.

Временной ряд y_t , $t = \overline{1, n}$ может быть представлен в следующем виде:

$$y_t = x_t + \varepsilon_t,$$

где x_t – детерминированная компонента процесса,



ε_t – стохастическая (случайная) компонента процесса.

Детерминированная компонента, иначе ее называют трендом, x_t характеризует общую тенденцию изменения изучаемого показателя. Стохастическая компонента ε_t отражает случайные колебания или шумы.

Отрезок времени k , отмеряемый от момента времени t , т.е. точки $n+1, n+2, \dots, n+k$, называют периодом упреждения или прогнозным периодом. Если период упреждения находится в пределах трех лет, то прогноз называют краткосрочным. Когда мы прогнозируем на период от трех до пяти лет, то выполняем среднесрочный прогноз, а при периоде более пяти лет – долгосрочный прогноз.

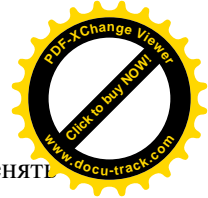
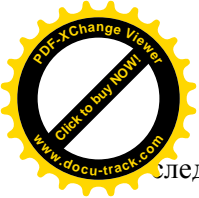
При прогнозировании оценивается математическое ожидание исследуемого процесса (точечный прогноз) и величина интервала, в который с заданной вероятностью попадет прогнозируемое значение процесса (интервальный прогноз).

Экстраполяционные методы можно классифицировать на следующие группы:

- 1) методы авторегрессии;
- 2) методы, основанные на разложении временного ряда на компоненты;
- 3) методы, позволяющие учесть неравнозначность исходных данных;
- 4) методы прямой экстраполяции;
- 5) методы, основанные на построении многофакторных корреляционно-регрессионных моделей.

Дадим краткую характеристику вышеприведенным группам.

Авторегрессионные модели используют в случае, когда влияние факторов на исследуемый параметр четко не определено, т.е. невозможно выделить стабильные во времени причинно-



следственные связи. Такие модели целесообразно применять для сильно автокоррелированных динамических рядов.

При построении моделей авторегрессии предполагается, что во временных рядах существует связь между недавно реализованными значениями и значением, которое наступит в будущем, т.е. имеется корреляция. Модель авторегрессии имеет следующий вид:

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_m y_{t-m},$$

где $a_l, l = \overline{0, m}$ – параметры уравнения регрессии, y_t – значение показателя в прогнозном периоде ($t = \overline{n+1, n+k}$).

При составлении данного уравнения часто осуществляют элиминирование мультиколлинеарности в матрице приведенных рядов данных ($n - m > 2$):

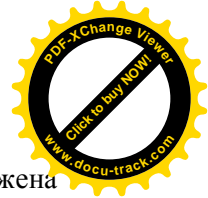
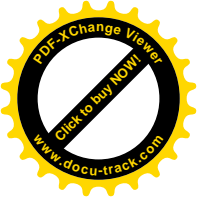
$$\begin{pmatrix} y_{m+1} & y_{m+2} & \cdots & y_n \\ y_m & y_{m+1} & \cdots & y_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \cdots & y_{n-m} \end{pmatrix}.$$

Для улучшения свойств прогнозной модели в нее вводят фактор времени в виде самостоятельной переменной t . Увеличение точности прогнозирования в ряде случаев достигается за счет учета линейного тренда.

Методы прогнозирования, основанные на разложении временного ряда на компоненты – главная тенденция, сезонные колебания и случайная составляющая, – позволяют описать многие (практически любые) экономические процессы независимо от их характера. При аддитивной связи компонент модель имеет вид:

$$y_t = y_t' + y_t'',$$

где y_t' – трендовая составляющая, y_t'' – сезонные колебания и случайная составляющая.



При этом последняя ($y_t'' = y_t - y_t'$) может быть разложена в ряд Фурье:

$$y_t'' = a_0 + \sum_{i=1}^m (a_i \cos it_r + b_i \sin it_r) + \xi_t.$$

Смысл переменной t_r и правила вычисления коэффициентов $a_0, a_i, b_i, i = \overline{1, m}$ приведены в работе [3].

Качество прогнозирования напрямую зависит от качества и полноты исходных данных. Вышеприведенные методы могут давать сбои при неравнозначности исходных данных (например, разные размерности и порядки значений факторов).

Преодолеть данную проблему можно с помощью соответствующих методов:

- 1) метод авторегрессии с последующей адаптацией коэффициентов уравнения;
- 2) метод взвешенных отклонений.

Мы рассмотрим один из подходов адаптации коэффициентов.

Для адаптации коэффициентов модели авторегрессии может быть использован принцип метода градиентного (наискорейшего) спуска:

$$a_n = a_o - k - \text{grad}(l_{t+\tau}^2), \quad (10)$$

где a_n – вектор новых коэффициентов,

a_o – вектор старых коэффициентов,

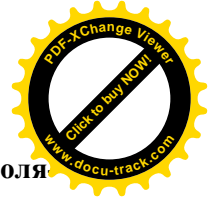
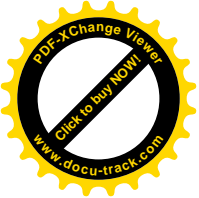
$k > 0$ – управляемый коэффициент,

$l_{t+\tau}$ – ошибка прогноза в момент времени $t + \tau$.

После математических преобразований получим откорректированные оценки коэффициентов:

$$a_n = a_o + 2kl_{t+\tau}x_t.$$

Для нахождения коэффициента k можно использовать итеративную процедуру, описанную в работе [4].



8.1. Прогнозирование с помощью методов экстраполяции

Прогнозирование с помощью методов экстраполяции включает в себя следующие этапы:

- 1) определение цели и задачи исследования, анализ прогнозируемого объекта;
- 2) подготовка исходных данных;
- 3) фильтрация исходного временного ряда;
- 4) логический отбор видов аппроксимирующей (прогнозной) функции;
- 5) оценка математической модели прогнозирования (вычисление коэффициентов прогнозной функции);
- 6) выбор математической модели прогнозирования.

8.1.1. Определение цели и задачи исследования, анализ прогнозируемого объекта

Данный этап предусматривает детальное логическое изучение системы и формулирование взаимодействий ее компонент.

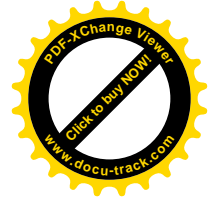
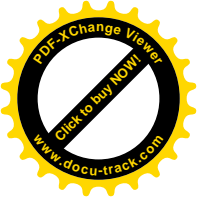
8.1.2. Подготовка исходных данных

На данном этапе устанавливается полнота исходных рядов данных. Недостающие значения вычисляются с помощью методов интерполяции.

Например, если значение $y_k, k = \overline{1, n}$ в поквартальном ряде отсутствует, то его можно вычислить (оценить) с помощью следующей формулы:

$$y_k = (y_{k-1} + y_{k+1}) \frac{y_{k-4}}{y_{k-5} + y_{k-3}}. \quad (11)$$

Вычисление недостающего значения по формуле (11) позволяет учесть соседние значения и прошлогоднюю тенденцию изменения ряда.



8.1.3. Фильтрация временного ряда

На данном этапе осуществляется выявление аномальных значений и замена их значениями, соответствующими тенденциям динамики прогнозируемого параметра.

Аномальным значением называют такую величину члена ряда, которая существенно отклоняется от значений близких элементов.

Устранение аномальных явлений осуществляется двумя способами, которые могут быть использованы по отдельности или последовательно: сглаживание и выравнивание.

Сглаживание применяется для устранения случайных отклонений (шума) из экспериментальных значений исходного ряда.

Сглаживание проводится с помощью многочленов, приближающих группы опытных точек (методом наименьших квадратов). Чаще всего применяют линейную зависимость. При сглаживании пересчет ведется не только для текущих значений y_k , $k = \overline{1, n}$, но и для соседних.

Для трех точек используются следующие формулы:

$$y'_k = \frac{1}{3}(y_{k-1} + y_k + y_{k+1}),$$

$$y'_{k-1} = \frac{1}{6}(5y_{k-1} + 2y_k - y_{k+1}),$$

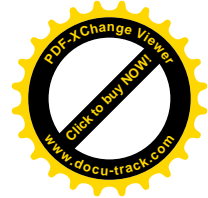
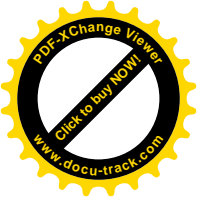
$$y'_{k+1} = \frac{1}{6}(-y_{k-1} + 2y_k + 5y_{k+1}).$$

Для пяти точек используются следующие формулы:

$$y'_k = \frac{1}{5}(y_{k-2} + y_{k-1} + y_k + y_{k+1} + y_{k+2}),$$

$$y'_{k-1} = \frac{1}{10}(4y_{k-2} + 3y_{k-1} + 2y_k + y_{k+1}),$$

$$y'_{k+1} = \frac{1}{10}(y_{k-1} + 2y_k + 3y_{k+1} + 4y_{k+2}),$$



$$y'_{k-2} = \frac{1}{5}(3y_{k-2} + 2y_k + y_{k+1} - y_{k+2}),$$

$$y'_{k+2} = \frac{1}{5}(-y_{k-2} + y_0 + 2y_{k+1} + 3y_{k+2}).$$

Сглаживание (даже в простом линейном варианте) является во многих случаях эффективным средством выявления тренда при наличии в экспериментальных точках случайных помех и ошибок измерения.

Выравнивание применяется для представления исходного ряда без изменения его числовых значений. Для этого исходную эмпирическую формулу $y = f(t, a, b)$, где t – момент времени, a, b – параметры, приводят к следующему виду:

$$Y = a_1 T + b_1.$$

Здесь используется двухпараметрическая зависимость, которая широко распространена и проста в получении. Использование большего числа параметров влечет за собой получение громоздких формул и, зачастую, неудачу в выравнивании.

Наиболее часто при выравнивании используется логарифмирование и замену переменной.

Ниже приводятся примеры выравнивания для некоторых эмпирических формул.

Пример 1:

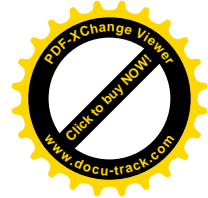
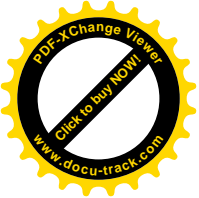
$$y = at^b.$$

Логарифмирование: $\lg y = \lg a + b \lg t$. Замена переменных $T = \lg t$, $Y = \lg y$, $b_1 = \lg a$, $a_1 = b$.

Пример 2:

$$y = ae^{bt}.$$

Логарифмирование: $\lg a = \lg a + b t \lg e$. Замена переменных $a_1 = b \lg e$, $b_1 = \lg a$, $T = t$.



Пример 3:

$$y = \frac{1}{a + be^{-t}}.$$

Выравнивание: $Y = \frac{1}{y}$, $T = e^{-t}$, $a_1 = b$, $b_1 = a$.

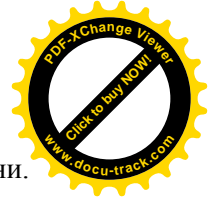
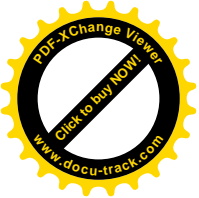
С помощью выравнивания мы имеем линейную зависимость, для которой значительно проще определить коэффициенты. После прогнозирования полученные значения необходимо пересчитать по формулам, обратным исходному преобразованию.

На практике метод выравнивания используется для определения приближенных значений параметров аппроксимирующей функции.

8.1.4. Логический отбор аппроксимирующей функции

На данном этапе осуществляется анализ исходного ряда данных на предмет изменения его значений во времени. При этом учитываются условия протекания рассматриваемого процесса и требования, предъявляемые к математической модели. Этап предусматривает рассмотрение и решение следующих вопросов:

- 1) какой функции соответствует поведение значений исследуемого параметра (монотонной, стабильно монотонной, периодической, имеющей один или несколько экстремумов, функции детерминированного хаоса);
- 2) имеются ли ограничения для показателя сверху и снизу, связанные с постановкой задачи;
- 3) имеет ли функция, определяющая процесс, точки перегиба;
- 4) обладает ли анализируемая функция свойством симметричности (такие случаи могут проявляться, например, при рассмотрении жизненного цикла товара);



5) накладываются ли на процесс ограничения во времени.

Ниже приведены классы наиболее часто используемых функций.

Степенной полином:

$$y(t) = a_0 + \sum_{i=1}^m a_i t^i;$$

экспоненциальный полином:

$$y(t) = \exp\left(a + \sum_{i=1}^m a_i t^i\right);$$

гиперболический полином:

$$y(t) = a_0 + \sum_{i=1}^m \frac{1}{a_i t^i},$$

где y – прогнозируемый показатель,

t – момент времени,

a_0, a_1, \dots, a_m – параметры (коэффициенты) экстраполирующих функций.

На практики хорошие результаты показывают линейная, квадратичная и степенная ($y(t) = at^b$) функции. Они содержат малое количество параметров и в прогнозном периоде показывают большую точность.

Также хорошие результаты получаются, когда используются следующие функции:

Экспоненциальная:

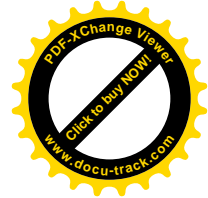
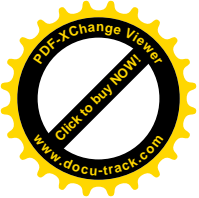
$$y(t) = ae^{bt};$$

модифицированная экспоненциальная:

$$y(t) = k - ae^{-bt};$$

гиперболическая:

$$y(t) = a + \frac{b}{c+t};$$



логистическая:

$$y(t) = \frac{k}{1 + be^{-et}}.$$

При выборе вида аппроксимирующей функции, когда есть возможность, прибегают к графическому способу подбора по виду точек временного ряда. Иногда вместо значений ряда используют их первые разности $z_k = y_k - y_{k-1}$, $k = \overline{2, n}$ или производную аппроксимирующей функции.

Для программной реализации выбора аппроксимирующей функции осуществляют верификацию прогноза по ретроспективному периоду, в которой используются остатки:

$$e_i = y_i - f(a, t_i),$$

где t_i – i -й момент времени, $i = \overline{1, n}$;

$f(a, t_i)$ – значение экстраполирующей функции в момент времени t_i ;

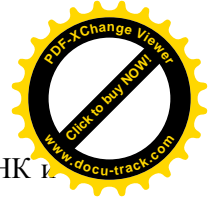
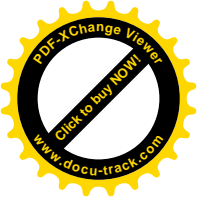
a – вектор параметров экстраполирующей функции;

y_i – фактическое значение параметра в момент времени t_i .

Когда средняя величина остатков равна или близка нулю, то функция может быть внесена в список прогнозных функций. Обычно при прогнозировании выбирают сразу несколько аппроксимирующих функций. Окончательный выбор осуществляется путем сравнения точности прогнозирования по ретроспективному периоду.

8.1.5. Оценка математической модели прогнозирования

На данном этапе определяются параметры различных аппроксимирующих функций. Для этого используется метод наименьших квадратов (МНК) и его модификации, метод экспоненциального сглаживания, метод вероятностного моделирования, метод адаптивного сглаживания.



Мы подробно рассмотрим 2 метода: модификацию МНК и метод экспоненциального сглаживания.

Метод наименьших квадратов, как мы уже знаем, сводится к минимизации квадратов разности расчетного значения параметра и фактического на ретроспективном периоде:

$$S = \sum_{i=1}^n (\hat{y}_i - y_i)^2 \rightarrow \min, \quad (12)$$

где \hat{y}_i – расчетные значения исходного ряда,

y_i – фактические значения исходного ряда, n – число наблюдений.

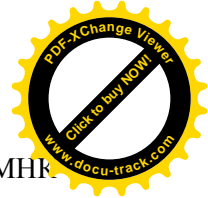
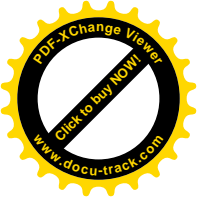
МНК с соотношением (12) называют классическим МНК. В этом случае предполагается, что все значения ряда являются равноправными. Если это не так, то каждое слагаемое умножается на коэффициент дисконтирования $\beta_i < 1, i = \overline{1, n}; \sum_{i=1}^n \beta_i = 1$.

Такая модификация носит название «Метод взвешенных наименьших квадратов» (МВНК). В данном методе соотношения (12) принимает вид:

$$S_w = \sum_{i=1}^n \beta_i (\hat{x}_i - x_i)^2 \rightarrow \min.$$

Коэффициенты β_i могут быть заданы в числовой форме или в виде функции. Обычно строят возрастающие последовательности, показывая, что будущее поведение системы в большей степени определяется поздними наблюдениями, чем ранними.

Метод наименьших квадратов является одним из самых простых с точки зрения реализации на ЭВМ. За этой простотой, к сожалению, стоят три серьезных недостатка:



1) Модель тренда жестко фиксируется, и с помощью МНК можно получить прогноз на небольшой период. Из этого следует, что МНК относят к методам краткосрочного прогнозирования.

2) Выбор вида модели весьма затруднителен. Также сложно обосновать выбор весов в МВНК.

3) Метод МНК хорошо реализуется для линейных и реализуемых зависимостей. В этом случае нахождение коэффициентов модели сводится к решению системы линейных уравнений. Задача значительно усложняется, если для прогноза используется функциональная зависимость, не сводимая к линейной.

Метод экспоненциального сглаживания является эффективным и надежным методом среднесрочного прогнозирования.

В данном методе предполагается, что чем ближе точка (значение по времени) к прогнозному периоду, тем выше их ценность. Такое положение лежит в основе метода экспоненциального сглаживания.

Сущность метода заключается в сглаживании исходного динамического ряда взвешенной скользящей средней, веса которой ($w_i, i = \overline{1, n}$) подчиняются экспоненциальному закону (рис. 9).

Пусть исходный динамический ряд описывается полиномом следующего вида:

$$y_t = b_0 + b_1 t + \frac{b_2}{2!} t^2 + \dots + \frac{b_p}{p!} t^p + \varepsilon_t,$$

где b_0, b_1, \dots, b_p – коэффициенты полинома,

p – порядок (степень) полинома,

ε_t – случайная составляющая.

Величина $S_t^k(y)$ называется экспоненциальной средней k -го порядка для ряда y_t .

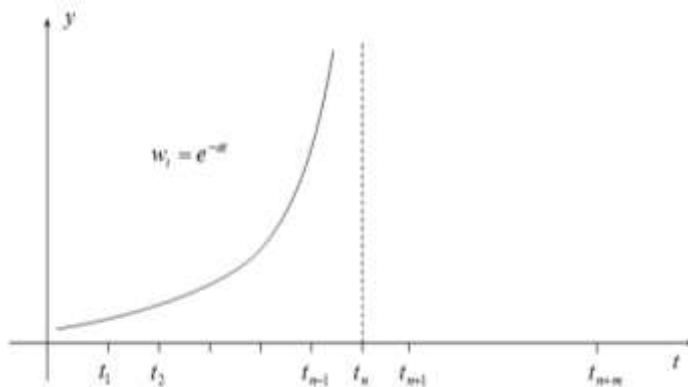
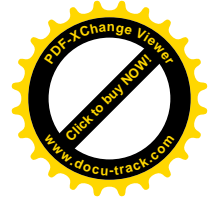
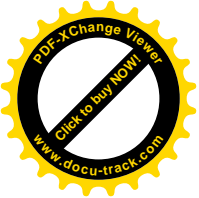


Рис. 9. Коэффициент экспоненциального сглаживания

Параметр сглаживания α определяет влияние исходного ряда наблюдений. Чем больше α , тем больше вклад последних наблюдений в формирование тренда, а влияние начальных условий убывает быстро. При малом значении параметра α прогнозные оценки учитывают все наблюдения, при этом уменьшение влияния более ранней информации происходит медленно.

В расчетах экспоненциальная средняя определяется с помощью рекуррентной формулы Брауна:

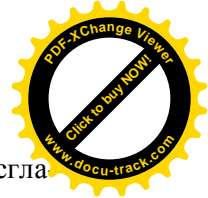
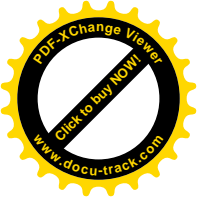
$$S_t^k(y) = \alpha S_t^{k-1}(y) + (1 - \alpha) S_{t-1}^k(y). \quad (13)$$

Для того чтобы использовать соотношение (13), необходимо задать начальные условия $S_0^l, l = \overline{1, k}$ по формуле Брауна-Мейера:

$$S_0^l = \sum_{i=0}^n (-1)^i \frac{\hat{b}_i}{n!} \frac{\alpha(1-\alpha)}{(l-1)!} \sum_{j=0}^{\infty} j^n (1-\alpha)^j \frac{(n-1+j)!}{j!}, \quad (14)$$

где $\hat{b}_i, i = \overline{0, n}$ – оценки коэффициентов.

Оценку коэффициентов прогнозирующего полинома определяют через экспоненциальные средние по фундаментальной теореме Брауна-Мейера.



Рассмотрим применение метода экспоненциального сглаживания для линейного и параболического тренда, которые наиболее популярны.

8.1.6. Линейная модель Брауна

Она имеет следующий вид:

$$y_t = b_0 + b_1 t + \varepsilon_t. \quad (15)$$

Начальные приближения (условия) для случая линейного тренда равны (по формуле (14)):

экспоненциальная средняя первого порядка:

$$S_1^1(y) = b_0 - \frac{1-\alpha}{\alpha} b_1;$$

экспоненциальная средняя второго порядка:

$$S_1^2(y) = b_0 - \frac{2(1-\alpha)}{\alpha} b_1;$$

Экспоненциальные средние первого и второго порядков в момент времени t (см. формулу (15)):

$$S_2^1(y) = \alpha y_1 + (1-\alpha) S_1^1(y);$$

$$S_2^2(y) = \alpha S_2^1(y) + (1-\alpha) S_1^2(y).$$

Для $t > 2$:

$$S_t^1(y) = \alpha y_t + (1-\alpha) S_{t-1}^1(y);$$

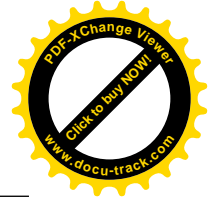
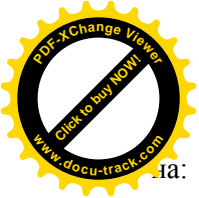
$$S_t^2(y) = \alpha S_t^1(y) + (1-\alpha) S_{t-1}^2(y).$$

Оценки коэффициентов линейного тренда:

$$\hat{b}_0 = 2 S_t^1(y) - S_t^2(y),$$

$$\hat{b}_1 = \frac{\alpha}{1-\alpha} [S_t^1(y) - S_t^2(y)].$$

Прогноз на m шагов (в моменты времени $t = \overline{t_{n+1}, t_{n+m}}$) равен $y_t = \hat{b}_0 + \hat{b}_1 t$. При этом оценка ошибки прогноза будет рав-



$$\sigma = \sigma_{\varepsilon_t} \sqrt{\frac{\alpha}{(2-\alpha)^3} \left[1 + 4(1-\alpha) + 5(1-\alpha)^2 + 2\alpha(4-3\alpha)t_m + 2\alpha^2 t_m^2 \right]}.$$

8.1.7. Параболический тренд

$$y_t = b_0 + b_1 t + \frac{1}{2} b_2 t^2 + \varepsilon_t.$$

Начальные приближения

$$S_0^1(y) = b_0 - \frac{1-\alpha}{\alpha} b_1 + \frac{(1-\alpha)(2-\alpha)}{2\alpha^2} b_2;$$

$$S_0^2(y) = b_0 - \frac{2(1-\alpha)}{\alpha} b_1 + \frac{(1-\alpha)(3-2\alpha)}{2\alpha^2} b_2;$$

$$S_0^3(y) = b_0 - \frac{3(1-\alpha)}{\alpha} b_1 + \frac{(1-\alpha)(4-3\alpha)}{2\alpha^2} b_2.$$

Экспоненциальные средние:

$$S_t^1(y) = \alpha y_t + (1-\alpha) S_{t-1}^1(y);$$

$$S_t^2(y) = \alpha S_t^1(y) + (1-\alpha) S_{t-1}^2(y);$$

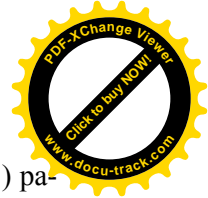
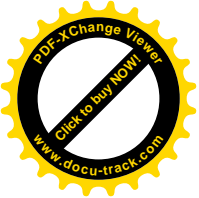
$$S_t^3(y) = \alpha S_t^2(y) + (1-\alpha) S_{t-1}^3(y).$$

Оценки коэффициентов параболической зависимости для тренда:

$$\hat{b}_0 = 3[S_t^1(y) - S_t^2(y)] + S_t^3(y);$$

$$\hat{b}_1 = \frac{\alpha}{(1-\alpha)^2} \left[(6-5\alpha) S_t^1(y) - 2(5-4\alpha) S_t^2(y) + (4-3\alpha) S_t^3(y) \right];$$

$$\hat{b}_2 = \frac{\alpha}{(1-\alpha)^2} \left[S_t^1(y) - 2S_t^2(y) + S_t^3(y) \right].$$



Прогноз на m шагов (в моменты времени $t = t_{n+1}, t_{n+m}$) равен $y_t = \hat{b}_0 + \hat{b}_1 t + \frac{1}{2} \hat{b}_2 t^2$. При этом оценка ошибки прогноза будет равна: $\sigma_y^* \approx \sigma_{\varepsilon_t} \sqrt{2\alpha + 3\alpha^2 + 3\alpha^3 t_1}$.

Для метода экспоненциального сглаживания основным и наиболее трудным моментом является выбор параметра сглаживания α . Он определяет оценки коэффициентов модели, а, следовательно, и результаты прогноза.

Приближенная оценка параметра метода α осуществляется двумя способами:

1) *Соотношение Брауна*, выведенное из условия равенства скользящей и экспоненциальной средней:

$$\alpha = \frac{2}{N+1}, \text{ где } N - \text{число точек ряда, для которых динамика}$$

считается однородной и устойчивой (число точек в интервале сглаживания). Иногда $\alpha = \frac{2}{n+1}$, где n – число наблюдений (точек) в ретроспективном периоде.

2) *Соотношение Мейера*:

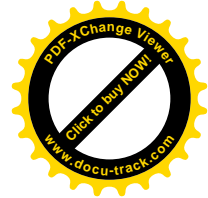
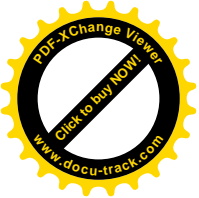
$$\alpha = \frac{\sigma_n}{\sigma_\varepsilon}, \text{ где } \sigma_n - \text{средняя квадратическая ошибка модели;}$$

σ_ε – средняя квадратическая ошибка исходного ряда.

Очевидно, что выбор параметра α нужно связывать с точностью прогнозирования. Для более обоснованного выбора α можно использовать процедуру обобщенного сглаживания, в результате которой получаются соотношения, связывающие оценку ошибки прогноза σ и параметр сглаживания:

для линейной модели:

$$\sigma^2 = \frac{\alpha}{(1+\beta)^2} [1 + 4\beta + 5\beta^2 + 2\alpha(1+3\beta)\tau + 2\alpha^2\tau^3] \sigma_\varepsilon^2;$$



для квадратичной модели:

$$\sigma^2 = \sigma_\varepsilon^2 (2\alpha + 3\alpha^3 + 3\alpha^2\tau),$$

где $\beta = 1 - \alpha$,

τ – горизонт прогнозирования,

σ_ε – дисперсия остатков на ретроспективном диапазоне.

Очевидно, что при уменьшении параметра сглаживания α снижается оценка ошибки прогноза, но при этом увеличивается риск получения некачественной экстраполяционной модели.

Качество прогноза во многом зависит от выбора порядка полинома. Известно, что превышение второго порядка модели не приводит к существенному увеличению точности прогноза, но значительно усложняет расчет.

В заключение отметим, что метод экспоненциального сглаживания является одним из наиболее эффективных, надежных и широко применяемых методов прогнозирования. Он позволяет получить оценку параметров тренда, характеризующих не средний уровень процесса, а тенденцию, сложившуюся к моменту последнего наблюдения, и при этом отличается простотой вычислительных операций.

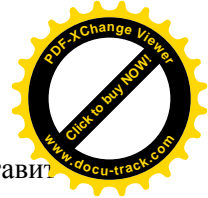
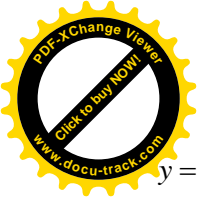
Проиллюстрируем применение метода экспоненциального сглаживания на конкретном примере.

Ниже, в таблице 3 приведены исходные данные.

Таблица 3.

Год	1997	1998	1999	2000
t	1	2	3	4
y_t	40	43	46	48

Очевидно, что представленный ряд исходных значений можно аппроксимировать с помощью линейной функции. Ее коэффициенты, вычисленные с помощью метода наименьших квадратов будут следующие: $a = 2,7$; $b = 37,5$, т.е.



$y = 2,7t + 37,5$. При этом относительная ошибка составит $\sigma_\varepsilon = 0,3$.

Параметр сглаживания при $N = 4$ будет равен $\alpha = \frac{2}{4+1} = 0,4$.

Найдем экспоненциальные средние.

Начальные условия:

$$S_1^1 = 37,5 - \frac{1-0,4}{0,4} 2,7 = 33,45;$$

$$S_1^2 = 37,5 - \frac{2(1-0,4)}{0,4} 2,7 = 29,4.$$

Вычислим экспоненциальные средние, начиная с момента времени $t = 2$ (мы делаем прогноз на ретроспективном периоде для $m = 1$ при $t = 2$) и с использованием вышеприведенных начальных условий.

$$S_2^1 = 0,4 \cdot 40 + 0,6 \cdot 33,45 = 36,07,$$

$$S_2^2 = 0,4 \cdot 36,07 + 0,6 \cdot 29,4 = 32,07.$$

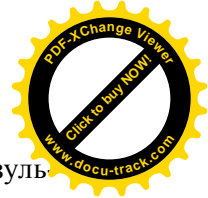
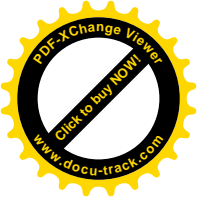
Соответственно пересчитываются коэффициенты регрессии a и b :

$$\hat{a}_2 = \frac{0,4}{1-0,4} (36,07 - 32,07) = 2,67,$$

$$\hat{b}_2 = 2 \cdot 36,07 - 32,07 = 40,07,$$

а прогнозируемое значение равно $\hat{y}_2 = 40,07 + 2,67 \cdot 1 = 42,74$. Мы умножаем на единицу, т.к. осуществляем прогноз на один период $m = 1$.

Аналогичным образом вычисляются необходимые величины для $t = \overline{3,5}$.



В таблице 4 сведены исходные, промежуточные и результирующие значения прогнозирования с помощью метода экспоненциального сглаживания.

Таблица 4.

t	y_t	$S_t^{[1]}$	$S_t^{[2]}$	\hat{a}_t	\hat{b}_t	\hat{y}_t	$\hat{y}_t - y_t$
1	40						
2	43	36,07	32,07	40,07	2,67	42,74	-0,26
3	46	38,84	34,78	42,90	2,71	45,62	-0,38
4	48	41,71	37,55	45,86	2,77	48,63	0,63
5 ($l=1$)	-	44,22	40,21	48,23	2,67	51,89	-

Окончательная модель прогнозирования

$$y_{t+m} = 2,67m + 48,23.$$

Оценка ошибки прогноза

$$\sigma = 0,46.$$

8.1.8. Выбор математической модели прогнозирования

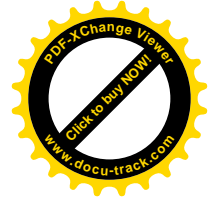
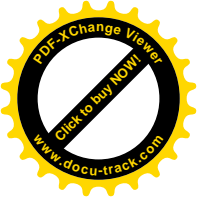
Выбор моделей прогнозирования базируется на оценке их качества. Качество модели определяется их точность и адекватностью. Адекватность характеризуется наличием и учетом определенных статистических свойств. Точность модели определяется степенью ее близости к фактическим данным.

Адекватность модели

Модель прогнозирования считается адекватной, если она учитывает существенную закономерность исследуемого процесса. В ином случае ее нельзя применять.

Закономерность исследуемого процесса находит отражение в наличии определенных статистических свойств остаточной компоненты. Их всего три:

- 1) независимость остатков;
- 2) случайность остатков и подчинение их нормальному закону распределения;



3) равенство суммы остатков нулю.

Независимость остаточной компоненты означает отсутствие автокорреляции между остатками $e_i = y_i - y_{it}$, $i = \overline{1, n}$. Наличие автокорреляции может привести к следующим последствиям:

- 1) дисперсия остатков регрессии будет недооцененной;
- 2) выборочная дисперсия параметров регрессии будет рассчитана с ошибкой, что станет препятствием к корректному применению МНК при построении модели исходного динамического ряда.

Для идентификации наличия автокорреляции остатков используется критерий Дарбина-Уотсона, в соответствии с которым вычисляется d - статистика:

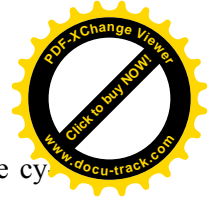
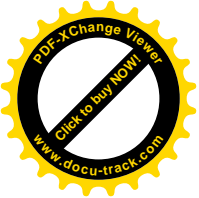
$$d = \frac{\sum_{i=2}^n [(y_i - y_{it}) - (y_{i-1} - y_{it-1})]^2}{\sum_{i=1}^n (y_i - y_{it})^2},$$

где y_i , $i = \overline{1, n}$ – фактические значения исходного ряда, y_{it} , $it = \overline{1, n}$ – теоретические значения ряда, n – объем выборки.

Установлено, что область значений величины d является отрезок $[0, 4]$. Согласно методу Дарбина-Уотсона существуют нижний d_n и верхний d_g пределы значений статистики d . Такие значения, зависящие от уровня значимости, объема выборки и числа объясняющих переменных (факторов), называют критическими. С ними сравнивается полученная величина d . При сравнении возникает один из четырех ниже перечисленных случаев (см. таблица 5, рис. 10).

Критерий Дабрина-Уотсона обладает двумя недостатками:

- 1) наличие области неопределенности, в которой нельзя прийти ни к какому выводу;



2) при объеме выборки менее 15 для статистики d не существуют критические значения d_H и d_ϵ .

Таблица 5.

№ п/п	Случай	Принимаемое решение
1	$d_\epsilon \leq d \leq 4 - d_\epsilon$	принимается гипотеза отсутствия автокорреляции остатков
2	$0 \leq d \leq d_H$	принимается гипотеза существования положительной автокорреляции остатков
3	$d_H \leq d \leq d_\epsilon$ или $4 - d_\epsilon \leq d \leq 4 - d_H$	при выбранном уровне значимости нельзя прийти к определенному выводу
4	$4 - d_H \leq d \leq 4$	принимается гипотеза о существовании отрицательной автокорреляции остатков

Случай:



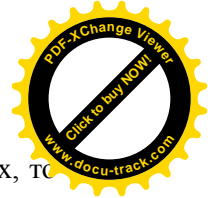
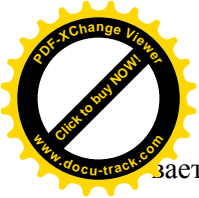
Рис. 10. Возможные случаи для значения статистики d

В этом случае для оценки независимости остатков можно использовать коэффициент автокорреляции:

$$r_a = 1 - \frac{d}{2}.$$

Расчетное значение r_a сравнивается с табличным r_{aT} . При этом степень свободы f равна n . Если $r_a \leq r_{aT}$, то остатки независимы.

Проверка случайности остатков осуществляется с помощью расчета количества так называемых точек поворота (критерий пиков и впадин). Здесь каждое значение ряда остатков сравни-



зается с двумя соседними. Если оно больше или меньше их, то данная точка ряда является поворотной.

Пусть всего поворотных точек K . Тогда, если остатки случайны, то должно выполняться следующее строгое неравенство:

$$K > \left[\frac{(2n-1)}{3} - 2\sqrt{\left(16n - \frac{29}{90}\right)} \right].$$

Соответствие ряда остатков нормальному закону распределения можно проверить с помощью критерия Колмогорова-Смирнова. Также такую проверку можно выполнить, используя коэффициенты асимметрии – A_c (мера «скошенности») и эксцесса – E_k (мера «скупенности»):

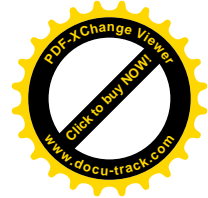
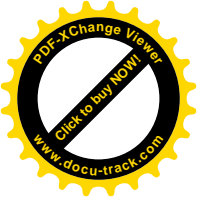
$$A_c = \frac{\frac{1}{n} \sum_{i=1}^n e_i^3}{\sqrt{\left[\frac{1}{n} \sum_{i=1}^n e_i^2 \right]^3}};$$
$$E_k = \frac{\frac{1}{n} \sum_{i=1}^n e_i^4}{\left(\frac{1}{n} \sum_{i=1}^n e_i^2 \right)^2} - 3.$$

Данные коэффициенты оцениваются на близость к нулю. Для этого вычисляются среднеквадратические отклонения:

$$S_a = \frac{6(n-2)}{(n+1)(n+3)};$$
$$S_e = \frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}.$$

Если выполняются соотношения:

$$A_c \leq 1,5S_a;$$



$$E_k \leq 1,5S_e,$$

то считается, что распределение ряда остатков не противоречит нормальному закону. В случае $A_c > 2S_a$ или $E_k > 2S_e$, то распределение ряда не соответствует нормальному закону распределения, и построение прогноза и его доверительных интервалов не правомочно. Если $1,5S_a < A_c \leq 2S_a$ или $1,5S_e < E_k \leq 2S_e$, т.е. коэффициенты попадают в зону неопределенности, то можно использовать RS - критерий:

$$RS = (E_{\max} - E_{\min}) / S,$$

где $E_{\max} = \max_{i=1,n} e_i$, $E_{\min} = \min_{i=1,n} e_i$, S – среднее квадратическое

отклонение остатков.

Если значение RS попадает между табулированными границами с заданным уровнем значимости, то гипотеза о нормальном распределении ряда остатков принимается.

Равенство нулю средней ошибки можно проверить с помощью критерия Стьюдента:

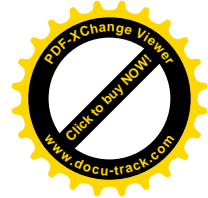
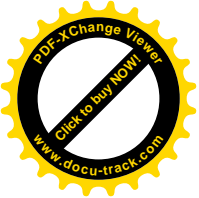
$$t_p = \left| \frac{1}{n} \sum_{i=1}^n e_i \right| \sqrt{\frac{n}{S}}.$$

Если t_p больше табличного значения при числе степеней свободы статистики t $f = n - 1$ и заданном уровне значимости α , то гипотеза равенства нулю средней ошибки отклоняется.

Оценка точности модели f выполняется с помощью ряда характеристик. Ниже приведены наиболее часто используемые показатели:

- 1) оценка стандартной ошибки;
- 2) средняя относительная ошибка модели;
- 3) среднее линейное отклонение;
- 4) ширина доверительного интервала.

Оценка стандартной ошибки осуществляется по формуле:



$$S_{1,f(x)} = \sqrt{\frac{\sum_{i=1}^n [y_i - f(x_i)]^2}{n - p}},$$

где n – число наблюдений, p – число определяемых коэффициентов модели.

Средняя относительная ошибка оценки вычисляется по формуле:

$$\bar{m}_\alpha = \frac{1}{n} \sum_{i=1}^n \frac{y_i - f(x_i)}{f(x_i)} 100\%.$$

Среднее линейное отклонение вычисляется по формуле:

$$B = \frac{\sum_{i=1}^n |y_i - f(x_i)|}{\sqrt{n(n-1)}}.$$

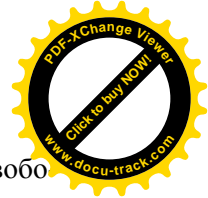
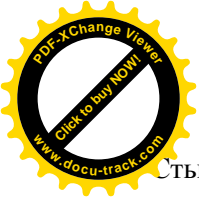
Ширина доверительного интервала в точке прогноза

Данный показатель позволяет определить, в каком диапазоне может находиться прогнозное значение. При вычислении ширины доверительного интервала оценивают верхнее и нижнее значение посредством определения стандартных ошибок коэффициентов модели.

Для решения задач используются интервальные оценки коэффициентов модели:

$$a_i^I = a_i \pm t_p S_{a_i}, \quad (16)$$

где $a_i, i = \overline{0, m}$ – коэффициенты модели (предполагается, что в ней присутствует свободный член, в противном случае i изменяется от единицы до m), $S_{a_i}, i = \overline{0, m}$ – стандартное отклонение коэффициента (см. статистическую значимость коэффициентов регрессии), t_p – теоретическое значение критерия



Стьюдента при уровне значимости 0,05 и числе степеней свободы $f_1 = n - m - 1$.

Если работа ведется с аппроксимацией вида $y_t = a_1 x + a_0$, то стандартные отклонения вычисляются по следующим формулам:

$$S_{a_0} = \sqrt{\frac{\sigma^2}{n}}, \quad S_{a_1} = \sqrt{\frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}},$$

где σ^2 – среднеквадратическое отклонение исходного ряда. Следует обратить внимание на то, что S_{a_0} для многофакторной модели вычисляется точно также как для вышеуказанной аппроксимации.

Несмещенная оценка дисперсии случайной составляющей:

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - y_{it})^2.$$

Используя соотношение вида (16) вычисляются верхняя и нижняя границы доверительного интервала Y_u, Y_d :

$$Y_u = f(a^u, x); \quad Y_d = f(a^d, x),$$

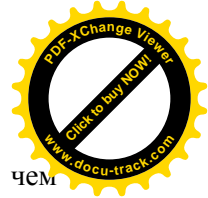
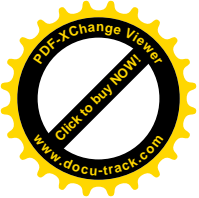
где $a_i^u = a_i + t_p S_{a_i}$, $a_i^d = a_i - t_p S_{a_i}$, x – вектор факторов, используемый для прогнозирования.

По найденным границам вычисляется ширина доверительного интервала:

$$\Delta = Y_u - Y_d.$$

Из вышесказанного следует, что ширина доверительного интервала зависит:

1) от числа степеней свободы, и как следствие, от объема выборки, т.е. чем больше объем выборки, тем, при прочих равных условиях, меньше ширина доверительного интервала;



2) от величины стандартного отклонения $S_{a_i}, i = \overline{0, m}$: чем больше данная величина, тем шире интервал.

Итогом работ по выбору вида математической модели прогноза является формирование ее обобщенных характеристик. В них должны входить:

- 1) вид уравнения модели и значения его параметров;
- 2) оценка точности и адекватности модели;
- 3) точечные и интервальные прогнозные оценки.

8.2. Модель Хольта-Уинтерса

Модель Хольта-Уинтерса относится к классу мультипликативных моделей с линейным ростом и позволяет проводить экстраполяционный прогноз с учетом сезонного спектра исходного ряда и общей тенденции поведения параметра. При этом предполагается, что временной ряд разбит на равные интервалы, которые называют сезонами.

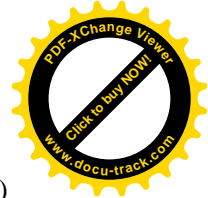
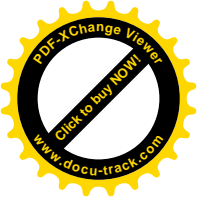
Пусть ряд Y разбит на k сезонов по L значений. Всего ряд содержит kL значений. Модель Хольта-Уинтерса для таких обозначений имеет следующий вид:

$$y_t(t + \tau) = [a(t) + b(t)]F(t + \tau - L),$$

где $y_t(t + \tau)$ – прогнозная оценка, τ – период упреждения (шаг прогнозирования), $a(t), b(t), F(t + \tau - L)$ – адаптируемые коэффициенты модели.

Последний множитель модели F называют коэффициентом сезонности. Все коэффициенты адаптируются (т.е. пересчитываются) при переходе между сезонами, т.е. от членов ряд с номером $t - L$ к t .

Задача прогнозирования параметра с помощью модели Хольта-Уинтерса сводится к нахождению ее коэффициентов, которые вычисляются по следующим рекуррентным формулам:



$$b(t) = \alpha_1 \frac{y(t)}{F(t-L)} + (1 - \alpha_1)[b(t-1) + a(t-1)], \quad (17)$$

$$a(t) = \alpha_2[b(t) - b(t-1)] + (1 - \alpha_2)a(t-1), \quad (18)$$

$$F(t) = \alpha_3 \frac{y(t)}{b(t)} + (1 - \alpha_3)F(t-L). \quad (19)$$

В данных формулах присутствуют переменные α_1 , α_2 , α_3 . Такие переменные называются параметрами сглаживания. Они задаются специалистом в диапазоне от нуля до единицы. Их нахождение для модели, описывающей тот или иной процесс – отдельная задача.

Очевидно, что для вычисления необходимых коэффициентов нам потребуется вычислить их начальные значения $a(0)$, $b(0)$, $F(-3)$, $F(-2)$, $F(-1)$, $F(0)$.

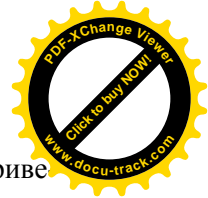
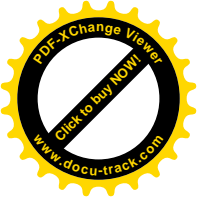
Величины $a(0)$ и $b(0)$ являются коэффициентами линейной аппроксимации исходного временного ряда $y = (y_1, y_2, \dots, y_{kL})$.

Обозначим через y_{tl} значения аппроксимации в момент времени $t = \overline{1, kL}$, т.е. $y_{tl} = a(0)t + b(0)$. Тогда оставшиеся начальные значения вычисляются по следующим формулам:

$$F(-p) = \frac{1}{k} \sum_{i=1}^k \frac{y_{z(p,i)}}{y_{z(p,i),l}}, \quad p = \overline{0, L-1}, \text{ где } z_i = iL - p.$$

Имея начальные коэффициенты, мы можем вычислять коэффициенты модели от сезона к сезону и перейти, тем самым, к прогнозированию. Очевидно, что горизонт прогнозирования кратен величине сезона L . При этом рекомендуют осуществлять прогноз только на один сезон $\tau = \overline{1, L}$.

Оценку адекватности модели можно осуществить с помощью ранее приведенной методики (см. прогнозирование с помощью экстраполяции).



Для иллюстрации работы модели Хольта-Уинтерса приведем пример.

Даны следующие значения параметра y (см. таблица 6).

Таблица 6.

t	1	2	3	4	5	6	7	8	9	10	11	12
y	49	78	84	94	65	99	111	142	103	189	152	181

Пусть сезон состоит из четырех значений $L=4$, тогда $k=3$. Аппроксимация для данного ряда имеет следующий вид:

$$y_{it} = 10,98t + 40,86.$$

Отсюда $a(0)=10,98$; $b(0)=40,86$.

Вычислим начальные значения коэффициента сезонности.

Для этого рассчитаем значения аппроксимации y_{it} , $t=\overline{1,12}$ в каждый момент времени исторического периода (см. таблица 7):

Таблица 7.

t	1	2	3	4	5	6
y_{it}	51,79	62,72	73,65	84,58	95,51	106,44
t	7	8	9	10	11	12
y_{it}	117,37	128,3	139,23	150,16	161,09	172,02

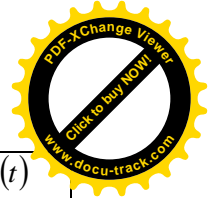
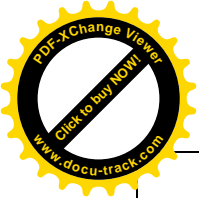
$$\text{Отсюда } F(-3) = \frac{1}{3} \left(\frac{49}{51,79} + \frac{65}{95,51} + \frac{103}{139,23} \right) = 0,79,$$

$$F(-2) = 1,14, \quad F(-1) = 1,01, \quad F(0) = 1,09.$$

Пусть заданы следующие параметры сглаживания: $\alpha_1 = 0,4$, $\alpha_2 = 0,3$, $\alpha_3 = 0,6$. Тогда по формулам (17)–(19) можно последовательно находить коэффициенты модели (см. таблица 8).

Таблица 8.

t	$a(t)$	$b(t)$	$F(t)$	$y(t)$	$y_t(t)$
-3	—	—	0,787091	—	—
-2	—	—	1,141154	—	—



t	$a(t)$	$b(t)$	$F(t)$	$y(t)$	$y_t(t)$
-1	—	—	1,007124	—	—
0	10,98	40,864	1,086843	—	—
1	8,276464	42,83221	1,001236	49	40,22719
2	7,366116	48,07418	1,429957	78	67,76762
3	8,070179	57,78718	1,275015	84	68,14361
4	7,640718	64,42582	1,310162	94	82,53274
5	4,032448	60,03897	1,050072	65	64,50543
6	3,200137	61,29705	1,541034	99	103,0521
7	4,755359	69,68126	1,465787	111	101,628
8	7,117056	82,30894	1,559189	142	114,44
9	5,594413	84,35052	1,152686	103	93,90377
10	7,504419	96,31162	1,793842	189	148,7762
11	4,788733	94,76375	1,548708	152	151,9436
12	5,048805	100,4194	1,70514	181	172,7226

Прогнозные значения $y_t(12+\tau)$, $\tau = \overline{1,4}$ приведены в таблице 9.

Таблица 9.

τ	1	2	3	4
$y_t(12+\tau)$	115,752	189,1933	171,1586	197,0559



ЛЕКЦИЯ 9. ПАРАМЕТРЫ И ХАРАКТЕРИСТИКИ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

Случайный характер поступления заявок и продолжительности их обслуживания приводит к тому, что в СМО протекает *случайный процесс*. Для эффективной организации работы СМО необходимо изучить данный процесс, описав с использованием аппарата теории вероятностей и математической статистики.

Рассмотрим основные характеристики СМО.

1. Входной поток заявок, поступающих на обслуживание.

Для наиболее точного расчета характеристик СМО обычно используется простейший пуассоновский поток заявок.

Простейшим называется поток событий (в данном случае поступления заявок), обладающий свойствами стационарности, ординарности, отсутствия последствия.

Стационарность потока означает, что вероятностные характеристики такого потока не должны меняться в зависимости от времени. В частности, так называемая интенсивность (или «плотность») потока событий — среднее число событий в единицу времени — для стационарного потока должна оставаться постоянной.

Примечание. Постоянство интенсивности не означает, что фактическое число событий, появляющихся в единицу времени, постоянно — нет, поток может иметь местные сгущения и разрежения. Важно, что для стационарного потока эти сгущения и разрежения *не носят закономерного характера*, а среднее число событий, попадающих на единичный участок времени, остается постоянным для всего рассматриваемого периода. На практике часто встречаются потоки событий, которые (по крайней мере, на ограниченном участке времени) могут рассматриваться как стационарные. Например, поток вызовов, поступающих на телефонную станцию, скажем, на интервале от 12 до 13 часов может считаться стационарным. Тот же поток в течение целых суток уже не будет стационарным (ночью интенсивность потока вызовов гораздо



меньше, чем днем). Заметим, что так же обстоит дело и с большинством физических процессов, которые мы называем «стационарными» — в действительности они стационарны только на ограниченном участке времени, а распространение этого участка до бесконечности — лишь удобный прием, применяемый в целях упрощения.

Для потока характерно *отсутствие последействия*, если события, образующие поток, появляются в последовательные моменты времени независимо друг от друга.

Примечание. Поток пассажиров, входящих на станцию метро, можно считать потоком без последействия, потому что причины, обусловившие приход отдельного пассажира именно в данный момент, а не в другой, как правило, не связаны с аналогичными причинами для других пассажиров. Если такая зависимость появляется, условие отсутствия последействия оказывается нарушенным.

Ординарность потока означает, что события в потоке приходят поодиночке, а не парами, тройками и т. д.

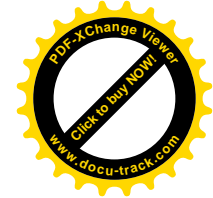
Примечание. Например, поток клиентов, направляющихся в парикмахерскую, практически можно считать ординарным, чего нельзя сказать о потоке клиентов, направляющихся в ЗАГС для регистрации брака. Поток атак истребителей по бомбардировщику, находящемуся над вражеской территорией, ординарен, если они атакуют цель поодиночке, и не ординарен, если они идут в атаку парами или тройками.

Для простейшего пуассоновского потока вероятность поступления k заявок на обслуживание в течение интервала времени t определяется по формуле:

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k \geq 0, \quad t \geq 0, \quad (9.1)$$

λ - интенсивность потока заявок (количество заявок, которые поступили в систему за единицу времени).

Плотность распределения интервалов времени между событиями в этом потоке определяется по экспоненциальному закону (это означает, что интервалы



времени между заявками могут быть как очень короткими, так и очень длинными)

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0. \quad (9.2)$$

Примечание.

Рассмотрим самый простой поток с интенсивностью 1 и обозначим моменты поступления требований на осе $(0, t)$ как показано на рис. 9.1. Определим, какое распределение имеют промежутки времени T между моментами поступления двух соседних требований.

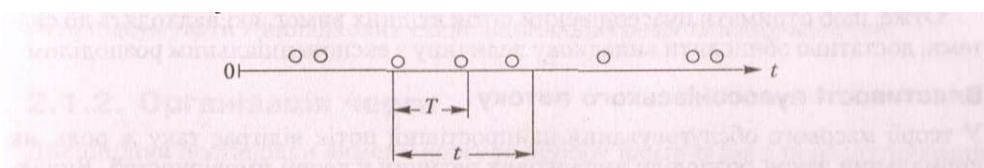


Рис. 9.1. Моменты поступления требований для пуассоновского потока

Очевидно, что промежутки времени T - случайные величины. Найдем закон их распределения. Функция распределения $F(t)$ определяет вероятность того, что случайная величина приобретет значение, которое меньше за t , то есть

$$F(t) = P, \quad T < t.$$

Пусть t_0 - начало промежутка времени T . Найдем вероятность того, что случайная величина T будет меньше за t . Для этого нужно, чтобы на промежуток длиной t , который начинается с точки t_0 , попало хотя бы одно требование. Вычислим функцию $F(t)$ через вероятность противоположного события, то есть через вероятность P_0 того, что за промежуток времени t к системе не поступит ни одного требования:

$$F(t) = 1 - P_0.$$

Значение вероятности P_0 найдем за формулой (9.1) при условии, что $t=0$:

$$F(t) = 1 - e^{-\lambda t}, \quad t > 0.$$

Тогда функция распределения случайной величины T будет иметь вид:

Чтобы найти функцию плотности распределения $f(t)$ случайной величины T , продифференцируем функцию $F(t)$ за t :

$$f(t) = \lambda e^{-\lambda t}, \quad t > 0.$$

Это и является функцией плотности показательного или экспоненциального закона распределения, ее графики и графики функции распределения при условии $\lambda = 1$ изображено на рис. 9.2.

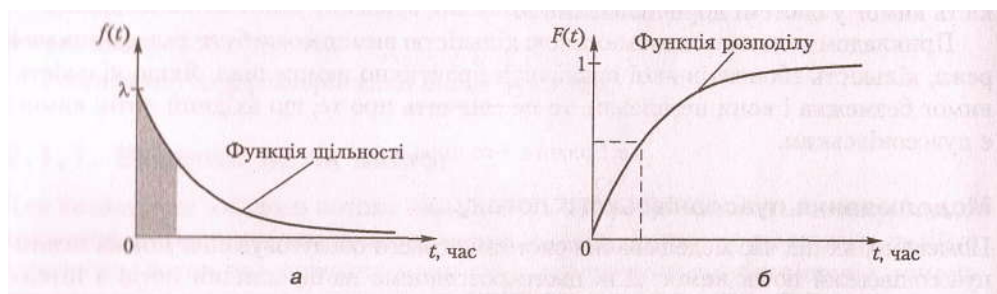


Рис. 9.2. Графика функций плотности (а) и распределения (б) экспоненциального закона:

$$\lambda = 1; \rho = 0.5$$

Следовательно, чтобы получить пуассоновский поток входных заявок, которые поступают к системе, достаточно вычислить случайную величину с экспоненциальным распределением.

2. Выходной поток обслуженных заявок.

Выходной поток требований для СМО с простейшим потоком входных заявок также является простейшим. Время обслуживания заявок, определяющее *интенсивность выходного потока* μ , также распределено по экспоненциальному закону:

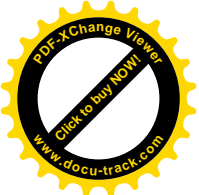
$$f(t) = \mu e^{-\mu t} \quad (t > 0) \quad (9.3)$$

3. Организация работы СМО

Обслуживание осуществляется с использованием одного или нескольких идентичных приборов – каналов обслуживания. В соответствии с количеством каналов (m) выделяют:

- одноканальные СМО;
- многоканальные СМО.

Кроме того, различают два типа СМО:



- системы с отказами – в таких системах заявка, поступившая в момент, когда все каналы заняты, получает «отказ» и покидает СМО;
- системы с ожиданием (очередью) – в таких системах заявка, поступившая в момент, когда все каналы заняты, становится в очередь и ожидает освобождения одного из каналов.

Системы с очередью делятся также на два типа:

- системы с неограниченным ожиданием – заявка, поступившая в СМО всегда будет обслужена;
- системы с ограниченным ожиданием – на пребывание заявки в очереди накладываются ограничения (например, длина очереди, время пребывания заявки в очереди).

Рассмотренные характеристики относятся к Марковским СМО, т.е. СМО типа М/М/т.

Оценка эффективности работы СМО

Для оценки эффективности работы СМО используется ряд статистических и экономических величин.

Статистическая оценка эффективности

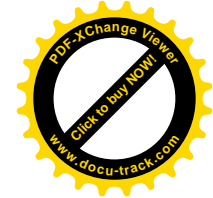
1. *Вероятность простоя СМО (P_0)* показывает, какую часть от общего времени работы СМО все ее каналы свободны, т.е. простаивают из-за отсутствия заявок.

2. *Вероятность отказа ($P_{отк}$)* характеризует, какая доля всех поступающих заявок не обслуживается системой из-за занятости ее каналов или большого количества заявок в очереди. Для СМО без ограничений на очередь $P_{отк} = 0$.

3. *Вероятность обслуживания ($P_{обсл}$)* показывает, какая доля всех поступающих заявок обслуживается системой.

Очевидно, что $P_{обсл} = 1 - P_{отк}$.

Для СМО без отказов $P_{обсл} = 1$.



4. *Вероятности состояний* обычно используются в качестве промежуточных величин для вычисления других характеристик СМО (например, P_j – вероятность пребывания в СМО ровно j заявок).

5. *Нагрузка на СМО* определяется как отношение интенсивности потока заявок к интенсивности, с которой СМО может их обслуживать:

$$\rho = \frac{\lambda}{\mu m} \quad (9.4)$$

где λ – интенсивность поступления заявок;

μ – интенсивность обслуживания заявок;

m – количество обслуживающих каналов.

Любая СМО без ограничений на очередь может нормально работать (т.е. обслуживать все поступающие заявки) только при условии, что $\rho < 1$. Величина $\rho > 1$ означает, что количество заявок, поступающих в СМО в единицу времени (λ), превышает количество заявок, которые СМО может обслужить в единицу времени ($m\mu$). В таких условиях в СМО без ограничений на очередь количество заявок, ожидающих обслуживания, будет постоянно возрастать, так как заявки будут поступать в СМО быстрее, чем она может их обслуживать. Для СМО с ограничениями на очередь и без очереди возможны любые значения ρ , так как в таких СМО часть заявок получает отказ, т.е. не допускается в СМО.

Коэффициент загрузки СМО:

$$U = \rho(1 - P_{отк}) \quad (9.5)$$

1. Среднее число занятых каналов:

$$\bar{S} = mU \quad (9.6)$$

2. Среднее число заявок в СМО:

$$\bar{k} = \bar{q} + \bar{S} \quad (9.7)$$

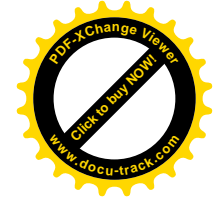
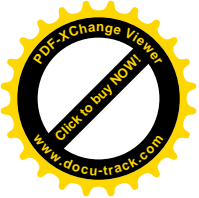
где \bar{q} – средняя длина очереди.

3. Пропускная способность СМО:

$$\gamma = \mu \bar{S} \quad (9.8)$$

или

$$\gamma = \lambda(1 - P_{отк}) \quad (9.9)$$



4. Среднее время пребывания заявки в очереди (формула Литтла):

$$\bar{w} = \frac{\bar{q}}{\gamma} \quad (9.10)$$

5. Среднее время пребывания заявки в СМО:

$$\bar{t} = \bar{w} + \bar{x} \quad (9.11)$$

или

$$\bar{t} = \frac{\bar{k}}{\gamma} \quad (9.12)$$

Формулы (9.5)–(9.12) могут применяться для расчета характеристик *любых* разомкнутых СМО, независимо от количества каналов, потока заявок, закона распределения времени обслуживания и т.д. Для разомкнутых СМО *без ограничений на очередь* верны следующие формулы:

- коэффициент загрузки: $U = \rho$;
- пропускная способность: $\gamma = \lambda$.

Примечание. Эти формулы представляют собой частные случаи для $P_{отк} = 0$.

Вероятность простоя (P_0), вероятность отказа ($P_{отк}$) и средняя длина очереди (\bar{q}) рассчитываются по-разному в зависимости от типа СМО.

Экономическая оценка эффективности

1. Выручка от обслуживания заявок в СМО в течение времени T :

$$V = \gamma CT \quad (9.13)$$

где γ – пропускная способность СМО;

C – выручка от обслуживания одной заявки.

2. Затраты, связанные с обслуживанием заявок в СМО в течение времени T :

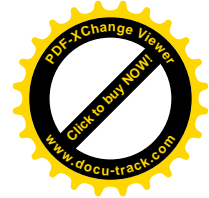
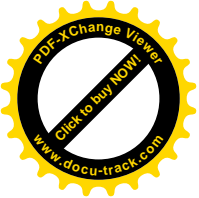
$$Z_{обс} = \gamma C_{обс} T \quad (9.14)$$

где $C_{обс}$ – затраты, связанные с обслуживанием одной заявки.

3. Затраты, связанные с эксплуатацией СМО в течение времени T :

$$Z_{эксп} = (\bar{S} C_{раб} + (m - \bar{S}) C_{пр}) T \quad (9.15)$$

где m – количество каналов в СМО;



\bar{S} – среднее число заявок на обслуживании (в каналах), или среднее число занятых каналов;

$C_{\text{раб}}$ – затраты, связанные с работой одного канала в течение единицы времени;

$C_{\text{пр}}$ – затраты, связанные с простоем одного канала в течение единицы времени.

4. Убытки, связанные с отказами в обслуживании за время T :

$$Z_{\text{отк}} = \lambda C_{\text{отк}} P_{\text{отк}} T \quad (9.16)$$

где λ – интенсивность потока заявок;

$C_{\text{отк}}$ – убытки, связанные с отказом в обслуживании одной заявки;

$P_{\text{отк}}$ – вероятность отказа.

5. Убытки за время T , связанные с пребыванием заявок в СМО (как в очереди, так и на обслуживании):

$$Z_{\text{пр}} = \bar{k} C_{\text{пр}} T \quad (9.17)$$

где \bar{k} – среднее число заявок в СМО;

$C_{\text{пр}}$ – убытки, связанные с пребыванием заявки в СМО в течение единицы времени.

ЛЕКЦИЯ 10. ОДНОКАНАЛЬНЫЕ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ (СМО)

Рассмотрим СМО типа $M/M/1$ с одним каналом обслуживания S и очередью к нему q (рис. 10.1).

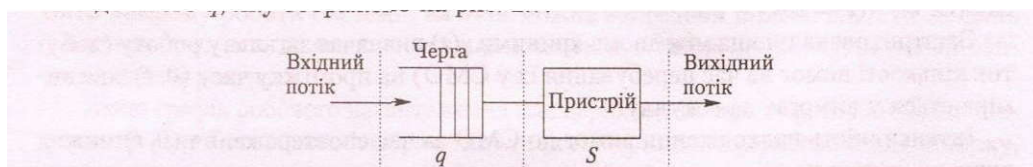


Рис. 10.1. СМО с одним каналом обслуживания

Оценка эффективности работы одноканальной СМО марковского типа без ограничений на очередь осуществляется с использованием следующих формул.

1. Вероятность простоя:

$$P_o = 1 - \rho \quad (17)$$

где ρ - нагрузка на СМО.

2. Вероятность отказа: $P_{отк} = 0$.

3. Вероятность обслуживания: $P_{обсл} = 1$.

4. Средняя длина очереди:

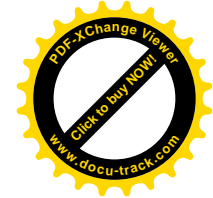
$$\bar{q} = \frac{\rho^2(v^2 + \varepsilon^2)}{2(1 - \rho)} \quad (10.1)$$

где v - коэффициент вариации интервалов времени между заявками;

ε - коэффициент вариации времени обслуживания.

Для пуассоновского потока заявок интервалы времени между их поступлением, а также продолжительность обслуживания - случайные величины, распределенные по экспоненциальному закону. Коэффициент вариации экспоненциального закона распределения равен 1, откуда формула (18) может быть представлена в упрощенном виде:

$$\bar{q} = \frac{\rho^2}{1 - \rho} \quad (10.2)$$



Примечание. Формула (10.1) позволяет точно рассчитать среднюю длину очереди только для СМО типа $M/M/1$. Для других СМО величина \bar{q} , найденная по формуле (10.1), является приближенной.

5. Вероятности пребывания в СМО k заявок:

$$P_k = \rho^k (1 - \rho), \quad k = 1, 2, \dots \quad (10.3)$$

1. Вероятность того, что время пребывания заявки в СМО превысит некоторую заданную величину T :

$$P(t < T) = e^{-\mu(1-\rho)T} \quad (10.4)$$

Основные характеристики работы одноканальных СМО сведены в таблицу

Характеристики	Одноканальная СМО	
	с ожиданием и ограниченной очередью	с ожиданием и неограниченной очередью
Количество каналов	m	
Интенсивность – среднее количество заявок, поступающих в СМО в единицу времени	λ или $\lambda = \frac{1}{(t_1 + t_2) \frac{1}{2}}$, если интервалы между моментами поступления деталей на обработку составляют от t_1 до t_2 минут	
Среднее время обслуживания заявки в канале	\bar{x}	
Интенсивность обслуживания заявок - среднее количество заявок, которое может быть обслужено одним каналом СМО в единицу времени	$\mu = \frac{l}{x}$	
Нагрузка на СМО	$\rho = \frac{\lambda}{\mu m}$	
Коэффициент вариации интервалов времени между заявками при равномерном распределении	-	$v = \frac{t_2 - t_1}{(t_1 + t_2)\sqrt{3}}$
Коэффициент вариации времени обслуживания	-	$\varepsilon = \frac{\sigma}{x}$ σ – стандартное отклонение случайной величины
Вероятность простоя	$P_0 = \frac{1 - \rho}{1 - \rho^{N+2}}, \rho \neq 1, n = 1..N$	$P_0 = 1 - \rho$
Вероятность отказа	$P_{отк} = \frac{\rho^{N+1}(1 - \rho)}{(1 - \rho^{N+2})}$	$P_{отк} = 0$
Вероятность обслуживания	$P_{обсл} = 1 - P_{отк}$	$P_{обсл} = 1 - P_{отк} = 1$
Средняя длина очереди	$\bar{q} = \frac{\rho^2(1 + \rho^N(N\rho - N - 1))}{(1 - \rho)(1 - \rho^{N+2})}$	$\bar{q} = \frac{\rho^2}{(1 - \rho)} \quad \bar{q} = \frac{\rho^2(v^2 + \varepsilon^2)}{2(1 - \rho)}$
Вероятности пребывания в СМО n заявок	$P_n = \rho^n P_0, n = 1, 2, \dots, N, \rho \neq 1$ $P_n = \frac{1}{N + 1}, \rho = 1$	$P_n = \rho^n P_0, n = 1, 2, \dots, N$

ЛЕКЦИЯ 11. МНОГОКАНАЛЬНЫЕ СМО

Анализ многоканальных СМО (рис. 11.1), в отличие от одноканальных, значительно сложнее. С помощью теории массового обслуживания можно получать аналитические зависимости в замкнутом виде для расчетов характеристик работы многоканальной СМО в стационарном режиме работы, однако лишь для модели типа $M/M/m$. Для СМО с другими законами распределения времени поступления и обслуживания требований используют численные методы.

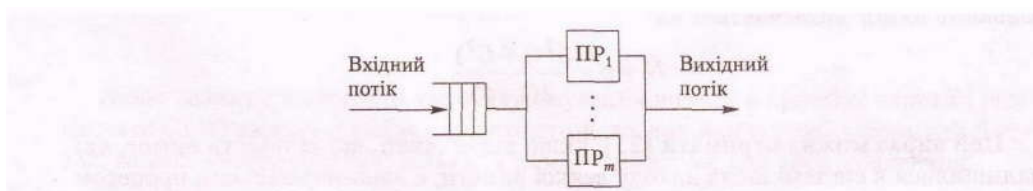


Рис. 11.1 Багатоканальна СМО

Для оценки эффективности работы марковской многоканальной СМО без ограничений на очередь используются следующие формулы.

1. Вероятность простоя:

$$P_0 = \left[\sum_{k=0}^{m-1} \frac{(m\rho)^k}{k!} + \frac{(m\rho)^m}{m!(1-\rho)} \right]^{-1} \quad (11.1)$$

где m – количество каналов (т.е. количество заявок, которые могут обслуживаться в СМО одновременно).

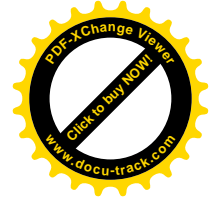
2. Вероятность отказа: $P_{отк} = 0$.

3. Вероятность обслуживания: $P_{обсл} = 1$.

4. Средняя длина очереди:

$$\bar{q} = \frac{\rho(m\rho)^m}{m!(1-\rho)^2} P_0 \quad (11.2)$$

5. Вероятности пребывания в СМО k заявок:



$$P_k = \begin{cases} \frac{(m\rho)^k}{k!} P_0, & k = 1, \dots, m \\ \frac{(m\rho)^k}{m^{k-m} m!} P_0, & k > m \end{cases} \quad (11.3)$$

$$(11.4)$$

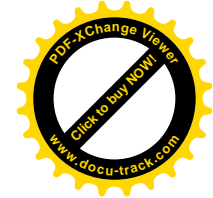
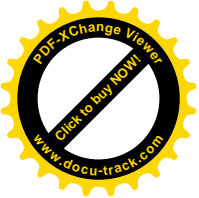
Формула (11.3) позволяет найти вероятности состояний СМО, при которых очередь отсутствует (количество заявок, обслуживаемых в СМО, не превышает количества каналов), а формула (11.4) – вероятности состояний при наличии очереди.

Примечание. Приведенные формулы могут применяться и для приближенного расчета характеристик немарковских многоканальных СМО (т.е. СМО типа $M/G/m$, $G/M/m$ или $G/G/m$).

При $\rho = 1$ расчет вероятности простоя и средней длины очереди выполняется по следующим формулам:

$$P_0 = \left[\sum_{i=0}^m \frac{(m\rho)^i}{i!} + \frac{n(m\rho)^m}{m!} \right]^{-1}, \quad (11.5)$$

$$\bar{q} = \frac{(m\rho)^m P_0}{m!} \cdot \frac{n(n+1)}{2}. \quad (11.6)$$



ЛЕКЦИЯ 12. МОДЕЛЬ ОБСЛУЖИВАНИЯ МАШИННОГО ПАРКА

Модель обслуживания машинного парка можно описать используя, теорию массового обслуживания, считая данный процесс замкнутой системой массового обслуживания.

Замкнутые СМО – это СМО с фиксированным количеством заявок, периодически требующих обслуживания.

Будем обозначать количество заявок в замкнутой СМО как N , а среднее время между окончанием обслуживания заявки и ее следующим обращением за обслуживанием (время между обращениями) – как T . Как и для других СМО, количество каналов будем обозначать как m , а среднее время обслуживания заявки – как \bar{x} .

Точный расчет характеристик замкнутых СМО возможен только в случае, если и время обслуживания заявки, и время между обращениями представляют собой случайные величины, распределенные по экспоненциальному закону.

Характеристики эффективности работы замкнутых СМО в основном те же, что и для разомкнутых СМО. Однако расчет характеристик для замкнутых и разомкнутых

СМО существенно различается.

Для расчета характеристик замкнутых СМО применяются следующие формулы.

Вероятность простоя:

$$P_0 = \left[1 + \sum_{j=1}^m \frac{N! (\bar{x}/T)^j}{(N-j)! j!} + \sum_{j=m+1}^N \frac{N! (\bar{x}/T)^j}{(N-j)! m! m^{j-m}} \right]^{-1} \quad (12.1)$$

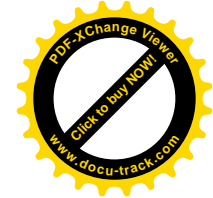
где P_0 – вероятность простоя;

N - количество заявок в замкнутой СМО;

T – время между обращениями в СМО;

m - количество каналов обслуживания;

Среднее число заявок в очереди – средняя длина очереди:



$$\bar{q} = P_0 \sum_{j=m+1}^N (j-m) \frac{N!(\bar{x}/T)^j}{(N-j)!m!m^{j-m}} \quad (12.2)$$

где \bar{q} – средняя длина очереди.

Среднее число заявок на обслуживании – среднее число занятых каналов:

$$\bar{S} = \begin{cases} 1 - P_0, & m = 1 \\ m - mP_0 - P_0 \sum_{j=1}^{m-1} (m-j) \frac{N!(\bar{x}/T)^j}{(n-j)!j!}, & m > 1 \end{cases} \quad (12.3)$$

где \bar{S} – среднее число заявок на обслуживании.

Среднее число заявок в СМО, т.е. на обслуживании и в очереди:

$$\bar{k} = \bar{q} + \bar{S} \quad (12.4)$$

где \bar{k} – среднее число заявок в СМО.

Среднее время пребывания заявки в СМО, т.е. в очереди и на обслуживании:

$$\bar{t} = \frac{\bar{k}T}{N - \bar{k}} \quad (12.5)$$

где \bar{t} – среднее время пребывания заявки в СМО.

Среднее время пребывания заявки в очереди – среднее время ожидания обслуживания:

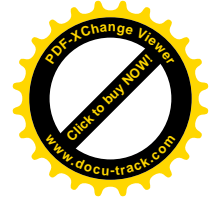
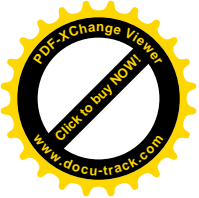
$$\bar{w} = \bar{t} + \bar{x} \quad (12.6)$$

где \bar{w} – среднее время ожидания обслуживания.

Коэффициент загрузки СМО. Эта величина показывает, какую часть от общего времени своей работы СМО выполняет обслуживание заявок, рассчитывается так:

$$U = \frac{\bar{S}}{m} \quad (12.7)$$

где U - Коэффициент загрузки СМО.



Пропускная способность – среднее количество заявок, обслуживаемых в единицу времени:

$$\gamma = \frac{\bar{S}}{\bar{x}} \quad (12.8)$$

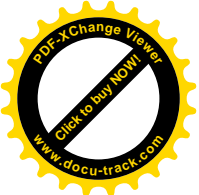
где γ – пропускная способность СМО.

Вероятности пребывания в СМО j заявок:

$$P_j = \begin{cases} \frac{N!(\bar{x}/T)^j P_0}{(N-j)!j!}, & j = 1, \dots, m \\ \frac{N!(\bar{x}/T)^j}{(N-j)!m!m^{j-m}} P_0, & j = m+1, \dots, N \end{cases} \quad (12.9)$$

где P_j – вероятности пребывания в СМО j заявок

Формула (10) позволяет найти вероятности состояний СМО, при которых очередь отсутствует (количество заявок, обслуживаемых в СМО, не превышает количества каналов), и – вероятности состояний при наличии очереди.



ЛЕКЦИЯ 13. СМО С ПРИОРИТЕТАМИ

При обслуживании с приоритетами более высокий приоритет должны иметь сигналы, требующие меньшего времени обработки. Требуется выбрать дисциплину обслуживания, обеспечивающую минимальное среднее время обработки всех сигналов.

В СМО с приоритетами каждой заявке назначается некоторый приоритет. Если в очереди находятся заявки с разными приоритетами, то первыми на обслуживание поступают заявки с более высоким приоритетом.

Для такой дисциплины обслуживания заявок точный расчет характеристик возможен только при следующих условиях:

поток заявок является пуассоновским;

СМО является одноканальной;

нет ограничений на очередь.

Другими словами, тип СМО – $M/M/1$ или $M/G/1$ без ограничений на очередь.

Приоритеты заявок могут быть относительными или абсолютными.

В СМО с *относительными приоритетами* обслуживается в первую очередь заявки с более высоким приоритетом. Обслуживание заявки доводится до конца, даже в случае поступления заявки с более высоким приоритетом.

В СМО с *абсолютными приоритетами* обслуживание заявки прерывается, если поступает заявка с более высоким приоритетом. Заявка, обслуживание которой было прервано, возвращается в очередь и поступает на дообслуживание только тогда, когда в очереди не останется ни одной заявки с более высоким приоритетом.

Пусть в СМО имеется R значений (уровней) приоритета. Будем обозначать номером 1 высший приоритет, а номером R – низший. Будем обозначать характеристики

СМО для заявок с i -м уровнем приоритета индексом, обозначающим приоритет (например, – среднее время пребывания в СМО заявок с первым уровнем приоритета).



Средние характеристики СМО для заявок всех уровней приоритета будем указывать без индексов (например, \bar{t} – среднее время пребывания в СМО всех заявок).

В расчетах характеристик СМО с приоритетами используются величины нагрузки на СМО, создаваемой заявками каждого уровня приоритета:

$$\rho_i = \frac{\lambda_i}{\mu_i}, \quad (13.1)$$

где λ_i – интенсивность потока заявок с i -м уровнем приоритета;

μ_i – интенсивность обслуживания заявок с i -м уровнем приоритета, определяемая как $\mu_i = \frac{1}{\bar{x}_i}$, где x_i – среднее время обслуживания заявок с i -м уровнем приоритета.

Нагрузка на СМО, создаваемая всеми заявками, определяется следующим образом:

$$\rho = \sum_{i=1}^R \rho_i, \quad (13.2)$$

Расчет характеристик СМО с приоритетами во многих случаях удобно начинать с вычисления среднего времени пребывания в очереди для заявок с различными уровнями приоритета. Для СМО с **относительными приоритетами** эти величины вычисляются следующим образом:

1) для заявок с высшим приоритетом (с приоритетом 1):

$$\bar{w}_1 = \frac{\sum_{j=1}^R \rho_j \bar{x}_j (1 + \varepsilon_j^2)}{2(1 - \rho_1)} \quad (13.3)$$

где ε_j – коэффициент вариации времени обслуживания заявок с j -м уровнем приоритета;

2) для заявок с приоритетами 2, 3, ..., R :

$$\bar{w}_i = \frac{\sum_{j=1}^R \rho_j \bar{x}_j (1 + \varepsilon_j^2)}{2(1 - \sum_{j=1}^{i-1} \rho_j)(1 - \sum_{j=1}^i \rho_j)}, \quad i = 2, \dots, R. \quad (13.4)$$



Для СМО с абсолютными приоритетами средние времена пребывания заявок в очереди рассчитываются по следующим формулам:

1) для заявок с высшим приоритетом (с приоритетом 1):

$$\bar{w}_1 = \frac{\rho_1 \bar{x}_1 (1 + \varepsilon_1^2)}{2(1 - \rho_1)} \quad (13.5)$$

2) для заявок с приоритетами 2, 3, ..., R:

$$\bar{w}_i = \frac{\bar{x}_i \sum_{j=1}^{i-1} \rho_j}{1 - \sum_{j=1}^{i-1} \rho_j} + \frac{\sum_{j=1}^i \rho_j \bar{x}_j (1 + \varepsilon_j^2)}{2(1 - \sum_{j=1}^{i-1} \rho_j)(1 - \sum_{j=1}^i \rho_j)}, \quad i = 2, \dots, R. \quad (13.6)$$

Другие характеристики СМО определяются по следующим формулам (для СМО как с относительными, так и с абсолютными приоритетами).

Среднее время пребывания заявки в очереди:

$$\bar{w} = \frac{\sum_{i=1}^R \lambda_i \bar{w}_i}{\lambda}, \quad (13.7)$$

или

$$\bar{w} = \sum_{i=1}^R P_i \bar{w}_i, \quad (13.8)$$

где λ – интенсивность потока всех заявок в СМО: $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_R$.

P_i – доля заявок с i -м уровнем приоритета в потоке заявок, поступающих в СМО, $P_1 + P_2 + \dots + P_R = 1$.

Среднее время пребывания заявки в СМО:

$$\bar{t}_i = \bar{w}_i + \bar{x}_i, \quad i = 1, \dots, R, \quad (13.9)$$

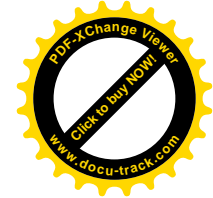
$$\bar{t} = \frac{\sum_{i=1}^R \lambda_i \bar{t}_i}{\lambda}, \quad (13.10)$$

или

$$\bar{t} = \sum_{i=1}^R P_i \bar{t}_i. \quad (13.11)$$

Среднее число заявок в СМО:

$$\bar{k}_i = \lambda_i \bar{t}_i, \quad i = 1, \dots, R, \quad (13.12)$$



$$\bar{k} = \sum_{i=1}^R \bar{k}_i . \quad (13.13)$$

Среднее число заявок на обслуживании (среднее число занятых каналов):

$$\bar{S}_i = \rho_i , i = 1, \dots, R, \quad (13.14)$$

$$\bar{S} = \rho . \quad (13.15)$$

Среднее число заявок в очереди:

$$\bar{q}_i = \bar{k}_i - \bar{S}_i , i = 1, \dots, R, \quad (13.16)$$

$$\bar{q} = \sum_{i=1}^R \bar{q}_i . \quad (13.17)$$

Пропускная способность СМО:

$$\gamma_i = \lambda_i , i = 1, \dots, R, \quad (13.18)$$

$$\gamma = \lambda . \quad (13.19)$$

Вероятность простоя СМО:

$$P_0 = 1 - \rho . \quad (13.20)$$

Коэффициент загрузки СМО:

$$U_i = \rho_i , i = 1, \dots, R, \quad (13.21)$$



ЛИТЕРАТУРА

1. Качала В.В. Основы теории систем и системного анализа. Учебное пособие для вузов. – М.: Горячая линия – Телеком, 2007. – 216 с.
2. Теория систем и системный анализ в управлении организациями: Справочник. / Под ред. В.Н. Волковой и А.А. Емельянова. – М.: Финансы и статистика, 2009 – 846 с.
3. Томашевський В.М. Моделювання систем. – К.: Видавнича група ВНУ, 2005. – 352 с.
4. Кобринский Н.Е., Майминас Е.З., Смирнов А.Д. Экономическая кибернетика. – СПб.: Питер, 2006. – 408 с.
5. Советов Б.Я., Яковлев С.А. Моделирование систем: Учебное пособие для вузов. 3-е изд., перераб. и доп. – М.: Высшая школа, 2006. – 343 с.
6. Доугерти К. Введение в эконометрику : Учебник. 3-е изд. / Пер. с англ. – М.: ИНФРА-М, 2009. – XIV, 465 с.
7. Смородинский С.С. Оптимизация решений на основе методов и моделей математического программирования: Учеб. пособие по курсу «Системный анализ и исследование операций» для студ. спец. «Автоматизир. системы обраб. информ.» дневн. и дистанц.форм обуч. / С.С. Смородинский, Н.В. Батин. – Мн.: БГУИР, 2005. – 136 с.
8. Бережная Е.В., Бережной В.И. Математические методы моделирования экономических систем: Учеб. пособие. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2006. – 432 с.
9. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – М.: ЮНИТИ – ДАНА, 2007. – 573 с.
10. Сурмин Ю.П. Теория систем и системный анализ: Учебное пособие. – К.: МАУП, 2005. – 368 с.
11. Экономическая кибернетика/ Под ред. Ю.Г. Лысенко. – Донецк: ООО «Юго-Восток Лтд», 2007. – 287с.
12. Шарапов О.Д., Дербенцев В.Д., Семьонов Д.С. Економічна кібернетика. – К.: КНЕУ, 2005. – 231 с.
13. Горелова Г.В., Кацко И.А. Теория вероятностей и математическая статистика в примерах и задачах с применением Excel: Учебное пособие для вузов, - 2-е изд. испр.и доп. – Ростов н/Д.: Феникс, 2006. – 475с.
14. Макаркина А.В., Аносов В.Л.. Учебно-методическое пособие по курсу «Прогнозирование социально-экономических процессов» для студентов экономических специальностей. – Краматорск: ДГМА, 2006. – 108 с.
15. Горбань С.Ф., Снижко Н.В. Теория вероятностей и математическая статистика: Учебное пособие. – К.: МАУП, 2006. – 168 с.
16. Глаголев В.В. Основы теории систем. Методы дискретной математики: Учебное пособие. – Тула, 2005. – 90 с.
17. Погостинская Н.Н., Погостинский Ю.А. Системный анализ финансовой отчетности: Учебное пособие. – СПб., 2008. – 96 с.