МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

І. І. Обод, Г. Е. Заволодько, І. В. Свид

# МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ СИСТЕМ

Навчальний посібник
для студентів спеціальностей
«Комп'ютерна інженерія»,
«Комп'ютерні науки та інформаційні технології»

За редакцією І. І. Обода

Харків 2019

УДК 681.5.015
О-21

Рецензенти:

*В. М. Карташов,* д-р техн. наук, проф., завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Харківський національний університет радіоелектроніки;

*С. Ю. Леонов,* д-р техн. наук, проф. кафедри обчислювальної техніки та програмування, Національний технічний університет «Харківський політехнічний інститут»

Автори: І. І. Обод, д-р техн. наук, проф.; Г. Е. Заволодько, к.т.н.; І. В. Свид, к.т.н., доц.

**Обод І. І.**

О-21 Математичне моделювання систем: навчальний посібник/ І. І. Обод, Г. Е. Заволодько, І. В. Свид. – Харків : Друкарня МАДРИД, 2019. – 268 с.

ISBN 978-617-7683-93-2

У навчальному посібнику викладено та системно наведено: методологічні основи, принципи дослідження, опис та моделювання систем, на основі яких розглянуто методи та засоби математичного моделювання систем з урахуванням реальних умов функціонування.

Призначено для студентів, які навчаються за напрямом підготовки «Інформаційні технології».

Іл. 108. Табл. 18. Бібліогр. 17 назв.

УДК 681.5.015

ISBN 978-617-7683-93-2

# ЗМІСТ

# ВСТУП

Моделювання (в широкому сенсі) є основним методом досліджень у всіх сферах знань і науково обґрунтованим методом оцінок характеристик складних систем, що використовуються для прийняття рішень у різних сферах інженерної діяльності. Існуючі і проектовані системи можна ефективно досліджувати за допомогою математичних моделей (аналітичних, статистичних та імітаційних), що реалізуються на сучасних ЕОМ, які в цьому випадку виступають як інструмент експериментатора з моделлю системи.

Сьогодні не можна назвати сферу людської діяльності, в якій тією чи іншою мірою не використовувалися б методи моделювання. Особливо це стосується сфери керування різними системами, де основними є процеси прийняття рішень на основі отриманої інформації.

Також неможливо уявити собі сучасну науку без широкого вживання математичного моделювання, суть якого полягає в заміні вихідного об'єкта математичною моделлю і подальшому вивченні моделі за допомогою обчислювально-логічних алгоритмів, що реалізовуються на комп'ютерах. Цей метод поєднує в собі достоїнства як теорії, так і експерименту, оскільки робота не з самим об'єктом (явищем, процесом), а з його моделлю дає можливість відносно швидко і без істотних витрат досліджувати його властивості і поведінку в різних ситуаціях. В той же час обчислювальні експерименти з моделями об'єктів дозволяють, спираючись на потужність сучасних обчислювальних методів і технічних засобів інформатики, детально і глибоко вивчати об'єкти в достатній повноті, недоступній чисто теоретичним підходам.

Моделювання поділяють на фізичне і математичне. Перевага фізичного моделювання перед натурним експериментом очевидна, завдяки економії часу і коштів, але воно має більш обмежене застосування порівняно з математичним моделюванням.

Математичне моделювання розуміють як спосіб дослідження різних процесів шляхом вивчення явищ, що мають різний фізичний зміст, але описуваних однаковими математичними співвідношеннями. Важливою є та обставина, що при вивченні будь-якого процесу методом математичного моделювання необхідно, в першу чергу, побудувати його математичний опис, тобто математичну модель. Математичну модель реальної системи розуміють як сукупність співвідношень (наприклад, формул, рівнянь, нерівностей, логічних умов, операторів і т. п.), що визначають характеристики станів системи (а через них і вихідні характеристики) залежно від її параметрів, вхідних характеристик, початкових умов і часу.

Математична модель реальної системи є тим абстрактним формально описаним об'єктом, вивчення якого можливе математичними методами, в тому числі і за допомогою математичного моделювання.

Зазначене вище визначило структуру цього навчального посібника, що складається з трьох головних частин, присвячених основам статистичного, імітаційного й аналітичного моделювання.

Перший розділ навчального посібника присвячено методологічним основам дослідження систем, де подано основні поняття теорії систем, визначення та властивості систем, методологія системного підходу, класифікація систем, принципи і закономірності дослідження та моделювання систем.

Другий розділ навчального посібника присвячено методологічним основам опису систем, тобто функціональному, морфологічному (структурному) та інформаційному опису систем, надано формування загального і детального подання системи та розглянуто показники і критерії оцінки систем.

У третьому розділі описано методологічні основи моделювання, визначені поняття моделі та моделювання, стисло надано теорію моделювання та описаний системний підхід до нього, викладено принципи побудови математичних моделей, дана класифікація моделей та основні етапи математичного моделювання.

Четвертий розділ акцентує увагу на моделюванні систем у вигляді «чорного ящика». Розглянуто модель «чорного ящика» та приклади моделей деяких систем на його основі.

У п'ятому розділі наведено статистичне моделювання систем, де подані загальні відомості про статистичне моделювання, загальну схему методу Монте-Карло та основні вимоги до складових методу моделювання систем, що вивчається.

Розділ шостий присвячено розгляду імітаційного моделювання систем та на прикладі моделей систем масового обслуговування наведено методологію побудови моделюючих алгоритмів зазначених систем і приклад імітаційного моделювання системи масового обслуговування.

Сьомий розділ присвячено аналітичному моделюванню систем масового обслуговування.

Матеріал, який наведено у посібнику, буде корисним студентам та аспірантам, що навчаються як за спеціальностю «Комп'ютерна інженерія», так і «Телекомунікаційні та інфокомунікаційні технології».

# 1. МЕТОДОЛОГІЧНІ ОСНОВИ ДОСЛІДЖЕННЯ СИСТЕМ

Створення складних систем, виробничих комплексів, систем керування цими комплексами часто вимагає використання знань про кількісні і якісні закономірності, що властиві розглянутим системам. Головними завданнями при цьому є загальносистемні питання, що включають визначення структури, організацію взаємозв'язку між елементами, взаємодію з зовнішнім середовищем, керування функціонуванням як усієї системи, так і її окремих елементів.

Наведений перелік завдань, що є далеко не повним, належить до завдань загальної теорії систем (ЗТС), центральними поняттями якої є: **система** та **системний підхід**.

Предметом вивчення ЗТС є система, незалежно від її природи, організації, способу існування і способу описання.

Метою розгляду системи є вирішення завдань аналізу, керування і проектування.

Під час розгляду реальної системи доводиться стикатися з сукупністю задач, розв'язання яких може бути під силу лише колективу професіоналів різного профілю. До таких належать задачі, починаючи з виділення системи з середовища, її формального опису, взаємодії із зовнішнім середовищем, вибору або розробки оптимального алгоритму керування, оптимального проектування, технічних засобів керування і тому подібне, закінчуючи кадрами й організацією колективу для вирішення цих проблем. Для вирішення названих задач системний аналіз залучає широкий спектр різних наук і різні сфери практичної діяльності. При цьому він надає велике значення методичним аспектам будь-якого системного дослідження.

## 1.1. Основні поняття загальної теорії систем

Загальна теорія систем – наукова дисципліна, що вивчає фундаментальні поняття та аспекти систем. Вона вивчає різні явища, відволікаючись від їх конкретної природи і ґрунтуючись лише на формальних взаємозв'язках між різними складовими їх чинниками і на характері їх зміни під впливом зовнішніх умов, при цьому результати всіх спостережень пояснюються лише взаємодією їх компонентів, наприклад, характером їх організації та функціонування, а не за допомогою безпосереднього звернення до природи залучених в явища механізмів (якби вони були фізичними, біологічними, екологічними, соціологічними, або концептуальними).

Для ЗТС об'єктом дослідження є не «фізична реальність», а «система», тобто абстрактний формальний взаємозв'язок між основними ознаками і властивостями.

При системному підході об'єкт дослідження подається як система. Саме поняття «система» може бути віднесено до одного з методологічних понять, оскільки розгляд об'єкта досліджується як система або відмова від такого розгляду залежить від завдання дослідження і самого дослідника.

Існує багато визначень ***системи***.

1. ***Система*** є комплексом елементів, що знаходиться у взаємодії.
2. ***Система*** – це множина об'єктів разом з відношенням цих об'єктів.
3. ***Система*** – множина елементів, що знаходяться у відношеннях або зв'язках один з одним, утворює цілісність або органічну єдність.

Терміни «відношення» і «взаємодія» використовуються в самому широкому сенсі, включаючи весь набір родинних понять, таких як обмеження, структура, організаційний зв'язок, з'єднання, залежність і т.д.

Таким чином, система $S$ являє собою упорядковану пару $S = (A, R)$, де $A$ – множина елементів; $R$ – множина відношень між $A$.

***Система*** – це повний, цілісний набір елементів (компонентів), взаємозалежних і взаємодіючих між собою так, щоб могла реалізуватися функція системи.

Дослідження об'єкта як системи передбачає використання ряду системних уявлень (категорій), серед яких основними є:

1. Структурне уявлення, пов'язане з виділенням елементів системи і зв'язків між ними.
2. Функціональне уявлення систем – виділення сукупності функцій (цілеспрямованих дій) системи та її компонентів, спрямоване на досягнення певної мети.
3. Макроскопічне уявлення – розуміння системи як нероздільного цілого, що взаємодіє із зовнішнім середовищем.
4. Мікроскопічне уявлення, засноване на розгляді системи як сукупності взаємопов'язаних елементів. Воно передбачає розкриття структури системи.
5. Ієрархічне уявлення, засноване на понятті підсистеми, одержуваному при розкладанні (декомпозиції) системи, що має системні властивості, які слід відрізняти від її елемента – неподільного на більш дрібні частини (з точки зору розв'язуваної задачі). Система може бути зображена як сукупність підсистем

різних рівнів, складова системної ієрархії, що замикається знизу тільки елементами.

6. Процесуальне уявлення, що передбачає розуміння системного об'єкта як динамічного об'єкта, що характеризується послідовністю його станів у часі.

Розглянемо визначення інших понять, тісно пов'язаних із системою та її характеристиками.

***Об'єкт.*** Об'єктом пізнання є частина реального світу, що виділяється і сприймається як єдине ціле протягом тривалого часу. Об'єкт може бути матеріальним і абстрактним, природним і штучним. Реально об'єкт володіє нескінченним набором властивостей різної природи. Практично в процесі пізнання взаємодія здійснюється з обмеженою множиною властивостей, що знаходяться в межах можливості їх сприйняття і необхідності для мети пізнання. Тому система як образ об'єкта задається на скінченній множині відібраних для спостереження властивостей.

***Зовнішнє середовище.*** Поняття «система» виникає там і тоді, де і коли ми матеріально або умоглядно проводимо замкнуту межу між необмеженою або деякою обмеженою множиною елементів. Ті елементи з їх відповідною взаємною обумовленістю, які потрапляють всередину, утворюють систему.

Ті елементи, які залишилися за межами, утворюють множину, що називається в теорії систем «системним оточенням», або просто «оточенням», або «зовнішнім середовищем».

З цих міркувань випливає, що немислимо розглядати систему без її зовнішнього середовища. Система формує і проявляє свої властивості в процесі взаємодії з оточенням, будучи при цьому провідним компонентом цього впливу.

Залежно від впливу на оточення і характеру взаємодії з іншими системами функції систем можна розташувати за зростанням таким чином:

- пасивне існування;
- матеріал для інших систем;
- обслуговування систем більш високого порядку;
- протистояння іншим системам (виживання);
- поглинання інших систем (експансія);
- перетворення інших систем і середовищ (активна роль).

Будь-яка система може розглядатися, з одного боку, як підсистема більш високого порядку (надсистеми), а з іншого – як надсистема системи більш низького порядку (підсистема). Наприклад, система «виробничий цех» входить як підсистема в систему більш високого рангу – «фірма». У свою чергу, надсистема «фірма» може бути підсистемою «корпорації».

Зазвичай як підсистема фігурують більш-менш самостійні частини систем, які виділяються за певними ознаками, що володіють відносною самостійністю, певним степенем вільності.

***Компонент*** – будь-яка частина системи, що вступає в певні відношення з іншими частинами (підсистемами, елементами).

***Елементом*** системи є частина системи з однозначно визначеними властивостями, що виконують певні функції і не підлягають подальшому розбиттю в рамках розв'язуваної задачі (з точки зору дослідника).

Поняття «елемент», «підсистема», «система» – взаємоперетворювані, система може розглядатися як елемент системи вищого порядку (метасистема), а елемент, при поглибленому аналізі – як система. Та обставина, що будь-яка підсистема є одночасно і відносно самостійною системою, приводить до двох аспектів вивчення систем: на макро- та мікрорівнях.

При вивченні на макрорівні основна увага приділяється взаємодії системи із зовнішнім середовищем. Причому системи більш високого рівня можна розглядати як частину зовнішнього середовища. При такому підході головними чинниками є цільова функція системи (мета), умови її функціонування. При цьому елементи системи вивчаються з точки зору організації їх в єдине ціле, впливу на функції системи в цілому.

На мікрорівні основними стають внутрішні характеристики системи, характер взаємодії елементів між собою, їх властивості та умови функціонування.

***Структура системи.*** Структуру системи розуміють як стійку множину відносин, яка зберігається тривалий час незмінною, принаймні протягом інтервалу спостереження. Структура системи випереджає певний рівень складності за складом відносин на множини елементів системи або, що еквівалентно, рівень різноманітності проявів об'єкта.

***Зв'язки*** – це елементи, що здійснюють безпосередню взаємодію між елементами (або підсистемами) системи, а також з елементами і підсистемами оточення.

Зв'язок – одне з фундаментальних понять у системному підході. Система як єдине ціле існує саме завдяки наявності зв'язків між її елементами, тобто, іншими словами, зв'язки виражають закони функціонування системи. Зв'язки розрізняють за характером взаємозв'язку як прямі і зворотні, а за видом прояву (опису) – як детерміновані та імовірнісні.

***Прямі зв'язки*** призначені для заданої функціональної передачі речовини, енергії, інформації або їх комбінацій – від одного елемента до іншого в напрямку основного процесу.

***Зворотні зв'язки*** здебільшого виконують інформуючі функції, відображаючи зміну стану системи в результаті керуючого впливу на неї. Відкриття принципу зворотного зв'язку стало визначною подією в розвитку техніки і мало виключно важливі наслідки. Процеси керування, адаптації, саморегулювання, самоорганізації, розвитку не можливі без використання зворотних зв'язків.

За допомогою зворотного зв'язку сигнал (інформація) з виходу системи (об'єкта керування) передається до органа керування. Тут цей сигнал, що містить інформацію про роботу, виконану об'єктом керування, порівнюється з сигналом, що задає зміст і обсяг роботи (наприклад, план). У разі виникнення неузгодженості між фактичним і плановим станом роботи вживаються заходи щодо його усунення.

Основними функціями зворотного зв'язку є:

1) протидія тому, що робить сама система, коли вона виходить за встановлені межі (наприклад, реагування на зниження якості);

2) компенсація збурень і підтримка стану стійкої рівноваги системи (наприклад, неполадки в роботі обладнання);

3) синтезування зовнішніх і внутрішніх збурень, які прагнуть вивести систему зі стану стійкої рівноваги, зведення цих збурень до відхилень однієї або декількох керованих величин (наприклад, вироблення керуючих команд та одночасна поява нового конкурента і зниження якості продукції, що випускається);

4) вироблення керуючих впливів на об'єкт керування згідно з законом, що погано формалізується. Наприклад, встановлення більш високої ціни на енергоносії викликає в діяльності різних організацій складні зміни, міняють кінцеві результати їх функціонування, вимагають внесення змін у виробничо-господарський процес шляхом впливів, які неможливо описати за допомогою аналітичних виразів.

Порушення зворотних зв'язків у соціально-економічних системах з різних причин веде до тяжких наслідків. Окремі локальні системи втрачають здатність до еволюції і тонкого сприйняття. Намічаються нові тенденції, перспективний розвиток та науково обґрунтоване прогнозування своєї діяльності на тривалий період часу, ефективного пристосування до постійно мінливих умов зовнішнього середовища.

Детермінований (жорсткий) зв'язок, як правило, однозначно визначає причину і наслідок, дає чітко обумовлену формулу взаємодії елементів. Імовірнісний (гнучкий) зв'язок визначає неявну, непряму залежність між елементами системи. Теорія ймовірності пропонує математичний апарат для дослідження цих зв'язків, що називається «кореляційними залежностями».

***Критерії*** – ознаки, за якими проводиться оцінка відповідності функціонування системи бажаного результату (мети) при заданих обмеженнях.

***Ефективність системи*** – співвідношення між заданим (цільовим) показником результату функціонування системи і фактично реалізованим.

***Функціонування*** будь-якої довільно обраної системи полягає у переробці вхідних (відомих) параметрів і відомих параметрів впливу навколишнього середовища в значення вихідних (невідомих) параметрів з урахуванням факторів зворотного зв'язку.

***Вхід*** – все, що змінюється при перебігу процесу (функціонування) системи.

***Вихід*** – результат кінцевого стану процесу.

Система здійснює свій зв'язок із середовищем таким чином. Вхід цієї системи є в той же час виходом попередньої, а вихід цієї системи – входом наступної. Таким чином, вхід і вихід розташовуються на межі системи і виконують одночасно функції входу і виходу попередніх і наступних систем.

Керування системою пов'язане з поняттями прямого і зворотного зв'язку, обмеженнями.

***Зворотний зв'язок*** призначений для виконання таких операцій:

- порівняння даних на вході з результатами на виході, з виявленням їх якісно-кількісної відмінності;
- оцінення змісту і сенсу відмінності;
- вироблення рішення, що випливає з відмінності;
- вплив на введення.

***Обмеження*** забезпечує відповідність між виходом системи і вимогою до нього як до входу в наступну систему – споживач. Якщо задана вимога не виконується, обмеження не пропускає його через себе. Обмеження, таким чином, відіграє роль узгодження функціонування цієї системи з цілями (потребами) споживача.

Визначення функціонування системи пов'язане з поняттям «проблемної ситуації», яка виникає, якщо є відмінність між необхідним (бажаним) виходом і існуючим (реальним) входом.

***Станом системи*** називається сукупність істотних властивостей, якими система володіє в кожен момент часу.

## 1.2. Властивості систем

***Властивість*** розуміють як бік об’єкта, що зумовлює його відмінність від інших об’єктів або схожість з ними і виявляється при взаємодії з іншими об’єктами.

***Характеристика*** – те, що відображає деяку властивість системи.

З визначення «системи» випливає, що головною властивістю системи є цілісність, єдність, що досягається за допомогою певних взаємозв’язків і взаємодій елементів системи, крім виникнення нових властивостей, якими елементи системи не володіють. Це властивість емерджентності (від анг. emerge – виникати, з’являтися).

1. Емерджентність – ступінь незвідності властивостей системи до властивостей елементів, з яких вона складається.
2. Емерджентність – властивість систем, що обумовлює появу нових властивостей і якостей, не властивих елементам, що входять до складу системи.

Емерджентність – принцип протилежний редукціонізму, згідно з яким ціле можна вивчати, розчленувавши його на частини і потім, визначаючи їх властивості, визначити властивості цілого.

Властивості емерджентності близькі до властивості цілісності системи, проте їх не можна ототожнювати.

***Цілісність*** системи означає, що кожен елемент системи робить внесок у реалізацію цільової функції системи.

Цілісність і емерджентність – інтегративні властивості системи. Наявність інтеграційних властивостей є однією з найважливіших рис системи. Цілісність виявляється в тому, що система володіє власною закономірністю функціональності, своєю метою.

**Організованість** – складна властивість систем, які полягають у наявності структури та функціонування (поведінки). Неодмінною належністю систем є їх компоненти, саме ті структурні утворення, з яких складається ціле і без чого воно неможливе.

**Функціональність** – це прояв певних властивостей (функцій) при взаємодії із зовнішнім середовищем. Тут же визначається мета (призначення системи) як бажаний кінцевий результат.

**Структурність** – це впорядкованість системи, певний набір і розташування елементів зі зв’язками між ними. Між функцією і структурою системи існує взаємозв’язок, як між філософськими категоріями – зміст і форма. Зміна змісту (функцій) спричиняє зміну форми (структури), але і навпаки.

Важливою властивістю системи є наявність поведінки – дії, змін, функціонування та ін.

Вважається, що ця поведінка системи пов'язана з середовищем (оточенням), тобто з іншими системами, з якими вона входить у контакт або вступає в певні відношення.

Процес цілеспрямованої зміни в часі стану системи називається **поведінкою**. На відміну від керування, коли зміна стану системи досягається за рахунок зовнішніх впливів, поведінка реалізується виключно самою системою, виходячи з власних цілей.

Поведінка кожної системи пояснюється структурою систем нижчого порядку, з яких складається ця система, і наявністю ознак рівноваги (гомеостазу). Відповідно за ознакою рівноваги система має певний стан (стану), який є для неї кращим. Тому поведінка систем описується в термінах відновлення цих станів, коли вони порушуються в результаті зміни навколишнього середовища.

Ще однією властивістю є властивість зростання (розвитку). Розвиток можна розглядати як складову частину поведінки (при цьому найважливішим).

Одним із первинних, а отже, основоположних атрибутів системного підходу є неприпустимість розгляду об'єкта поза його розвитком, який розуміється як необоротна, спрямована, закономірна зміна матерії і свідомості. В результаті виникає нова якість або стан об'єкта. Ототожнення (може бути і не зовсім строге) термінів «розвиток» і «рух» дозволяє висловитися в такому сенсі, що поза розвитком немислиме існування матерії, в цьому випадку – системи. Наївно уявляти собі розвиток, що відбувається стихійно. В множині процесів, що здаються на перший погляд чимось на зразок броунівського (випадкового, хаотичного) руху, при пильній увазі і вивченні спочатку ніби проявляються контури тенденцій, а потім і досить стійкі закономірності. Ці закономірності за природою своєю діють об'єктивно, тобто не залежать від того, чи бажаємо ми їх прояву, чи ні. Незнання законів і закономірностей розвитку – це блукання в темряві.

Поведінка системи визначається характером реакції на зовнішні впливи.

Фундаментальною властивістю систем є **стійкість**, тобто здатність системи протистояти зовнішнім впливам. Від неї залежить тривалість життя системи.

Прості системи мають пасивні форми стійкості: міцність, збалансованість, регульованість, гомеостаз. А для складних визначальними є активні форми: надійність, живучість і адаптованість.

Якщо перераховані форми стійкості простих систем (крім міцності) стосуються їхньої поведінки, то визначальна форма стійкості складних систем має переважно структурний характер.

**Надійність** – властивість збереження структури систем, незважаючи на загибель окремих її елементів за допомогою їх заміни або дублювання, а **живучість** – як активне подавлення шкідливих якостей. Таким чином, надійність є більш пасивною формою, ніж живучість.

**Адаптованість** – властивість змінювати поведінку або структуру з метою збереження, поліпшення або набуття нових якостей в умовах зміни зовнішнього середовища. Обов'язковою умовою можливості адаптації є наявність зворотних зв'язків.

Будь-яка реальна система існує в середовищі. Зв'язок між ними буває настільки тісним, що визначати межу між ними стає складно. Тому виділення системи із середовища пов'язано з тим або іншим ступенем ідеалізації.

Можна виділити два аспекти взаємодії: у багатьох випадках набуває характеру обміну між системою і середовищем (речовиною, енергією, інформацією); середовище зазвичай є джерелом невизначеності для систем.

Вплив середовища може бути пасивним або активним (антоганістічно, цілеспрямовано протидіє системі).

Тому в загальному випадку середовище слід розглядати не тільки як байдуже, а й як антагоністичне по відношенню до досліджуваної системи.

## 1.3. Системний підхід

***Системний підхід*** – це напрям дослідження об'єкта (системи) з різних боків, комплексно, на відміну від тих, що раніше застосовувалися (фізичних, структурних і так далі).

Методологія системного підходу при вирішенні завдань аналізу систем зводиться до того, що дослідження об'єкта орієнтуються на розкриття його інтеграційних якостей, на виявлення різноманітних зв'язків і механізмів, що забезпечують ці якості.

***Методологія системного підходу при вирішенні завдань проектування і синтезу систем*** полягає в такому. Завдання проектування системи поділяється на підзадачі проектування її елементів. Причому кожен з елементів повинен розглядатися не сам по собі, а у взаємодії з іншими елементами. Вирішення підзадач повинне відбуватися за умови забезпечення інтеграційних якостей функціонування всієї системи. Для виконання цієї вимоги необхідний єдиний ідеологічний і організаційний план проектування, що пов'язує всі фази в цілому, починаючи від дослідницького опрацювання до фази виготовлення та експлуатації. Основні межі методики проектування – системність і оптимізаційність, використання імітаційного моделювання й обчислювальної техніки. Зазвичай завдання проектування на сучасному рівні розвитку науки й обчислювальної техніки найчастіше здійснюється як багаторазове вирішення завдання аналізу безлічі варіантів проекту системи.

Суть системного підходу можна більш чітко описати за допомогою формалізованої структури, яка може бути застосована на практиці при розв'язанні завдань аналізу, синтезу і проектування:

$$S =< G, W, M, Q, Str(Org), ier, P, R, a, E, B, I, C >,$$

де $S$ – сукупність методологічних вимог системного підходу;

$G$ – формулювання мети проектування, синтезу системи або її виявлення при розв'язанні задачі аналізу;

$W$ – визначення інтеграційних якостей системи як цілого і (або) методів їх встановлення;

$M$ – поділ системи на безліч її складових підсистем;

$Q$ – встановлення мети функціонування властивостей кожної підсистеми і вивчення утворення механізму забезпечення мети системи як цілого і її інтеграційних властивостей;

$Str(Org)$ – аналіз структури (організації) системи, вивчення її впливу на інтеграційні якості системи в цілому;

$Ier$ – визначення рівня ієрархії цієї системи та її підсистем в ієрархічній структурі систем, куди входить ця система;

$P, R, a$ – вплив властивостей ($P$) системи на інші системи, а також виявлення відношень ($R$) зв'язків ($a$) цієї системи та її підсистем з іншими системами (зовнішнім середовищем);

$E$ – вивчення впливу зовнішнього середовища на систему;

$B$ – аналіз процесу функціонування системи, у тому числі її розвитку;

$I$ – аналіз інформаційних потоків, що циркулюють у системі і надходять ззовні з метою керування нею;

$C$ – опис принципів керування і процесу керування системою.

Наведена структура алгоритму системного підходу не є єдиною. Вони досить численні, проте принципових відмінностей немає, відмінності виявляються лише в деталях. Підкреслимо також, що на практиці використання алгоритму системного підходу можливий циклічний, ітераційний характер його вживання як у цілому, так і щодо окремих його етапів.

Системний підхід відрізняється від класичного (або індуктивного) підходу. Останній розглядає систему шляхом переходу від окремого до загального і синтезує систему шляхом злиття її компонентів, що розробляються окремо. На відміну від цього системний підхід передбачає послідовний перехід від загального до окремого, коли в основі розгляду є мета.

## 1.4. Класифікація систем

***Класифікацією*** називається розбиття на класи за найбільш істотними ознаками. Клас розуміється як сукупність об'єктів, що володіють деякими ознаками спільності. Ознака (чи сукупність ознак) є підставою (критерієм) класифікації.

Важливо зрозуміти, що класифікація — це лише модель реальності, тому до неї треба так і ставитись, не вимагаючи від неї абсолютної повноти. Ще необхідно підкреслити відносність будь-яких класифікацій.

Сама класифікація виступає як інструмент системного аналізу. За її допомогою структурується об'єкт (проблема) дослідження, а побудована класифікація є моделлю цього об'єкта.

Повної класифікації систем сьогодні немає, більш того, не вироблені остаточно її принципи. Різні автори пропонують різні принципи класифікації, а схожим по суті — дають різні назви. Деяка класифікація систем наведена на рис. 1.1, а критерії класифікації – в табл. 1.1.

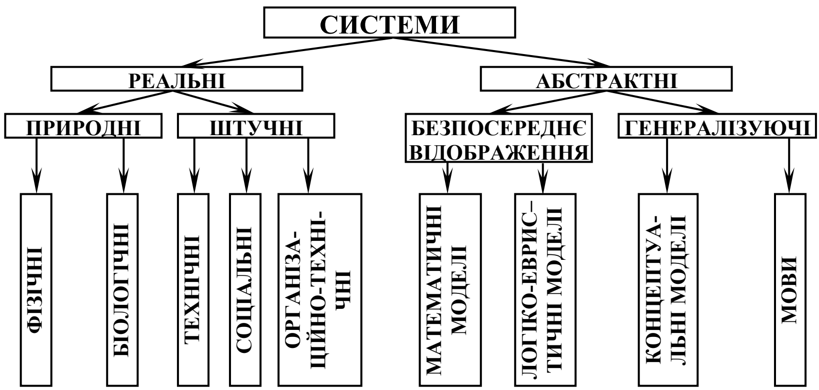


Рис. 1.1. Класифікація систем

Система може бути охарактеризована одним або декількома ознаками і відповідно їй може бути знайдено місце в різних класифікаціях, кожна з яких може бути корисною при виборі методології дослідження. Зазвичай мета класифікації – обмежити вибір підходів до відображення систем, виробити мову опису, відповіді для певного класу.

За змістом розрізняють системи **реальні** (**матеріальні**), об'єктивно існуючі і **абстрактні** (концептуальні, **ідеальні**), що є продуктом мислення.

Організаційна система, для ефективного функціонування якої істотним фактором є спосіб організації взаємодії людей з технічною підсистемою, називається людино-машинною системою.

Таблиця 1.1 – Критерії класифікації систем

| **Підстава (критерій) класифікації** | **Класи систем** |
|---|---|
| За взаємодією із зовнішнім середовищем | відкриті<br>закриті<br>комбіновані |
| За структурою | прості<br>складні<br>великі |
| За характером функцій | спеціалізовані<br>багатофункціональні (універсальні) |
| За характером розвитку | стабільні<br>такі, що розвиваються |
| За ступенем організованості | добре організовані<br>погано організовані (дифузні) |
| За складністю поведінки | автоматичні<br>вирішальні<br>такі, що самоорганізуються<br>такі, що передбачають<br>такі, що перетворюються |
| За характером зв'язку між елементами | детерміновані<br>стохастичні |
| За характером структури керування | централізовані<br>децентралізовані |
| За призначенням | такі, що виробляють<br>такі, що кирують<br>такі, що обслуговують |

Приклади людино-машинних систем: автомобіль – водій; літак – льотчик; ЕОМ – користувач і т.д.

Таким чином, технічні системи розуміють як єдину конструктивну сукупність взаємопов'язаних і взаємодіючих об'єктів, що призначена для цілеспрямованих дій із завданням досягнення в процесі функціонування заданого результату.

Відмітними ознаками технічних систем, порівняно з довільною сукупністю об'єктів або порівняно з окремими елементами, є конструктивність (практична якя є основою для відносин між елементами), орієнтованість і взаємопов'язаність складових елементів і цілеспрямованість.

Для того щоб система була стійкою до зовнішніх впливів, вона повинна мати стійку структуру. Вибір структури практично визначає технічний вигляд як всієї системи, такі її підсистем та елементів. Питання про доцільність застосування тієї чи іншої структури має вирішуватися, виходячи з конкретного призначення системи. Від структури залежить також здатність системи до перерозподілу функцій у разі повного або часткового відходу окремих елементів, а отже, надійність і живучість системи при заданих характеристиках її елементів.

**Абстрактні системи** є результатом відображення дійсності (реальних систем) у мозку людини.

Їх настрій – необхідний ступінь забезпечення ефективної взаємодії людини з навколишнім світом. Абстрактні (ідеальні) системи об'єктивні за джерелом походження, оскільки їх першоджерелом є об'єктивно існуюча дійсність.

Абстрактні системи поділяють на системи безпосереднього відображення (відображають певні аспекти реальних систем) та системи генералізуючого (узагальнюючого) відображення. До перших належать математичні та евристичні моделі, а до других – концептуальні системи (теорії методологічної побудови) і мови.

На основі поняття зовнішнього середовища системи поділяються на відкриті, закриті (замкнуті, ізольовані) і комбіновані. Розподіл систем на відкриті і закриті пов'язано з їх характерними ознаками: можливість збереження властивостей при наявності зовнішніх впливів. Якщо система не чутлива до зовнішніх впливів, її можна вважати закритою. В іншому випадку – відкритою.

Відкритою називається система, яка взаємодіє з навколишнім середовищем. Усі реальні системи є відкритими. Відкрита система є частиною більш загальної системи або кількох систем. Якщо виокремити з цього утворення власне розглянуту систему, то решта – її середовище.

Відкрита система пов’язана з середовищем певними комунікаціями, тобто мережею зовнішніх зв’язків системи. Виділення зовнішніх зв’язків та опис механізмів взаємодії «система – середовище» є центральним завданням теорії відкритих систем. Розгляд відкритих систем дозволяє розширити поняття структури системи. Для відкритих систем воно включає не тільки внутрішні зв’язки між елементами, а й зовнішні зв’язки з середовищем. При описі структури зовнішні комунікаційні канали намагаються розділити на вхідні (за якими середовище впливає на систему) і вихідні (навпаки). Сукупність елементів цих каналів, що належать власній системі, називається вхідними та вихідними полюсами системи. У відкритих систем, принаймні, один елемент має зв’язок із зовнішнім середовищем, щонайменше, один вхідний полюс і один вихідний, якими вона пов’язана з зовнішнім середовищем.

Для кожної системи зв’язку усі підпорядковані їй підсистеми є внутрішніми, а інші – зовнішніми. Зв’язки між системами і зовнішнім середовищем також, як і між елементами системи, мають, як правило, спрямований характер.

Важливо підкреслити, що в будь-якій реальній системі через законіи діалектики про загальний зв’язок явищ кількість усіх взаємозв’язків величезна, тому врахувати і досліджувати абсолютно всі зв’язки не можливо, отже їх кількість штучно обмежують. Разом з тим, враховувати всі можливі зв’язки недоцільно, оскільки серед них є багато несуттєвих, практично не впливає на функціонування системи і кількість отриманих рішень (з точки зору вирішуваних завдань). Якщо зміна характеристик зв’язку, її виключення (повний розрив) призводять до значного погіршення роботи системи, зниження ефективності, то такий зв’язок – істотний. Одне з найважливіших завдань дослідника – виділити суттєві для розгляду системи в умовах розв’язуваної задачі зв’язку і відокремити їх від несуттєвих. У зв’язку з тим, що вхідні і вихідні полюси системи не завжди вдається чітко виділити, доводиться вдаватися до певної ідеалізації дій. Найбільша ідеалізація має місце при розгляді закритої системи.

Закритою називається система, яка не взаємодіє з середовищем або взаємодіє з середовищем строго визначеним чином. У першому випадку передбачається, що система не має вхідних полюсів, а в другому, що вхідні полюси є, але вплив середовища має постійний характер і повністю (заздалегідь) відомий. Очевидно, що при останньому припущенні зазначені дії можуть бути віднесені власне до системи, та її можна розглядати як закриту. Для закритої системи будь-який її елемент має зв’язки тільки з елементами самої системи.

Зрозуміло, закриті системи являють собою деяку абстракцію реальної ситуації, оскільки, строго кажучи, ізольованих систем не існує. Однак, очевидно, спрощення опису системи, полягають у відмові від зовнішніх зв’язків, що може призвести до корисних результатів, спростити дослідження системи. Усі реальні системи тісно або слабо пов’язані з зовнішнім середовищем – відкриті. Якщо тимчасовий розрив або зміна характерних зовнішніх зв’язків не викликає відхилення у функціонуванні системи понад встановлені заздалегідь межі, то система пов’язана з зовнішнім середовищем слабо. В іншому випадку – тісно.

Комбіновані системи містять відкриті і закриті підсистеми. Наявність комбінованих систем свідчить про складноі комбінації відкритої і закритої підсистем.

Залежно від структури і просторово-часових властивостей системи діляться на прості, складні й великі.

Прості – це системи, що не мають розгалужених структур, що складаються з невеликої кількості взаємозв’язків і невеликої кількості елементів. Такі елементи служать для виконання найпростіших функцій, в них не можна виділити ієрархічні рівні. Відмітною особливістю простих систем є детермінованість (чітка визначеність) номенклатури, кількості елементів і зв’язків як усередині системи, так і з середовищем.

Складні – характеризуються кількість взаємозв’язаних елементів і внутрішніх зв’язків, їх неоднорідністю і різноякісністю, структурною різноманітністю, виконують складну функцію або ряд функцій. Компоненти складних систем можуть розглядатися як підсистеми, кожна з яких може бути деталізована ще більш простими підсистемами і т.д. до ти, поки не буде отриманий елемент.

На підставі викладеного можна перерахувати основні відмінні ознаки складних систем:

1. Наявність великої кількості взаємопов’язаних і взаємодіючих між собою елементів.

2. Складність функції, виконуваної системою і спрямованої на досягнення заданої мети функціонування.

3. Можливість розбиття системи на підсистеми, цілі функціонування яких підпорядковані спільній меті функціонування всієї системи.

4. Наявність керування (часто має ієрархічну структуру), розгалуженої інформаційної мережі й інтенсивних потоків інформації.

5. Наявність взаємодії із зовнішнім середовищем і функціонування в умовах впливу випадкових факторів.

Природно, що всі ознаки розглядаються у взаємозв’язку. Ієрархічна побудова – характерна ознака складних систем, при цьому рівні ієрархії можуть бути як однорідні, так і неоднорідні. Для складних систем властиві такі фактори, як неможливість передбачити їхню поведінку, тобто вони слабо передбачаються, їх потайність, різноманітність стану.

Складні системи можна поділити на такі факторні підсистеми:

1) ***вирішальну***, яка приймає глобальні рішення у взаємодії із зовнішнім середовищем і розподіляє локальні завдання між усіма іншими підсистемами;

2) ***інформаційну***, яка забезпечує збір, переробку і передачу інформації, необхідної для прийняття глобальних рішень і виконання локальних завдань;

3) ***керуючу*** для реалізації глобальних рішень;

4) ***гомеостазну***, що підтримує динамічну рівновагу всередині систем і регулює потоки енергії та речовини в підсистемах;

5) ***адаптивну***, що накопичують досвід у процесі навчання для поліпшення структури і функцій системи.

Великою називають систему, неспостережувану одночасно з позиції одного спостерігача в часі чи просторі, для якої суттєво просторовий фактор, кількість підсистем якої дуже велика, а склад різнорідний.

Система може бути і великою, і складною. Складні системи об’єднують більш велику групу систем, тобто великі – підклас складних систем.

Основоположними при аналізі та синтезі великих і складних систем є процедури декомпозиції та агрегування.

***Декомпозиція*** – поділ систем на частини, з подальшим самостійним розглядом окремих частин.

Очевидно, що декомпозиція являє собою поняття, пов’язане з моделлю, так як сама система не може бути розчленована без порушень властивостей. На рівні моделювання, розрізненого зв’язку заміняться відповідними еквівалентами, або моделі систем будуються так, що розкладання її на окремі частини при цьому виявляється природним.

Стосовно до великих і складних систем декомпозиція є потужним інструментом дослідження.

***Агрегування*** є поняттям, протилежним декомпозиції. У процесі дослідження виникає необхідність об’єднання елементів системи з метою розглянути її з більш загальних позицій.

Декомпозиція та агрегування являють собою дві протилежні сторони підходу до розгляду великих і складних систем, що застосовуються в діалектичній єдності.

Системи, для яких стан системи однозначно визначається початковими значеннями і може бути передбачений для будь-якого наступного моменту часу, називаються ***детермінованими***.

***Стохастичні*** – системи, зміни в яких мають випадковий характер. При випадкових впливах даних про стан системи недостатньо для передбачення в наступний момент часу.

За ступенем організованості: добре організовані, погано організовані (дифузні).

Уявити аналізований об’єкт або процес у вигляді добре організованої системи означає визначити елементи системи, їх взаємозв’язок, правила об’єднання в більш великі компоненти. Проблемна ситуація може бути описана у вигляді математичного виразу. Рішення завдання при поданні її у вигляді добре організованої системи здійснюється аналітичними методами формалізованого подання системи.

Приклади добре організованих систем: сонячна система, що описує найбільш істотні закономірності руху планет навколо Сонця; відображення атома у вигляді планетарної системи, що складається з ядра й електронів; опис роботи складного електронного пристрою за допомогою системи рівнянь, що враховує особливості умов його роботи (наявність шумів, нестабільності джерел живлення і т. п.).

Опис об’єкта у вигляді добре організованої системи застосовується в тих випадках, коли можна запропонувати детермінований опис і експериментально довести правомірність його застосування, адекватність моделі реальному процесу. Спроби застосувати клас добре організованих систем для подання складних багатокомпонентних об’єктів або багатокритеріальних задач погано вдаються: вони потребують неприпустимо великих витрат часу, практично не реалізовуються і не адекватні застосовуваним моделями.

***Погано організовані системи.*** При поданні об’єкта у вигляді погано організованої або дифузійної системи не ставиться завдання визначити всі компоненти, що враховуються, їх властивості та зв’язки між ними і цілями системи. Система характеризується деяким набором макропараметрів і закономірностями, які знаходяться на основі дослідження не всього об’єкта або класу явищ, а на основі визначеної за допомогою деяких правил вибірки компонентів, що характеризують досліджуваний об’єкт або процес. На основі такого вибіркового дослідження отримують характеристики чи закономірності (статистичні, економічні) і поширюють їх на всю систему в цілому. При цьому робляться відповідні застереження. Наприклад, при отриманні статистичних закономірностей їх поширюють на поведінку всієї системи з деякою надійною ймовірністю.

Підхід до відображення об’єктів у вигляді дифузних систем широко застосовується при описі систем масового обслуговування, визначенні чисельності штатів на підприємствах та установах, дослідженні документальних потоків інформації в системах керування і т. д.

З точки зору характеру функцій розрізняють спеціальні, багатофункціональні, й універсальні системи.

Для спеціальних систем характерна одиничність призначення і вузька професійна спеціалізація обслуговуючого персоналу (порівняно нескладна).

Багатофункціональні системи дозволяють реалізувати на одній і тій же структурі декілька функцій. Приклад: виробнича система, що забезпечує випуск різної продукції в межах певної номенклатури.

Універсальні системи передбачають реалізацію безлічі дій на одній і тій же структурі, однак склад функцій за видом і кількістю менш однорідний (менше визначений). Наприклад, комбайн.

За характером розвитку існують 2 класи систем: стабільні і ті, що розвиваються.

У стабільної системи структура і функції практично не змінюються протягом всього періоду її існування і, як правило, якість функціонування стабільних систем у міру зношування їх елементів тільки погіршується. Відбудовні заходи зазвичай можуть лише знизити темп погіршення.

Відмітною особливістю систем, що розвиваються, є те, що з плином часу їх структура і функції набувають істотних змін. Функції системи більш постійні, хоча часто і вони видозмінюються. Практично незмінним залишається лише їх призначення. Системи, що розвиваються, мають більш високу складність.

За порядком ускладнення поведінки розрізняють: автоматичні, вирішальні, самоорганізуючі, ті, що перетворюються.

***Автоматичні***: однозначно реагують на обмежений набір зовнішніх впливів, внутрішня їх організація пристосована до переходу в рівноважний стан при виведенні з нього (гомеостаз).

***Вирішальні***: мають постійні критерії розрізнення їх постійної реакції на широкі класи зовнішніх впливів. Сталість внутрішньої структури підтримується заміною елементів, що вийшли з ладу.

***Самоорганізуючі***: мають гнучкі критерії розрізнення та гнучкі реакції на зовнішні впливи, пристосовуються до різних типів впливу. Стійкість внутрішньої структури вищих форм таких систем забезпечується постійним самовідтворенням. Самоорганізуючі мають ознаки дифузних систем: стохастичність поведінки, нестаціонарність окремих параметрів і процесів. До цього додаються такі ознаки, як непередбачуваність поведінки; здатність адаптуватися до мінливих умов середовища, змінювати структуру при взаємодії системи з середовищем, зберігаючи при цьому властивості цілісності; здатність формувати можливі варіанти поведінки і вибирати з них найкращий та ін. Іноді цей клас розбивають на підкласи, виділяючи адаптивні або системи, що самопристосовуються, самовідтворювані, що самовідтворюються, та інші підкласи, які відповідають різним властивостям систем, що розвиваються.

Приклади: біологічні організації, колективна поведінка людей, організація керування на рівні підприємства, галузі, держави в цілому, тобто в тих системах, де обов'язково є людський фактор.

Якщо стійкість за своєю складністю починає перевершувати складні впливи зовнішнього світу, то це система, що передбачає – вона може передбачити подальший хід взаємодії.

Систему можна розділити на види за ознаками структури їх побудови та значущості тієї ролі, яку відіграють у них окремі складові частини порівняно з ролями інших частин.

У деяких системах одній з частин може належати домінуюча роль (її значущість >> (символ відношення «значної переваги») значущість інших частин). Такий компонент буде виступати як центральний, що визначає функціонування всієї системи. Такі системи називають централізованими.

В інших системах всі складові їх компоненти приблизно однаково значущі. Структурно вони розташовані не навколо деякого централізованого компонента, а взаємопов'язані послідовно або паралельно і мають приблизно однакові значення для функціонування системи. Це децентралізовані системи.

Системи можна класифікувати за призначенням. Серед технічних та організаційних систем виділяють: виробляючі, керуючі, обслуговуючі.

У виробляючих системах реалізуються процеси отримання деяких продуктів або послуг. Вони в свою чергу діляться на матеріально-енергетичні, в яких здійснюється перетворення природного середовища або сировини в кінцевий продукт речової або енергетичної природи, або транспортування такого роду продуктів; та інформаційні – для збору, передачі і перетворення інформації та надання інформаційних послуг.

Призначення керуючих систем – організація і керування матеріально-енергетичними та інформаційними процесами.

Обслуговуючі системи займаються підтримкою заданих меж працездатності виробляючих та керуючих систем.

Реальні системи діляться на **природні** (природні системи) і **штучні** (антропогенні).

**Природні** системи: системи неживої (фізичні, хімічні) і живої (біологічні) природи.

**Штучні** системи: створюються людством для своїх потреб або утворюються в результаті цілеспрямованих зусиль.

Штучні поділяються на технічні (техніко-економічні) і соціальні (громадські).

Технічна система спроектована і виготовлена людиною в певних цілях.

До соціальних систем належать різні системи людського суспільства.

Виділення систем, що складаються з одних тільки технічних пристроїв, майже завжди умовно, оскільки вони не здатні виробляти свій стан. Ці системи виступають як частини більш великих організаційно-технічних систем, що включають людей.

## 1.5. Принципи і закономірності дослідження та моделювання систем

Цілісність/емерджентність. Закономірність цілісності/емерджентності проявляється в появі в системі нових властивостей, відсутніх у елементів.

Для того щоб глибше зрозуміти закономірність цілісності, необхідно, насамперед, враховувати дві її сторони:

1. Властивості системи (цілого) $Q_s$ не є простою сумою властивостей складових її елементів (частин):

$$Q_s \neq \sum Q_i.$$

2. Властивості системи (цілого) залежать від властивостей складових її елементів (частин):

$$Q_s = f(q_i).$$

Крім цих двох основних сторін, слід мати на увазі, що об’єднані в систему елементи, як правило, втрачають частину своїх властивостей, притаманних їм поза системою, тобто система як би пригнічує ряд властивостей елементів. Але, з іншого боку, елементи, потрапивши в систему, можуть набути нових властивостей.

Звернемося до закономірності, двоїстої по відношенню до закономірності цілісності. Її називають фізичною адитивністю, незалежністю, сумативністю, відокремленістю. Властивість фізичної адитивності виявляється в системі, ніби вона розпалася на незалежні елементи; тоді стає справедливим

$$Q_s = \sum Q_i.$$

У цьому крайньому випадку і говорити про систему вже не можна.

Розглянемо проміжні варіанти – дві поєднані закономірності, які можна назвати прогресуючою факторизацією – прагненням системи до стану з дедалі більш незалежними елементами, і прогресуючою систематизацією – прагненням системи до зменшення самостійності елементів, тобто до більшої цілісності.

***Інтегративність.*** Цей термін часто вживається як синонім цілісності. Однак деякі дослідники виділяють цю закономірність як самостійну, прагнучи підкреслити інтерес не до зовнішніх чинників прояву цілісності, а до більш глибоких причин, що обумовлює виникнення цієї властивості, до чинників, що забезпечують збереження цілісності.

Інтегративними називають системоутворюючі, системозберігаючі фактори, серед яких важливу роль відіграють неоднорідність і суперечливість елементів (досліджувані більшістю філософів), з одного боку, і прагнення їх вступати в коаліції – з іншого.

***Закономірності ієрархічної впорядкованості систем.*** Ця група закономірностей (таблиця 1.2) характеризує і взаємодію системи з її оточенням – з середовищем (значущим або істотним для системи), надсистемою, підпорядкованими системами.

Таблиця 1.2 – Закономірності дослідження

| Закономірності взаємодії частини і цілого | Ступінь цілісності $\alpha$ | Коефіцієнт використання елементів $\beta$ |
|---|---|---|
| цілісність / емерджентність | 1 | 0 |
| прогресуюча систематизація | $\alpha > \beta$ | |
| прогресуюча факторизація | $\alpha < \beta$ | |
| адитивність (сумативність) | 0 | 1 |

***Комунікативність.*** Ця закономірність становить основу визначення системи, де система не ізольована від інших систем, вона пов’язана безліччю комунікацій із середовищем, що являє собою, в свою чергу, складне і неоднорідне утворення, що містить надсистему (метасистему – систему більш високого порядку, задані вимоги та обмеження досліджуваної системи), підсистеми (нижчі, підвідомчі системи) і системи одного рівня з розглянутою.

Таке складне єдність із середовищем називається закономірністю комунікативності, яка, в свою чергу, легко допомагає перейти до ієрархічності як закономірності побудови усього світу і будь-якої виділеної з нього системи.

***Ієрархічність.*** Необхідно враховувати не тільки зовнішню структурну сторону ієрархії, але й функціональні відношення між рівнями. Наприклад, у біологічних організаціях більш високий ієрархічний рівень надає спрямовуючий вплив на нижчий рівень, підпорядкований йому, і цей вплив проявляється в тому, що підпорядковані члени ієрархії набувають нових властивостей, відсутніх у них в ізольованому стані (підтвердження положення про вплив цілого на елементи, наведеного вище), а в результаті появи цих нових властивостей формується новий, інший «вигляд цілого» (вплив властивостей елементів на ціле). Нове ціле, що виникло таким чином, набуває здатності здійснювати нові функції, в чому і полягає мета утворення ієрархій.

Виділимо основні особливості ієрархічної впорядкованості з точки зору корисності їх використання як моделей системного аналізу:

1. Через закономірність комунікативності, яка проявляється не тільки між виділеною системою та її оточенням, а й між рівнями ієрархії досліджуваної системи, кожен рівень ієрархічної впорядкованості має складні відношення з вищим і нижчим рівнями. За метафоричним формулюванням, кожен рівень ієрархії має властивість «дволикого Януса»: «лик», спрямований у бік нижчого рівня, має характер автономного цілого (системи), а «лик», спрямований до вузла (вершини) вищого рівня, проявляє властивості залежної частини (елемента вищої системи). Ця конкретизація закономірності ієрархічності пояснює неоднозначність використання в складних організаційних системах понять «система» і «підсистема», «мета» і «засіб» (елемент кожного рівня ієрархічної структури цілей виступає як мета по відношенню до нижчого і як «підціль», а починаючи з деякого рівня, і як «засіб» по відношенню до вищої мети), що часто спостерігається в реальних умовах і призводить до некоректних термінологічних суперечок.

2. Найважливіша особливість ієрархічної впорядкованості як закономірності полягає в тому, що закономірність цілісності / емерджентності (тобто якісні зміни властивостей компонентів більш високого рівня порівняно з поєднуваними компонентами нижчого) проявляється в ній на кожному рівні ієрархії. При цьому об’єднання елементів у кожному з вузлів ієрархічної структури призводить не тільки до появи нових властивостей у вузла і втрати поєднуваними компонентами свободи прояву деяких своїх властивостей, але і до того, що кожен підпорядкований член ієрархії набуває нових властивостей, відсутніх у нього в ізольованому стані.

**Закономірності здійсненності систем.** Проблема здійсненності систем є найменш дослідженою. Розглянемо деякі з закономірностей, що допомагають зрозуміти цю проблему і враховувати її при визначенні принципів проектування та організації функціонування систем керування.

**Еквіфінальність.** Ця закономірність характеризує як би граничні можливості системи. Відповідно до цієї закономірності система може досягти необхідного кінцевого стану, не залежного від часу, і визначається виключно власними характеристиками системи при різних початкових умовах і різними шляхами. Це форма стійкості по відношенню до початкових і граничних умов.

**Закон «необхідної різноманітності».** На необхідність враховувати граничну здійсненність системи при створенні вперше в теорії систем звернув увагу У.Р. Ешбі. Він сформулював закономірність, відому за назвою закон «*необхідної різноманітності*».

Для задач прийняття рішень найбільш важливим є один із висновків цієї закономірності, яку можна спрощено пояснити на такому прикладі.

Коли дослідник (ОПР – особа, яка приймає рішення, спостерігач) $N$ стикається з проблемою $D$, вирішення якої для нього неочевидно, то має місце деяка розманітність можливих рішень $V_d$. Цій різноманітності протистоїть різноманітність думок дослідника (спостерігача) $V_n$. Завдання дослідника полягає в тому, щоб звести різноманітність $V_d - V_n$ до мінімуму, в ідеалі – до 0.

Ешбі довів теорему, на основі якої можливо сформулювати такий висновок: «Якщо $V_d$ дано постійне значення, то $V_d - V_n$ може бути зменшено лише за рахунок відповідного зростання $V_n$. Тільки різноманітність у $N$ може зменшити різноманітність, створювану в $D$; тільки різноманітність може знищити різноманітність».

Стосовно до систем керування закон «необхідної різноманітності» може бути сформульовано таким чином: різноманітність керуючої системи (системи керування) $V_{su}$ має бути більшою (або, принаймні, такою самою) від різноманітності керованого об’єкта $V_{ou}$:

$$V_{su} > V_{ou}.$$

Можливі такі шляхи вдосконалення керування при ускладненні виробничих процесів:

1) збільшення $V_{su}$, що може бути досягнуто шляхом зростання чисельності апарату керування, підвищення його кваліфікації, механізації та автоматизації управлінських робіт;

2) зменшення $V_{ou}$ за рахунок встановлення більш чітких і певних правил поведінки компонентів системи: уніфікація, стандартизація, типізація, введення потокового виробництва, скорочення номенклатури деталей, вузлів, технологічного оснащення і т.п.;

3) зниження рівня вимог до керування, тобто скорочення кількості постійно контрольованих і регульованих параметрів керованої системи;

4) самоорганізація об'єктів керування шляхом обмеження контрольованих параметрів за допомогою створення саморегульованих підрозділів (цехів, дільниць із замкнутим циклом виробництва, з відносною самостійністю і обмеженням втручання централізованих органів керування підприємством і т.п.).

**Закономірності розвитку систем.** Останнім часом дедалі більше починає усвідомлюватися необхідність врахування при моделюванні систем принципів їх зміни в часі, для розуміння яких можуть допомогти розглянуті нижче закономірності.

**Історичність**. Хоча, здавалося б, очевидно, що будь-яка система не може бути незмінною, що вона не тільки виникає, функціонує, розвивається, але й гине, і кожен легко може навести приклади становлення, розквіту, занепаду (старіння) та навіть смерті (загибелі) біологічних і соціальних систем, все ж для конкретних випадків розвитку організаційних систем і складних технічних комплексів важко визначити ці періоди. Не завжди керівники організацій та конструктори технічних систем враховують, що час є неодмінною характеристикою системи, що кожна система підпорядковується закономірності історичності, і що ця закономірність – така ж об'єктивна, як цілісність, ієрархічна впорядкованість і ін.

При цьому закономірність історичності можна враховувати не тільки пасивно, фіксуючи старіння, а й використовувати для попередження «смерті» системи, розробляючи «механізми» реконструкції, реорганізації системи для збереження її в новій якості.

**Закономірність самоорганізації.** Серед основних особливостей, що самоорганізуються з активними елементами, є здатність протистояти ентропійним (ентропія в цьому випадку – ступінь невизначеності, непередбачуваності стану системи і зовнішнього середовища) тенденціям, здатність адаптуватися до мінливих умов, перетворюючи за необхідності свою структуру і т.п. В основі цих здібностей, що виявляються зовні, лежить глибша закономірність, що базується на поєднанні в будь-якій реальній системі, що розвивається, двох суперечливих тенденцій: з одного боку, для всіх явищ, у тому числі і для тих, що розвиваються, відкритих систем справедливий другий закон термодинаміки («другий початок»), тобто прагнення до зростання ентропії; а з іншого боку, спостерігаються негентропійно (протилежні ентропійним) тенденції, що покладені в основі еволюції.

## Контрольні запитання

1. Що є загальною теорією систем?
2. Назвіть основні системні подання (категорії) при дослідженні об'єкта як системи.
3. Що розуміється як структура системи?
4. Що є ефективністю системи?
5. Назвіть основні властивості систем.
6. У чому полягає суть системного підходу при дослідженнях об'єкта?
7. Наведіть приклад класифікації систем за деяким критерієм класифікації.
8. Наведіть приклади абстрактних систем.
9. На які факторні підсистеми поділяються складні системи?
10. Яка різниця в декомпозиції та агрегуванні систем?
11. Назвіть принципи і закономірності дослідження та моделювання.

# 2. АНАЛІЗ ТА СИНТЕЗ СИСТЕМ

## 2.1. Загальні відомості про аналіз і синтез систем

Системний підхід сприяє виробленню правильного способу мислення про сам процес керування, але будь-яка система є частиною більшої системи і постійно змінюється. У разі відсутності достатньої кількості інформації про суть проблемної ситуації для організації процесу прийняття рішень застосовується системний аналіз.

У загальному вигляді процедури системного аналізу включають методики проведення дослідження та організацію процесу прийняття рішення. Предметом системного аналізу є «органічно цілісні системи, до складу яких потрапляють біологічні, психологічні, соціальні, економічні, складні технічні системи, а також комплексні кліматичні, географічні та геологічні утворення». Основу системного аналізу становить загальна теорія систем, яка дозволяє здійснювати дослідження проблем, що не вирішуються аналітично. Як правило, подібного роду проблеми містять невизначеність ситуації, що ускладнює прийняття рішень. Системний підхід об'єднує формальні знання та інтуїцію фахівців і стимулює цілеспрямоване аналітичне мислення. Він передбачає розбиття процесу дослідження на підпроцеси, моделює процеси цілеутворення і дозволяє виробити алгоритм прийняття рішення, спрямований на усунення проблем, що накопичилися.

У процесі системного аналізу здійснюється не тільки системне формулювання проблем, але й встановлення між ними причинно-наслідкових зв'язків і визначення найбільш значущих серед них, для того щоб потім сформулювати мету і визначити способи її досягнення. При цьому часто логічний аналіз супроводжується математичними, статистичними обчисленнями й вербальними оцінками як проблем, так і цілей та варіантів їх досягнення.

Суть аналізу (декомпозиції) полягає в поділі цілого на частини, в поданні складного у вигляді простих складових.

Особливість системного аналізу – використання формальних і неформальних процедур визначення цілей та функцій систем керування. Цей аналіз застосовується для вирішення проблем у ситуації невизначеності, коли слід використовувати експертні методи прийняття рішень.

***Аналіз*** розуміється, як процес дослідження систем, оснований на їх декомпозиції з подальшим визначенням статичних і динамічних характеристик елементів, розглянутих у взаємозв'язку з іншими елементами систем і навколишнім середовищем. Мета аналізу проявляється в прагненні підвищити ефективність функціонування системи, а також у визначенні найкращого варіанта серед усіх альтернативних.

Відносно систем керування задачі аналізу зводяться до таких процедур:

- ➢ визначення об'єкта аналізу;
- ➢ структурування системи;
- ➢ визначення функціональних особливостей системи керування;
- ➢ дослідження інформаційних характеристик системи;
- ➢ визначення кількісних і якісних показників системи керування;
- ➢ оцінка ефективності системи керування;
- ➢ узагальнення й оформлення результатів аналізу.

У цьому процесі дослідник може обрати один із двох напрямків аналізу: визначення стану системи, щоб позначити зони, які потребують поліпшення, і стимулювання змін або дослідження альтернативних варіантів знову створюваної системи з метою вибору кращого варіанта.

***Синтез*** (агрегування) є центральною ланкою створення систем; його суть полягає в поєднанні (уявному або реальному) простих складових об'єкта в єдине ціле.

Розглянемо аналітичні й синтетичні методи дослідження систем.

Свого часу Р. Декарт, французький філософ і математик, пропонував: розчленуйте досліджувану задачу на стільки частин, щоб легко й зручно було її вирішувати. Інший підхід відомий із міркувань прадавніх філософів: усі люди смертні; Каїн – людина, а значить Каїн смертний.

У першому випадку використовувалися методи аналізу, у другому – синтетичний метод дослідження.

Основні етапи розглянутих методів наведені в табл. 2.1.

Складність системного аналізу полягає в тому, що при розчленовуванні цілого на частини не втрачаються властивості системи (властивості цілого).

Таблиця 2.1 – Основні етапи розглянутих методів

| Аналіз | Синтез |
|---|---|
| Об'єкт розділяється на частини | Об'єкт розглядається як частина більшого цілого |
| Пояснюються частини | Пояснюється ціле, яке містить досліджувану частину |
| Знання про частини, агрегується в знання про цілий об'єкт | Ціле декомпозується для пояснення частин |

## 2.2. Моделі систем як основи декомпозиції

***Основи декомпозиції*** у цьому випадку розуміють як сукупність елементів системи (частин), углиб яких не проникає опис, тобто вони є умовно неподільними.

Відомо, що якість побудованих структур залежить від застосовуваної методики декомпозиції. При цьому набір частин, з одного боку, має бути повним, а з іншого – не повинен бути надлишковим. Таким чином, основою будь-якої декомпозиції є модель складу розглянутої системи.

Питання про повноту декомпозиції – це питання завершеності моделі: частин має бути стільки, скільки елементів містить модель, узята як основа.

Іноді корисно як основи декомпозиції не тільки перебирати різні моделі цільової системи, але й брати спочатку моделі надсистеми, потім – самої системи і, нарешті, модель підсистеми. Часто достатньо організувати простий перебір формальних типів моделей (фреймів): «чорного ящика», складу, структури, структурної схеми, моделі життєвого циклу, моделі масштабу і т.д.

Проблема повноти моделей полягає в тому, що змістовна модель будується за зразком формальної. Важливо відшукати компроміс між повнотою й простотою.

Набір повних моделей (фреймів), за великим рахунком, тільки відкриває перед дослідником поле можливих варіантів вивчення систем і спрямований на те, щоб викликати певні асоціації з приводу досліджуваної системи.

Якщо говорити про ресурси як про засоби, то формальний перелік типів ресурсів складається з енергії, матерії, часу, інформації, кадрів і фінансів.

При аналізі ресурсного забезпечення будь-якої конкретної системи цей перелік не дає можливості пропустити що-небудь важливе. Головна мета при цьому полягає в тому, щоб звести складний об'єкт аналізу до кінцевої сукупності простих підоб'єктів або пояснити конкретну причину неусувної складності.

Алгоритм декомпозиції як спосіб спрощення складного полягає в такому:

1) визначення об'єкта аналізу (будь-який вислів, розкриття змісту якого вимагає структурування);
2) визначення цільової системи (визначити, навіщо потрібно те, що ми збираємося робити; як цільова виступає система, в інтересах якої здійснюється аналіз);
3) вибір формальних моделей (набір фреймів і правил перебору);
4) визначення моделі основи (будується за допомогою класифікаторів на основі вивчення цільової системи);
5) аналіз чергового об'єкта декомпозиції;
6) здійснення процедури декомпозиції;
7) аналіз отриманих фрагментів;
8) перевірка чергового фрагмента на елементарність;
9) перевірка використання всіх фреймів;
10) перевірка деталізованості всіх основ;
11) звіт – остаточний результат у формі графа.

У реалізації наведеного алгоритму компроміс досягається за допомогою понять істотного (необхідного), елементарного (достатнього), а також поступовою наростаючою деталізацією базових моделей та ітеративності алгоритму декомпозиції.

## 2.3. Система методів аналізу

Системний аналіз застосовується для вирішення таких проблем, які не можуть бути сформульовані й вирішені за допомогою окремих формальних методів. У системному аналізі використовуються як формальні методи, так і методи якісного аналізу, спрямовані на активізацію творчого мислення експертів.

Системний аналіз можна розглядати не тільки одним із напрямів розвитку загальної теорії систем та ідей кібернетики: він досліджує загальні закономірності, що належать до складних систем, які вивчаються будь-якою наукою.

Системний аналіз сформувався в 60-х рр. XX ст., коли на основі теорії ефективності, теорії ігор, теорії масового обслуговування з'явилася синтетична дисципліна – «Дослідження операцій». Потім вона поступово переросла в системний аналіз, який став синтезом дослідження операцій і теорії керування. Він застосовується, насамперед, у дослідженні штучних соціотехнічних систем.

Виникаюча гостра проблема відповідно до системного підходу має бути розглянута як щось ціле, як система у взаємодії всіх її компонентів між собою й у взаємодії цілого із зовнішнім середовищем. Однак матеріальні системи настільки складні, що для цілей їх аналізу використовуються, як правило, моделі систем.

У цьому розумінні ***системний аналіз*** являє собою сукупність методів і засобів дослідження й конструювання складних об'єктів, методів обґрунтування рішень при створенні й керуванні технічними, економічними та соціальними системами.

Стосовно до соціальних систем системний аналіз використовується як один із найважливіших методів системного керування організацією. Побудова цих моделей починається зі збору інформації й аналізу розрізнених фактів, що дозволяють зробити узагальнення й виявити емпіричні закономірності. Далі переходять до визначення механізмів, які реалізують ці закономірності, оскільки, якщо існує якась підтверджена фактами закономірність, то існують і механізми, що забезпечують прояв цієї закономірності.

Суперечки про те, чи можна вважати системний аналіз наукою, тривають дотепер. Найбільші складності виникають із дослідженням систем, у яких задіяні люди. Подібні системи слабко формалізуються через багатофакторність зв'язків між елементами. Проте загальний алгоритм проведення системного аналізу полягає в такому: формулювання проблеми, виявлення цілей, формування критеріїв, генерування альтернатив і вибір варіанта розв'язку для наступної реалізації.

Можна зробити висновок про те, що системний аналіз – «це дисципліна, що займається проблемами прийняття рішень в умовах, коли вибір альтернативи потребує аналізу складної інформації різної фізичної природи». Звідси випливає висновок, що джерела системного аналізу і його методичні концепції є у дисциплінах, орієнтованих на проблеми прийняття рішень, у теорії дослідження операцій та загальної теорії керування.

Але, незважаючи на значну складову системного аналізу, орієнтовану на формальний інструментарій і точні методи, традиційні прийоми аналізу, основані на інтуїції людини і її схильності до асоціацій (і ще багато чого іншого, що знаходиться поза математикою й поки ще не властиве штучному інтелекту), продовжують активно використовуватися в системному аналізі.

Головне досягнення системного аналізу полягає в розробці методів переходу від неформальних задач до формальних, від моделей типу «чорного ящика» до моделей типу «білого ящика». Більша частина цих методів має неформальний характер, але вони досить конкретні й придатні для використання як технологія вирішення проблем.

У системному аналізі використовуються такі методи:

- строго формалізовані (експериментальні дослідження, побудови моделей);
- слабко формалізовані (експертні оцінки, колективний вибір);
- у принципі неформалізовані операції (формулювання проблем, виявлення цілей, визначення критеріїв, генерування альтернатив).

Якщо розглядати питання алгоритмізації системного аналізу, то необхідно зазначити, що будь-який процес дослідження за своєю природою є алгоритмічним. Алгоритм є планом цього процесу. Водночас очевидно, що для кожної проблеми може знадобитися особливий алгоритм аналізу.

Класифікація розділяє методи аналізу на чотири основні групи за принципом їх застосування в системних дослідженнях:

- неформальні;
- графічні;
- кількісні;
- моделювання.

**Аналітичні методи** дозволяють описати ряд властивостей багатовимірної й багатозв'язкової системи, відображеної у вигляді однієї-єдиної точки, що робить рух у $m$ -вимірному просторі. Це відображення здійснюється за допомогою функції $f(s)$ або точками оператора (функціонала) $F(S)$. Також можливо відобразити точками дві або більше системи (або їх частини) й розглядати їх взаємодію точок. Кожна з них робить рух і має свою поведінку в $m$ -вимірному просторі. Ця поведінка точок у просторі і їх взаємодія описуються аналітичними закономірностями й можуть бути подані у вигляді величин, функцій, рівнянь або системи рівнянь. Аналітичні методи є основою класичної математики й математичного програмування. Вони застосовуються лише в тому випадку, коли властивості системи можуть бути показані в детермінованих параметрах або у вигляді залежностей між ними.

**Статистичні методи** відображають систему за допомогою випадкових (стохастичних) подій, процесів, які описуються відповідними ймовірнісними (статистичними) характеристиками й статистичними закономірностями. У цьому випадку система подається у вигляді «розмитої» точки (області) у $m$ -вимірному просторі, в яку переводиться система, з урахуванням її властивостей, за допомогою оператора Ф[х]. Статистичні методи застосовуються для дослідження складних недетермінованих (таких, що саморозвиваються, самонавчаються) систем, а також у прикладній інформатиці для створення програм моделювання різних систем.

**Теоретико-множинні методи** подання систем є основою побудови загальної теорії систем. Ці методи дозволяють описувати систему в універсальних загальних поняттях: множина, елемент множини й відношення на множинах. Множини можуть задаватися двома способами: перерахуванням елементів $\alpha_1, \alpha_2, \dots, \alpha_n$ і назв характеристичної властивості (ім'я, що відображає цю властивість), наприклад: $\mathrm{A}, \mathrm{B}$. При використанні таких методів допускається введення будь-яких відношень між елементами на основі математичної логіки, яка є формальною мовою опису відношень між елементами, що належать до різних множин. Теоретико-множинні методи дозволяють описувати складні системи формальною мовою моделювання. Вони використовуються в тому випадку, коли більша й складна система не може бути подана лише методами однієї предметної області, а вимагає взаєморозуміння між фахівцями різних наук. Далі методи системного аналізу стають основою розвитку нових мов програмування й автоматизації проектування систем, які застосовуються в прикладній інформатиці.

**Логічні методи** є мовою опису систем у термінології алгебри логіки, яка лежить в основі функціонування мікроелементів будь-якого комп'ютера. Найбільшого поширення логічні методи набули за назвою булевої алгебри як бінарного уявлення про стан комп'ютерних схем. Кожний стан елемента розглядається як 1 або 0. Ці методи використовуються для створення моделей складних систем, адекватних законам математичної логіки побудови стійких структур.

**Лінгвістичні, семіотичні методи** призначені для створення спеціальних мов опису систем у вигляді понять тезауруса (множини смисловиражаючих елементів мови із заданими значеннєвими відношеннями і зв'язками). Лінгвістичні методи використовуються в прикладній інформатиці для формального подання правил (граматики) поєднання понять у зміст значеннєвих виразів. Семіотика базується на поняттях «символ» (знак), «знакова система», «знакова ситуація», тобто для символічного опису змісту в обчислювальній техніці.

Лінгвістичні й семіотичні методи стали широко застосовуватися в тих випадках, коли для першого етапу дослідження не можливо формалізувати прийняття рішень у погано формалізованих ситуаціях і не можна використати аналітичні й статистичні методи.

**Графічні методи** дозволяють наочно відображати об'єкт у вигляді образу системи, її структури і зв'язків в узагальненому вигляді. Графічні методи можуть бути лінійно-площинними й об’ємними. Найбільш уживані методи зображення системи – у вигляді графіка Ганта, діаграм, гістограм, рисунків і структурних схем. Графічні зображення найбільш наочно описують ситуацію або процес для ухвалення рішення в динамічно мінливих умовах. Такі методи застосовуються для структурно-функціонального аналізу складних систем та процесів, що відбуваються в них, особливо при моделюванні інформаційно-керуючих систем. У них необхідно враховувати взаємодію людини й структурних організацій, технічних обладнань. Графічні методи широко застосовуються на практиці для одержання керуючих рішень на основі мережевого планування.

У системному дослідженні, як правило, використовуються всі типи методів. На кожному етапі дослідження вибирають ті з них, які у разі найкращої комбінації дозволяють створити аргументовану й доказову платформу дослідження.

## 2.4. Структура системного аналізу

Загальний підхід до розв’язання задач може бути поданий як цикл (рис. 2.1).

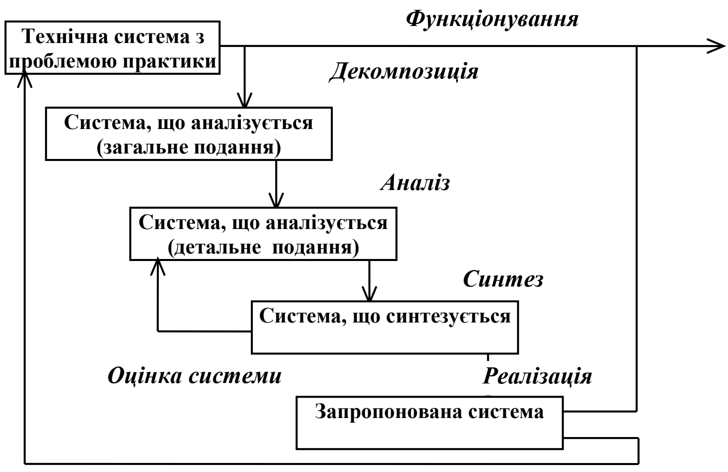


Рис. 2.1. Загальний підхід до вирішення проблем

При цьому в процесі функціонування реальної системи виявляється проблема практики як невідповідність існуючого стану справ необхідному. Для вирішення проблеми проводиться системне дослідження (декомпозиція, аналіз і синтез) системи, що знімає проблему. У процесі синтезу здійснюється оцінка аналізованої і синтезованої систем. Реалізація синтезованої системи у вигляді пропонованої фізичної системи дозволяє провести оцінку ступеня зняття проблеми практики і прийняти рішення на функціонування модернізованої (нової) реальної системи.

При такому поданні стає очевидним ще один аспект визначення системи: система є засобом вирішення проблем.

Основні завдання системного аналізу можуть бути наведені у вигляді трирівневого дерева функцій (рис. 2.2).


Рис. 2.2. Основні задачі системного аналізу

На етапі декомпозиції, що забезпечує загальне уявлення системи, процес здійснюється у послідовності наведеній нижче.

1. Визначення та декомпозиція загальної мети дослідження й основної функції системи як обмеження траєкторії в просторі станів системи або в області допустимих ситуацій. Найбільш часто декомпозиція проводиться шляхом побудови дерева цілей і дерева функцій.

2. Виділення системи з середовища (поділ на систему/«несистему») за критерієм участі кожного розглянутого елемента в процесі, що приводить до результату на основі розгляду системи як складової частини надсистеми.

3. Опис чинників, що впливають.

4. Опис тенденцій розвитку, невизначеностей різного роду.

5. Опис системи як «чорного ящика».

6. Функціональна (за функцією), компонентна (за видом елементів) і структурна (за видом відношень між елементами) декомпозиції системи.

Глибина декомпозиції обмежується. Декомпозиція повинна припинятися, якщо необхідно змінити рівень абстракції – подати елемент як підсистему. Якщо при декомпозиції з'ясовується, що модель починає описувати внутрішній алгоритм функціонування елемента замість закону його функціонування у вигляді «чорного ящика», то в цьому випадку відбулася зміна рівня абстракції. Це означає вихід за межі мети дослідження системи і, отже, викликає припинення декомпозиції.

В автоматизованих методиках типовою є декомпозиція моделі на глибину 5–6 рівнів. На таку глибину декомпозується зазвичай одна з підсистем. Функції, які вимагають такого рівня деталізації, часто дуже важливі, та їх детальний опис дає ключ до секретів роботи всієї системи.

У загальній теорії систем доведено, що більшість систем можуть бути декомпозовані на базові уявлення підсистем. До них належать:

- послідовне (каскадне) з'єднання елементів,
- паралельне з'єднання елементів,
- з'єднання за допомогою зворотного зв'язку.

Проблема проведення декомпозиції полягає в тому, що в складних системах відсутня однозначна відповідність між законом функціонування підсистем і алгоритмом його реалізації. Тому здійснюється формування декількох варіантів (або одного варіанта, якщо система відображена у вигляді ієрархічної структури) декомпозиції системи.

Розглянемо деякі найбільш часто вживані стратегії декомпозиції.

**Функціональна декомпозиція**. Декомпозиція базується на аналізі функцій системи. При цьому ставиться запитання: що робить система, незалежно від того, як вона працює. Підставою розбиття на функціональні підсистеми служить спільність функцій, які виконуються групами елементів.

**Декомпозиція за життєвим циклом**. Ознака виділення підсистем – зміна закону функціонування підсистем на різних етапах циклу існування системи «від народження до загибелі». Рекомендується застосовувати цю стратегію, коли метою системи є оптимізація процесів і коли можна визначити послідовні стадії перетворення входів у виходи.

**Декомпозиція за фізичномим процесом**. Ознака виділення підсистем – кроки виконання алгоритму функціонування підсистеми, стадії зміни станів. Хоча ця стратегія корисна при описі існуючих процесів, результатом її часто може стати занадто послідовний опис системи, який не буде повною мірою враховувати обмеження, продиктовані функціями один одному. При цьому може виявитися прихованою послідовність керування. Застосовувати цю стратегію слід, лише якщо метою моделі є опис фізичного процесу як такого.

Декомпозиція за підсистемами (структурна декомпозиція). Ознака виділення підсистем – сильний зв'язок між елементами за одним із типів відношень (зв'язків), що існують у системі (інформаційних, логічних, ієрархічних, енергетичних і т.п.). Силу зв'язку, наприклад, за інформацією можна оцінити коефіцієнтом інформаційного взаємозв'язку підсистем $k = N/N_0$, де $N$ – кількість взаємовикористовуваних інформаційних масивів у підсистемах, $N_0$ – загальна кількість інформаційних масивів. Для опису всієї системи має бути побудована складова модель, яка об'єднує всі окремі моделі. Рекомендується використовувати розкладання на підсистеми, тільки коли такий поділ на основні частини системи не змінюється. Нестабільність меж підсистем швидко знецінить як окремі моделі, так і їх об'єднання.

На етапі аналізу, що забезпечує формування детального подання системи, здійснюються:

1. Функціонально-структурний аналіз існуючої системи, що дозволяє сформулювати вимоги до створюваної системи. Він включає уточнення складу і законів функціонування елементів, алгоритмів функціонування та взаємовпливів підсистем, поділ керованих і некерованих характеристик, задання простору станів $Z$, задання параметричного простору $T$, в якому задано поведінку системи, аналіз цілісності системи, формулювання вимог до створюваної системи.
2. Морфологічний аналіз – аналіз взаємозв'язку компонентів.
3. Генетичний аналіз – аналіз передісторії, причин розвитку ситуації, наявних тенденцій, побудова прогнозів.
4. Аналіз аналогів.
5. Аналіз ефективності (за результативністю, ресурсомісткістю, оперативністю). Він включає вибір шкали вимірювання, формування показників ефективності, обґрунтування і формування критеріїв ефективності, безпосередньо оцінювання та аналіз отриманих оцінок.

6. Формування вимог до створюваної системи, включаючи вибір критеріїв оцінки і обмежень.

Етап синтезу системи, що вирішує технічну задачу, наведений у вигляді спрощеної функціональної діаграми на рис. 2.3. На цьому етапі здійснюються:

1. Розробка моделі необхідної системи (вибір математичного апарату; моделювання; оцінка моделі за критеріями адекватності, простоти, відповідності між точністю і складністю, балансу похибок, багатоваріантності реалізацій, блочності побудови).

2. Синтез альтернативних структур системи, що знімає проблему.

3. Синтез параметрів системи, що знімає проблему.

4. Оцінювання варіантів синтезованої системи (обґрунтування схеми оцінювання, реалізація моделі, проведення експерименту з оцінки, обробка результатів оцінювання, аналіз результатів, вибір найкращого варіанта).

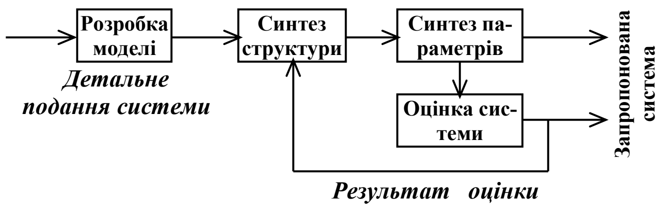


Рис. 2.3. Функціональна діаграма етапу синтезу системи

Оцінка ступеня зняття проблеми проводиться при завершенні системного аналізу.

Найбільш складними у виконанні є етапи декомпозиції та аналізу. Це пов'язано з високим ступенем невизначеності, який потрібно подолати під час дослідження.

Розглянемо процес формування загального і детального подання системи, що включає дев'ять основних стадій.

## 2.5. Формування загального подання системи

*Стадія 1.* Виявлення головних функцій (властивостей, цілей, призначення) системи. Формування (вибір) основних предметних понять, що використовуються в системі. На цій стадії мова йде про з'ясування основних виходів у системі. Саме з цього найкраще починати її дослідження. Повинен бути визначений тип виходу: матеріальний, енергетичний, інформаційний, вони повинні бути віднесені до будь-яких фізичних або інших понять (вихід виробництва – продукція (яка?): Вихід системи керування – командна інформація (для чого? В якому вигляді?); вихід автоматизованої інформаційної системи – відомості (про що?) і т.д.).

Стадія 2. Виявлення основних функцій і частин (модулів) у системі. Розуміння єдності цих частин у рамках системи. На цій стадії відбувається перше ознайомлення з внутрішнім змістом системи, виявляється, з яких великих частин вона складається і яку роль кожна частина відіграє в системі. Це стадія отримання первинних відомостей про структуру і характер основних зв'язків. Такі відомості слід подавати і вивчати за допомогою структурних або об'єктно-орієнтованих методів аналізу систем, де, наприклад, з'ясовується наявність переважно послідовного або паралельного характеру з'єднання частин, взаємної або переважно односторонньої спрямованості впливів між частинами і т.п. Уже на цій стадії слід звернути увагу на так звані системоутворюючі чинники, тобто на ті зв'язки, взаємообумовленості, які і роблять систему системою.

*Стадія 3*. Виявлення основних процесів у системі, їх ролі, умов здійснення; виявлення стадійності, стрибків, змін станів у функціонуванні; в системах з керуванням – виділення основних керуючих факторів. Тут досліджується динаміка найважливіших змін у системі, перебіг подій, уводяться параметри стану, розглядаються фактори, що впливають на ці параметри, що забезпечують перебіг процесів, а також умови початку і кінця процесів. Визначається, чи керовані процеси і сприяють вони здійсненню системою своїх головних функцій. Для керованих систем з'ясовуються основні керуючі впливи, їх тип, джерело і ступінь впливу на систему.

*Стадія 4*. Виявлення основних елементів «несистеми», з якими пов'язана досліджувана система. Виявлення характеру цих зв'язків. На цій стадії вирішується ряд окремих проблем. Досліджуються основні зовнішні впливи на систему (входи). Визначаються їх тип (речові, енергетичні, інформаційні), ступінь впливу на систему, основні характеристики. Фіксуються межі того, що вважається системою, визначаються елементи «несистеми», на які спрямовані основні вихідні впливи. Тут же корисно простежити еволюцію системи, шлях її формування. Нерідко саме це веде до розуміння структури та особливостей функціонування системи. В цілому ця стадія дозволяє краще усвідомити головні функції системи, її залежність і вразливість або відносну незалежність у зовнішньому середовищі.

*Стадія 5*. Виявлення невизначеностей і випадковостей в ситуації їх визначального впливу на систему (для стохастичних систем).

*Стадія 6*. Виявлення розгалуженої структури, ієрархії, формування уявлень про систему як про сукупність модулів, пов'язаних входами-виходами.

Стадією 6 закінчується формування загальних уявлень про систему. Як правило, цього достатньо, якщо мова йде про об'єкт, з яким ми безпосередньо працювати не будемо. Якщо ж мова йде про систему, якою потрібно займатися для її глибокого вивчення, поліпшення, керування, то нам доведеться піти далі спіралеподібним шляхом поглибленого дослідження системи.

## 2.6. Формування детального подання системи

*Стадія 7*. Виявлення всіх елементів і зв'язків, важливих для цілей розгляду. Їх віднесення до структури ієрархії в системі. Ранжування елементів і зв'язків за їх значущістю.

Стадії 6 і 7 тісно пов'язані одна з одною, тому їх обговорення корисно провести разом. Стадія 6 – це межа пізнання «всередину» досить складної системи для особи, яка оперує нею цілком. Більш поглиблені знання про систему (стадія 7) матиме вже тільки фахівець, що відповідає за її окремі частини. Для не надто складного об'єкта рівень стадії 7 – знання системи цілком – досяжний і для однієї людини. Таким чином, хоча суть стадій 6 і 7 одна і та ж, але в першій з них ми обмежуємося тим розумним обсягом відомостей, що доступний одному досліднику.

При поглибленій деталізації важливо виділяти саме істотні для розгляду елементи (модулі) та зв'язки, відкидаючи все те, що не становить інтересу для цілей дослідження. Пізнання системи передбачає не завжди тільки відділення істотного від несуттєвого, але також акцентування уваги на більш істотному. Деталізація повинна торкнутися і вже розглянутий в стадії 4 зв'язок системи з «несистемою». На стадії 7 сукупність зовнішніх зв'язків вважається проясненою настільки, що можна говорити про доскональне знання системи. Стадії 6 і 7 підбивають підсумок загального, цілісного вивчення системи. Подальші стадії вже розглядають тільки її окремі сторони. Тому важливо ще раз звернути увагу на системоутворюючі чинники, на роль кожного елемента і кожного зв'язку, на розуміння, чому вони саме такі або мають бути саме такими в аспекті єдності системи.

*Стадія 8*. Облік змін і невизначеностей у системі. Тут досліджуються повільна, зазвичай небажана зміна властивостей системи, яку прийнято називати «старінням», а також можливість заміни окремих частин (модулів) на нові, що дозволяють не тільки протистояти старінню, а й підвищити якість системи порівняно з початковим станом. Таке вдосконалення штучної системи прийнято називати розвитком. До нього також відносять поліпшення характеристик модулів, підключення нових модулів, накопичення інформації для кращого її використання, а іноді і перебудову структури, ієрархії зв'язків.

Основні невизначеності в стохастичній системі вважаються дослідженими на стадії 5. Однак недетермінованість завжди присутня і в системі, не призначеній працювати в умовах випадкового характеру входів і зв'язків. Додаємо, що облік невизначеностей в цьому випадку зазвичай перетворюється на дослідження чутливості найважливіших властивостей (виходів) системи. Чутливість розуміють як ступінь впливу зміни входів на зміну виходів.

*Стадія 9.* Дослідження функцій і процесів у системі з метою керування ними. Введення керування і процедур ухвалення рішення. Керуючі впливи як системи керування. Для цілеспрямованих та інших систем з керуванням ця стадія має велике значення. Основні керуючі фактори були з'ясовані при розгляді стадії 3, але там це мало характер загальної інформації про систему. Для ефективного введення управлінь або вивчення їх впливів на функції системи і процеси в ній необхідне глибоке знання системи. Саме тому ми говоримо про аналіз управлінь тільки зараз, після всебічного розгляду системи. Нагадаємо, що керування може бути надзвичайно різноманітним за змістом – від команд спеціалізованої керуючої ЕОМ до міністерських наказів.

Однак можливість однакового розгляду всіх цілеспрямованих утручань в поведінку системи дозволяє говорити вже не про окремі управлінські акти, а про систему керування, яка тісно переплітається з основною системою, але чітко виділяється у функціональному відношенні.

На цій стадії з'ясовується, де, коли і як (у яких точках системи, в які моменти, в яких процесах, стрибках, вибору з сукупності, логічних переходах і т.д.) система керування впливає на основну систему, наскільки це ефективно, прийнятно і зручно реалізовується. При введенні управлінь в системі повинні бути досліджені варіанти переведення входів і постійних параметрів у керовані, визначені допустимі межі керування і способи їх реалізації.

Після завершення стадій 6–9 дослідження систем триває на якісно новому рівні – специфічна стадія моделювання. Про створення моделі можна говорити тільки після повного вивчення системи.

## 2.7. Показники та критерії оцінки систем

Системи створюються для реалізації ряду операцій. Необхідний і реально досягнені результати можуть відрізнятися. Це залежить від умов перебігу операцій, якості системи, що реалізують операції, і способів досягнення необхідних результатів.

Тому при оціненні прийнято розрізняти якість систем і ефективність операцій, що реалізуються системою.

### *2.7.1. Співвідношення понять якості та ефективності*

Співвідношення понять якості та ефективності наведено в табл. 2.2.

**Методика визначення критеріїв якості.** Кожна $i$-та якість $j$-тої системи, $i = \overline{1,n}, i = \overline{1,m}$ описується за допомогою деякої вихідної змінної $y_i^j$, що відображає певну істотну властивість системи.

Значення $y_i^j$ характеризує міру цієї якості або частинний показник якості.

4. Показник $y_i^j$ може набувати значення з множини допустимих значень $< y_i^{dop} >$.

Узагальненим показником якості $j$-тої системи буде вектор $Y^j = \left\| y_1^j, y_2^j, \dots, y_n^j \right\|$, компонентами якого будуть показники його окремих властивостей.

Частинні показники мають різну розмірність. Тому при визначенні $Y^j$ слід оперувати нормованими значеннями.

Завдання нормування визначається у вигляді $y_i^n = {}^{y_i}/_{y_i^0}$, де $y_i^0$ – нормуюча величина.

Можливі кілька підходів до вибору нормуючої величини:

- $y_i^0$ задається ОПР;
- $y_i^0 = \max y_i^j$ ;
- $y_i^0 = \max y_i^j - \min y_i^j$.

Необхідна якість системи задається правилами, яким повинні задовольняти показники істотних властивостей, а перевірка їх виконання називається оцінюванням якості системи.

Таким чином, **критерій якості** – це показник істотних властивостей системи і правило його оцінювання.

Таблиця 2.2 – Співвідношення понять якості та ефективності

| Поняття | Якість | Ефективність |
|---|---|---|
| Визначає | Властивості або сукупність властивостей системи, що обумовлюють її придатність для використання за призначенням | Властивість процесів функціонування системи, що характеризує пристосованість до досягнення мети системи. |
| Сфера застосування | Об'єкт будь-якої природи | Цілеспрямовані системи |
| Основна характеристика | Сукупність властивостей системи, істотних для її використання за призначенням | Ступінь відповідності результатів функціонування системи її цілі |
| Фактор структурного аналізу | Структура системи | Алгоритм функціонування |
| Розмірність | Вектор | Показники результативності, ресурсомшсткості та оперативності |
| Спосіб оцінювання | Критерії придатності, оптимальності, зверхності | Критерії придатності, оптимальності |

Нехай $Y^* = \|y_1^*, y_2^*, \dots, y_n^*\|$ – вектор ідеальної системи.

Тоді область адекватності показника якості визначається

$$\delta \subseteq |Y^{dop}/Y^*|/|Y^*|,$$

де δ – радіус області адекватності.

### *2.7.2. Типи критеріїв якості. Критерій придатності*

$$K^{prid}: (\forall_i)\left(y_i^j \in \delta | \delta_i \to y_i^{dop}, i = \overline{1,n}\right).$$

Правило, згідно з яким $j$-та система вважається придатною, якщо значення всіх $i$-х частинних показників $y_i^j$ цієї системи належить області адекватності $\delta$, а радіус цієї області відповідає допустимим значенням усіх частинних показників.

**Критерій оптимальності**

$$K^{opt}\colon (\exists_i)\left(y_i^j \in \delta | \delta_i \to \delta^{dop}\right).$$

Правило, згідно з яким $j$-та система вважається оптимальною за $i$-м показником якості: якщо існує хоча б один частинний показник $y_i^j$, значення якого належить області адекватності, а радіус цієї області за цим показником оптимальний.

**Критерій переваги**

$$K^{\mathrm{per}}\colon (\forall_{\mathrm{i}})\left(\mathrm{y}_{\mathrm{i}}^{\mathrm{j}} \in \delta | \delta_{\mathrm{i}} \to \delta^{\mathrm{dop}}, \mathrm{i} = \overline{1, \mathrm{n}}\right).$$

Правило, згідно з яким $j$-та система вважається кращою, якщо всі значення частинних показників якості $y_i^j$ належать області адекватності, а радіус цієї області оптимальний за всіма показниками.

Формально $K^{per} \subset K^{opt} \subset K^{prid}$.

При оцінюванні якості систем з керуванням доцільно виділити декілька рівнів якості, проранжованих за зростанням складності розглянутих властивостей.

### *2.7.3. Шкала рівнів якості системи*

Порядкова шкала рівнів якості і дерево властивостей систем з керуванням наведено на рис. 2.4.

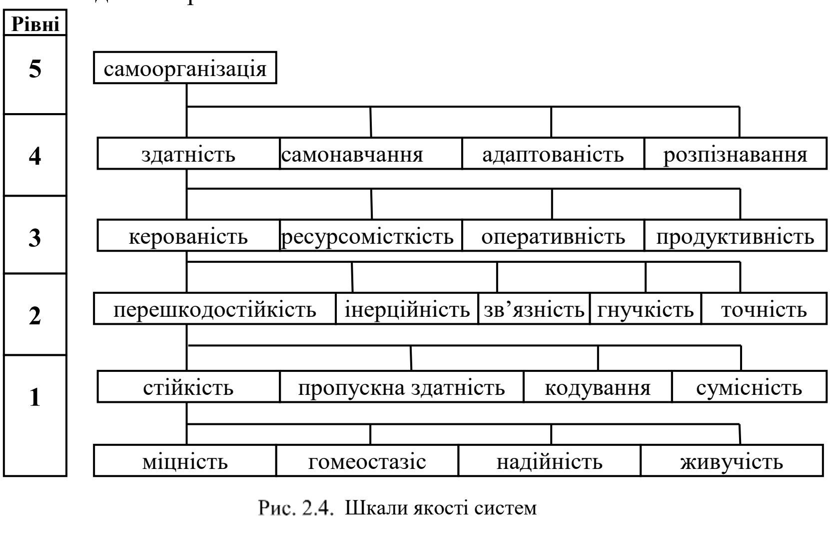


Рис. 2.4. Шкали якості систем

Система, що володіє якістю цього порядку, має і всі інші більш прості якості, але не має якостей вищого порядку.

Первинною якістю будь-якої системи є **стійкість** як здатність системи повертатися в рівноважний стан при виведенні з нього зовнішніми впливами. Для простих систем стійкість об'єднує такі властивості, як міцність, стабільність, стійкість і т.д. Для складних характерні різні форми структурної стійкості, такі як надійність, живучість і т.д.

Більш складним, ніж стійкість, є **перешкодостійкіст**ь, що розуміється як здатність системи без спотворень сприймати і передавати інформаційні потоки. До таких властивостей належать пропускна здатність, кодування інформації, електромагнітна сумісність.

Наступним рівнем шкали якості системи є **керованість** – здатність системи переходити за скінчений час в потрібний стан під впливом керуючих впливів. Керованість забезпечується, насамперед, прямим і зворотним зв'язками та об'єднує такі властивості системи, як оперативність, точність, інерційність і т.д. На цьому рівні якості для складної системи керованість передбачає здатність прийняття рішень щодо формування керуючих впливів.

Наступним рівнем за шкалою якостей є **здатність**, що визначає можливість системи з досягнення необхідного результату на основі наявних ресурсів у заданий період часу. Ця якість характеризується такими властивостями, як результативність, ресурсомісткість та оперативність. По суті, це **потенційна ефективність** функціонування системи – здатність отримати необхідний результат при ідеальному способі використання ресурсів та у відсутності впливу зовнішнього середовища.

Найбільш складною якістю системи є самоорганізація. Самоорганізуюча система здатна змінювати свою структуру, параметри, алгоритм функціонування і т.д. Властивостями самоорганізації є адаптованість, самонавчання, здатність до розпізнавання ситуації.

Рівень якості вибирає дослідник залежно від цілей дослідження, наявності інформації та умови застосування системи.

Суттєві властивості відповідно до подання системи як семантичної моделі можна умовно класифікувати не тільки за рівнем складності, але й за належністю до системоутворюючих (загальносистемних), структурних або функціональних груп. Нижче наведено характерні показники істотних властивостей систем:

- загальносистемні властивості: цілісність, стійкість, спостережуваність, керованість, детермінованість, відкритість, динамічність та ін.;

➢ структурні властивості: склад, зв'язність, організація, складність, масштабність, просторовий розмах, централізованість, обсяг та ін.;

➢ функціональні (поведінкові) властивості: результативність, ресурсомісткість, оперативність, активність, потужність, мобільність, продуктивність, швидкодія, готовність, працездатність, точність, економічність та ін.

При такому розгляді показники якості можна віднести до області загальносистемних і структурних властивостей систем. Властивості ж, які характеризують процес функціонування (поведінка) системи, можна назвати операційними властивостями або властивостями операції, оскільки штучні системи створюються для виконання конкретних операцій.

У загальному випадку оцінка операційних властивостей проводиться як оцінка двох аспектів:

➢ результату (результатів) операції;

➢ алгоритму, що забезпечує отримання результатів.

Якість результату операції та алгоритму, що забезпечує отримання результатів, оцінються за показниками якості операції, до яких належать результативність, ресурсомісткість та оперативність.

Результативність $E$ операції обумовлюється одержуваним цільовим ефектом, заради якого функціонує система.

Ресурсомісткість $R$ характеризується ресурсами всіх видів (людськими, матеріально-технічними, енергетичними, інформаційними, фінансовими тощо), використовуваними для отримання цільового ефекту.

Оперативність $O$ визначається витратою часу, потрібного для досягнення мети операції.

Оцінка результату операції (аспект 1) враховує, що операція проводиться для досягнення певної мети – результату операції. Результат операції розуміється як ситуація (стан системи і зовнішнього середовища), що виникає на момент її завершення. Для кількісної оцінки результату операції вводиться поняття показника результату операції (ПРО), вектора, $Y_{ish} = \|Y_E, Y_R, Y_O\|$, компоненти якого є окремі властивості, що відображають результативність, ресурсомісткість та оперативність операції.

Оцінка алгоритму функціонування (аспект 2) є провідною при оціненні ефективності. Таке твердження ґрунтується на теоретичному постулаті, підтвердженому практикою: наявність гарного «алгоритму» функціонування системи підвищує впевненість в отриманні необхідних результатів. У принципі, необхідні результати можуть бути отримані і без хорошого алгоритму, але ймовірність цього не велика. Це положення особливо важливо для організаційно-технічних систем і систем, в яких результати операції використовуються у режимі реального часу.

У сукупності результативність, ресурсомісткість і оперативність породжують комплексну властивість – ефективність процесу $Y_{ef}$ – ступінь його пристосованості до досягнення мети. Ця властивість, притаманна тільки операціям, проявляється при функціонуванні системи і залежить як від властивостей самої системи, так і від зовнішнього середовища.

В літературі термін «ефективність» пов’язується і з системою, і з операцією, і з рішенням. Утворені при цьому поняття можна вважати еквівалентними. У кінцевому підсумку, кожне з них відображає відповідність результату операції поставленій меті. Зазвичай потрібно мати на увазі, що одна або декілька операцій реалізуються системою. Для більшості операцій процедура оцінки ефективності рішень має характер прогнозування.

Вибір критерію ефективності – центральний, найвідповідальніший момент дослідження системи.

Вважається, що набагато краще знайти неоптимальне рішення правильно обраному критерію, ніж навпаки – оптимальне рішення при неправильно вибраному критерії.

Процес вибору критерію ефективності, як і процес визначення мети, є значною мірою суб’єктивним, творчим, що потребують у кожному окремому випадку індивідуального підходу. Найбільшою складністю відзначається вибір критерію ефективності рішень в операціях, що реалізуються ієрархічними системами.

Математичний вираз критерію ефективності називають цільовою функцією, оскільки її екстремізація є відображенням мети операції. Звідси випливає, що для формування критерію ефективності рішень в операції, насамперед, потрібно визначити поставлену мету. Потім потрібно знайти множини керованих і некерованих характеристик системи, що реалізує операцію. Наступний крок – визначення показників результатів операції. Тільки після цього можливі вибір і формування критерію ефективності. Показники (функції показників) фіналів операції, на основі яких формується критерій ефективності, прийнято називати показниками ефективності. В окремих операціях показник результату операції може прямо виступати критерієм ефективності.

Конкретний фізичний зміст показників визначається характером і цілями операції, а також якістю реалізації системи і зовнішніми впливами.

В окремих системах як показники результативності можуть розглядатися показники ресурсомісткості або оперативності, проте якість операції в цілому не може бути охарактеризована жодною з перерахованих частинних властивостей окремо, а визначається, подібно ПРО, їх сукупністю $Y_{ich} = \|Y_E, Y_R, Y_O\|$.

Хоча конкретні операції досить різноманітні, існує ряд загальних принципових положень, якими необхідно керуватися при формуванні системи критеріїв ефективності рішень.

Залежно від типу систем і зовнішніх впливів операції можуть бути детермінованими, ймовірнісними або невизначеними. Відповідно до цього виділяють три групи показників і критеріїв ефективності функціонування систем:

- в умовах визначеності, якщо ПРО відображають один строго певний результат детермінованої операції;
- в умовах ризику, якщо ПРО є дискретними або безперервними випадковими величинами з відомими законами розподілу в імовірнісній операції;
- в умовах невизначеності, якщо ПРО є випадковими величинами, закони розподілу яких не відомі.

**Критерій придатності** для оцінки детермінованої операції

$$K^{prid}: (\forall_i)\left(y_i^j \in \delta | \delta_i \to y_i^{dop}, i \in \|E, R, O\|\right)$$

визначає правило, за яким операція вважається ефективною, якщо всі частинні показники результату операції належать області адекватності.

**Критерій оптимальності** для оцінки детермінованої операції

$$K^{opt}: (\exists_i)\left(y_i^j \in \delta | \delta_i \to \delta^{dop}, \in \|E, R, O\|\right)$$

визначає правило, за яким операція вважається ефективною, якщо всі частинні показники результату операції належать області адекватності, а радіус області адекватності за цими показниками оптимальний.

**Критерій придатності** для оцінки ефективності ймовірнісної операції

$$K^{prid}: P_{dc}\left(Y_{ef}\right) \geq P_{dc}^{treb}\left(Y_{ef}\right)$$

визначає правило, за яким операція вважається ефективною, якщо ймовірність досягнення мети за показниками ефективності $P_{dc}\left(Y_{ef}\right)$ не менша від необхідної ймовірності досягнення мети за цими показниками $P_{dc}^{treb}\left(Y_{ef}\right)$.

Критерій оптимальності для оцінення ефективності ймовірнісної операції

$$K^{opt}: P_{dc}\left(Y_{ef}\right) = P_{dc}\left(Y_{ef}^{opt}\right)$$

визначає правило, за яким операція вважається ефективною, якщо ймовірність досягнення мети за показниками ефективності $P_{dc}\left(Y_{ef}\right)$ дорівнює ймовірності досягнення мети з оптимальними значеннями цих показників $P_{dc}\left(Y_{ef}^{opt}\right)$.

Основною проблемою оцінки ефективності імовірнісних операцій є неясність способу визначення необхідних ймовірностей. Це пов’язано з відсутністю достатньої статистики. Відомо, що застосування методів класичної теорії ймовірностей допустимо при повторюваності дослідів і однаковості умов. Ці вимоги в складних системах виконуються не завжди.

Найбільші труднощі виникають при оціненні ефективності систем в умовах невизначеності. Для вирішення цього завдання розроблено кілька підходів. Порядок оцінення ефективності систем у невизначених операціях становить один із розділів теорії прийняття рішень.

Вибір показників для конкретної системи пов’язаний з аналізом великого обсягу погано структурованої інформації, і тому в системному аналізі сформульовано вимоги, дотримання яких дозволяє обґрунтувати придатність показників у цьому завданніоцінювання.

Загальними вимогами до показників результату операції є:

- відповідність ПРО мети операції;
- повнота;
- вимірність;
- ясність фізичного змісту;
- ненадмірність;
- чутливість.

Однією з основних вимог є відповідність ПРО мети операції, реалізованої системою. Цілі операції значною мірою залежать від призначення системи. Наприклад, для такої ІС, як АСУ, цілями операції можуть бути забезпечення необхідних значень оперативності, достовірності, стійкості та безпеки вирішення завдань керування і передачі повідомлень та ін. Для кожної з висунутих цілей мають бути визначені одна або декілька складових ПРО.

До основних вимог до ПРО належить також його повнота. Суть цієї вимоги полягає в тому, що ПРО повинен відображати бажані (цільові) і небажані (побічні) наслідки операції за показниками результативності, ресурсомісткості та оперативності. Зауважимо, що одним із показників правильності вибору складових ПРО та їх повноти є монотонний характер функції корисності (цінності), побудованої для кожної складової. Якщо при цьому яка-небудь з функцій не монотонна, то це означає, що упущені одна або декілька складових ПРО.

Наступна важлива вимога до ПРО – вимірність його складових за допомогою або натурного експерименту, або моделей операції. Якщо розглянута операція не дозволяє це зробити, її доцільно розкласти на підоперації, що забезпечують вимірність складових. Процес декомпозиції операції на підоперації може бути багаторівневим. Наприклад, операцію «Розв'язання завдань керування» можна розділити на підоперації: «Розв'язання завдань планування» та «Розв'язання задач оперативного керування», а останні, в свою чергу, – на «Розв'язання задач обліку», «Розв'язання задач контролю» і т.д.

При визначенні завдань ПРО необхідно прагнути до ясності їх фізичного змісту, тобто щоб вони вимірювалися за допомогою кількісних заходів, доступних для сприйняття. Проте досягти цього вдається не завжди. Тоді доводиться вводити так звані суб'єктивні складові ПРО. Наприклад, така властивість людей, як навчальність, зазвичай не може бути визначена за допомогою характеристик, що мають фізичний зміст. У цьому випадку часто вводять деяку штучну шкалу. Інший спосіб забезпечення вимірності складових ПРО – перехід до показників-замінників, побічно характеризує аналізовану властивість. Вимога ясності фізичного змісту обмежує можливості агрегування частинних показників до одного критерію. Так, наприклад, не має фізичного змісту узагальнений скалярний показник, складений із частинних показників результативності, ресурсомісткості та оперативності.

Важливою вимогою до ПРО є мінімізація його розмірності, тобто забезпечення ненадлишкових набору складових. Зі збільшенням кількості складових різко зростає трудомісткість побудови функції ефективності.

І, нарешті, до групи основних вимог до складових ПРО зазвичай включають їх відносно високу чутливість до змін значень керованих характеристик.

Таким чином, набір складових ПРО може бути визначений різними способами, оскільки до теперішнього часу ще не існує формальної теорії, що забезпечує об'єктивне розв'язання цієї задачі. Дві особи, які приймають рішення на одну і ту саму операцію, можуть визначити різний склад ПРО. Важливо лише те, що, використовуючи різні ПРО, вони повинні вибрати однакове рішення – оптимальне.

## Контрольні запитання

1. У чому різниця між аналізом та синтезом систем?
2. Що є основами декомпозиції систем?
3. Поясніть алгоритм декомпозиції систем?

4. Дайте порівняльній аналіз методів аналізу систем.

5. Назвіть основні елементи структури системного аналізу.

6. Охарактеризуйте етап загального уявлення системи.

7. Охарактеризуйте етап детального подання системи.

8. Назвіть основні показники та критерії оцінення систем.

9. Яка різниця між поняттями якості систем та ефективністю операцій, що реалізуються системою?

10. Перерахуйте і роз’ясніть характеристики моделей систем.

# 3. МЕТОДОЛОГІЧНІ ОСНОВИ МОДЕЛЮВАННЯ СИСТЕМ

## 3.1. Поняття моделі та моделювання

Моделювання є основним методом дослідження великих і складних систем у теорії систем.

Кожна теорія – це теж модель розуміння змісту предмета дослідження. Моделі можуть створюватися на основі засобів пізнання (форми мислення) – евристичні, гіпотетичні, концептуальні й на основі раціонально-логічних засобів дослідження – емпіричні, теоретичні, математичні.

Існує багато визначень моделей. Найчастіше модель розуміють як якийсь об’єкт-замінник, який у певних умовах може заміняти об’єкт-оригінал, відтворюючи властивості, що цікавлять нас, і характеристики оригіналу. Причому тут істотну перевагу мають зручності, тобто модель являє собою відображення яким-небудь способом істотних характеристик об’єктів, процесів та їх взаємозв’язків із реальними системами. В основі моделювання покладено принцип аналогії.

Переконавшись в аналогічності двох об’єктів, припускають, що функції, властивості одного об’єкта притаманні іншому об’єкту, для яких вони не встановлені. Метод аналогій полягає в тому, що вивчає один об’єкт – модель, а висновки переносяться на інший – оригінал. Інакше кажучи, аналогія – висновок від моделі до оригіналу.

Модель є свого роду інструментом дослідження систем і дозволяє на основі зміни вихідних припущень прогнозувати поведінку системи. Крім того, модель являє собою засіб спрощення об’єкта і його вивчення, оскільки дозволяє досліджувати систему з погляду її істотних характеристик, абстрагуючись від побічних впливів середовища.

Таким чином, модель є матеріально або теоретично створеною системою, призначеною замінити або подавати об’єкт дослідження в процесі пізнання. Модель має бути зручнішою для дослідження. Вивчення моделі та реалізація за її допомогою різних завдань дозволяє отримати інформацію про реальний об’єкт дослідження. При цьому вимоги до моделі визначаються вирішуваним завданням і наявними засобами.

Серед методів спрощення, здійснюваних у процесі моделювання, можна назвати:

- виключення з розгляду ряду змінних – як виключення несуттєвих, так і за рахунок агрегування змінних;
- зміна природи змінних – за рахунок розгляду змінних як констант, а також і за рахунок розгляду дискретних величин як безперервних;
- зміна характеру зв'язки між елементами (заміни нелінійних залежностей на лінійні);
- зміна обмежень – як шляхом зняття обмежень, так і за рахунок уведення нових.

Будь-яка модель будується на основі деяких теоретичних принципів і реалізується певними інструментальними засобами прикладних наук.

Моделі класифікують за різними ознаками. Наведемо деякі приклади.

***Графічна модель*** – об'єкт, геометрично подібний оригіналу (географічна карта).

***Геометрична модель*** – об'єкт, подібний оригіналу за формою (зліпок).

***Функціональна модель*** – об'єкт, що відображає поведінку оригіналу (будь-яка діюча модель).

***Символічна модель*** – виражається за допомогою абстрактних символів (програма для ЕОМ).

***Статистична модель*** **–** описує взаємозв'язки між елементами, що мають випадковий характер (схема Бернуллі).

***Описова (дескриптивна) модель*** – словесний опис, порівняльні характеристики (різні визначення).

***Математична модель*** – сукупність рівнянь або нерівностей, таблиці, матриці й інші способи опису оригіналу.

Прикладом статичних моделей можуть служити світлина (модель конкретного об'єкта) або топографічна карта місцевості; динамічних моделей – процес обтікання моделі літака в аеродинамічній трубі на різних режимах польоту або демонстрація відеоролика, що зафіксував технологічний процес виготовлення якого-небудь продукту. Можна виділити абстрактні моделі (образи, що приходять у свідомість людини увісні), знакові (математичні моделі) і т.д.

Модель має системо- і змістоутворюючу роль у науковому пізнанні. На моделі вивчають невідомі властивості предметів. Модель прагне, якомога більш яскраво висловити структуру явища, його основні аспекти. Модель є концентрованим вираженням сутності предмета або процесу, виділяючи тільки його основні риси.

Знання – це моделі навколишнього світу, що фіксуються людиною в її мозку або на технічних носіях. Моделі володіють підвищеною наочністю, виділяючи основні аспекти сутності, й активно використовуються в процесах пізнання і навчання. Людина, вирішуючи, як їй вчинити в тій чи іншій ситуації, завжди намагається уявити собі наслідки рішення, для цього вона програє ситуацію, уявляє її собі подумки, будуючи модель у голові. Комп'ютер є підсилювачем для виробництва цієї діяльності, інструментом інформаційної технології. Комп'ютерні моделі прискорюють процес дослідження, роблять його більш точним.

Алгоритми – знання, які вибудовуються людиною (або, ширше, розумною істотою) в ланцюжок так, щоб з'єднати початковий стан із бажаним, метою; це один з варіантів низки заходів, кроків, що приводять до мети.

Таким чином, моделі – це основа розумної розумової діяльності; моделі виконують функцію базису, а моделювання – функцію інструменту для прогнозування.

***Моделювання*** є основним методом досліджень у всіх галузях знань і науково обґрунтованим методом оцінок характеристик складних систем, що використовується для ухвалення рішень у різних сферах інженерної діяльності. Існуючі і проектовані системи можна ефективно досліджувати за допомогою математичних моделей, що реалізуються на сучасних ЕОМ, які в цьому випадку виступають як інструмент експериментатора при роботі з моделлю системи.

Сьогодні не можна назвати сферу людської діяльності, в якій тією або іншою мірою не використовувалися б методи моделювання.

Моделювання можна розглядати як заміщення досліджуваного об'єкта (оригіналу) його умовним образом, описом або іншим об'єктом, що іменується моделлю і забезпечує близьку до оригінала поведінку в рамках деяких допущень і прийнятних похибок. Моделювання зазвичай виконується з метою пізнання властивостей оригіналу шляхом дослідження його моделі, а не самого об'єкта. Зрозуміло, моделювання виправдане у тому випадку, коли воно простіше за створення самого оригіналу або коли останній за якихось причин краще взагалі не створювати. Модель розуміється як фізичний або абстрактний об'єкт, властивості якого в певному значенні схожі з властивостями досліджуваного об'єкта. При цьому вимоги до моделі визначаються вирішуваним завданням і наявними засобами.

Про моделювання природно говорити лише при використанні моделі для пізнання оригіналу.

Заміщення одного об'єкта іншим з метою отримання інформації про найважливіші властивості об'єкта-оригіналу за допомогою об'єкта-моделі називається моделюванням.

Розглядаючи гносеологічну роль теорії моделювання, моделювання можна визначити як метод опосередкованого пізнання, при якому об'єкт-оригінал, що вивчається, знаходиться в деякій відповідності з іншим об'єктом-моделлю, причому модель здатна в тому або іншому відношенні замінювати оригінал на деяких стадіях пізнавального процесу.

Процес моделювання передбачає наявність об'єкта дослідження; дослідника, перед яким поставлено конкретне завдання; моделі, що створюється для здобуття інформації про об'єкт і необхідного для вирішення поставленого завдання. Причому по відношенню до моделі дослідник є, по суті справи, експериментатором, лише в цьому випадку експеримент проводиться не з реальним об'єктом, а з його моделлю. Такий експеримент для інженера є інструментом безпосереднього вирішення організаційно-технічних завдань.

Моделювання починається з формування предмета досліджень – системи понять, що відображає істотні для моделювання характеристики об'єкта. Це завдання є досить складним, що підтверджується різною інтерпретацією в науково-технічній літературі таких фундаментальних понять, як система, модель, моделювання. Подібна неоднозначність не свідчить про помилковість одних і правильність інших термінів, а відображає залежність предмета досліджень (моделювання) від цього об'єкта, а також і від цілей дослідника.

Система $S$ – цілеспрямована множина взаємопов'язаних елементів будь-якої природи. Зовнішнє середовище $E$ – множина елементів будь-якої природи, що існують поза системою, впливають на систему або знаходяться під її дією.

Залежно від мети дослідження можуть розглядатися різні співвідношення між самим об'єктом $S$ і зовнішнім середовищем $E$. Таким чином, залежно від рівня, на якому знаходиться спостерігач, об'єкт дослідження може виділятися по-різному і можуть мати місце різні взаємодії цього об'єкта із зовнішнім середовищем.

Прояв функцій системи в часі $S(t)$, тобто функціонування системи, означає перехід системи з одного стану в інший, тобто рух в просторі станів $Z$.

З розвитком системних досліджень, з розширенням експериментальних методів вивчення реальних явищ дедалі більшого значення набувають абстрактні методи, з'являються нові наукові дисципліни, автоматизуються елементи розумової праці. Важливе значення при створенні реальних систем $S$ мають математичні методи аналізу і синтезу, цілий ряд відкриттів базується на суто теоретичних дослідженнях.

Проте було б неправильно забувати про те, що основним критерієм будь-якої теорії є практика, і навіть суто математичні, абстрактні науки базуються в своїй основі на фундаменті практичних знань. Експеримент був і залишається одним з основних і істотних інструментів пізнання. Подібність і моделювання дозволяють по-новому описати реальний процес і спростити експериментальне його вивчення.

Процес моделювання є процесом переходу з реального у віртуальне (модельне) за допомогою формалізації, далі відбувається вивчення моделі (власне моделювання) і, нарешті, інтерпретація результатів як зворотний перехід з віртуальної сфери в реальну. Цей шлях замінює пряме дослідження об’єкта в реальній сфері, тобто лобове або інтуїтивне розв’язання задачі. Отже, в найбільш простому випадку в технології моделювання існує 3 етапи (рис. 3.1):

- формалізація (перехід від реального об’єкта до моделі);
- власне моделювання (дослідження і перетворення моделі);
- інтерпретація (переведення результатів моделювання в сферу реальності).

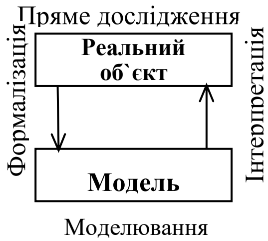


Рис. 3.1. Процес моделювання

Якщо потрібне уточнення, ці етапи повторюються знову й знову: формалізація (проектування), моделювання, інтерпретація.

Теорія заміщення одних об’єктів (оригіналів) іншими об’єктами (моделями) і дослідження властивостей об’єктів на їх моделях називається теорією моделювання.

## 3.2. Теорія моделювання

Теорія моделювання являє собою взаємопов'язану сукупність положень, визначень, методів і засобів створення та вивчення моделей. Ці положення, `визначення, методи та засоби, як і самі моделі, є предметом теорії моделювання.

Основне завдання теорії моделювання полягає в тому, щоб оснастити дослідника технологією створення таких моделей, які досить точно і повно фіксують властивості оригіналу, що є цікавішими: простіше і швидше піддаються дослідженню і допускають перенесення його результатів на оригінал.

Формально модель об'єкта моделювання $S$ зображують у вигляді множини величин, що описують процес функціонування реальної системи і утворюють такі підмножини:

- сукупність вхідних дій на систему

$$x_i \in X, i = \overline{1, n_X} \ ,$$

- сукупність вихідних характеристик системи

$$y_j \in Y, j = \overline{1, n_Y} \ ,$$

- сукупність внутрішніх (власних) параметрів системи

$$h_k \in H, k = \overline{1, n_H} \ ,$$

- сукупність дій зовнішнього середовища

$$v_l \in V, l = \overline{1, n_V} \ .$$

При цьому в перерахованих підмножинах можна виділити керовані і некеровані змінні. У загальному випадку $x_i, y_j, h_k, v_l$ є елементами підмножин, що не перетинаються, і містять як детерміновані, так і стохастичні складові.

При моделюванні системи $S$ вхідні дії, внутрішні параметри системи і дії зовнішнього середовища $E$ є незалежними (екзогенними) змінними, які у векторній формі мають відповідно вигляд

$$\vec{x}(t) = \left\| x_1(t), x_2(t), \dots, x_{n_X}(t) \right\|,$$

$$\vec{h}(t) = \left\| h_1(t), h_2(t), \dots, h_{n_H}(t) \right\|,$$

$$\vec{v}(t) = \left\| v_1(t), v_2(t), \dots, v_{n_V}(t) \right\|,$$

а вихідні характеристики системи є залежними (ендогенними) змінними і у векторній формі мають вигляд

$$\vec{y}(t) = \left\| y_1(t), y_2(t), \dots, y_{n_Y}(t) \right\|.$$

Процес функціонування системи $S$ описується в часі оператором $F_s$, який у загальному випадку перетворить екзогенні змінні в ендогенні відповідно до співвідношень вигляду

$$\vec{y}(t) = F_s\big(\vec{x}(t), \vec{h}(t), \vec{v}(t), t\big). \tag{3.1}$$

Сукупність залежностей вихідних характеристик системи від часу $y_j$ для всіх видів $j = \overline{1, n_Y}$ називається вихідною траєкторією в $\vec{y}(t)$. Залежність (3.1) називається законом функціонування системи $S$ і позначається $F_s$. У загальному випадку закон функціонування системи $F_s$ може бути заданий у вигляді функції, функціонала, логічних умов, в алгоритмічній і табличній формах або у вигляді словесного правила відповідності.

Вельми важливим для опису і дослідження системи $S$ є поняття алгоритму функціонування $A_s$, який розуміється як метод набуття вихідних характеристик з урахуванням вхідних дій $\vec{x}(t)$, дій зовнішнього середовища $\vec{v}(t)$ і власних параметрів системи $\vec{h}(t)$. Очевидно, що один і той самий закон функціонування $F_s$ системи $S$ може бути реалізований різними способами, тобто за допомогою множини різних алгоритмів функціонування $A_s$.

Співвідношення (3.1) може бути задане різними способами: аналітично (за допомогою формул), графічно, таблично і так далі. Такі співвідношення у ряді випадків можуть бути отримані через властивості системи $S$ в конкретні моменти часу, що називаються станами. Стан системи $S$ характеризується вектором

$$\vec{\dot{z}} = \|\dot{z}_1, \dot{z}_2, \dots, \dot{z}_m\|, m = \overline{1, n_z},$$

де $\dot{z}_1 = z_1(\dot{t}), \dot{z}_2 = z_2(\dot{t}), \dots, \dot{z}_m = z_{1m}(\dot{t})$, у деякий момент часу $\dot{t} \in (t_0, T)$ на інтервалі моделювання $(0, T)$.

Якщо розглядати процес функціонування системи $S$ як послідовну зміну станів, то вони можуть бути інтерпретовані як координати точки в багатовимірному фазовому просторі, причому кожній реалізації процесу відповідатиме деяка фазова траєкторія. Сукупність всіх можливих значень станів $\{\vec{z}\}$ називається простором станів об'єкта моделювання $Z$.

Стани системи $S$ у момент часу $t_0 < t^* \leq T$ повністю визначаються початковими умовами $\vec{z}^0 = \vec{z}(t_0)$, вхідними діями $\vec{x}(t)$ внутрішніми параметрами $\vec{h}(t)$ і діями зовнішнього середовища $\vec{v}(t)$, які мали місце в проміжку часу $t_0 \dots t^*$ за допомогою двох векторних рівнянь

$$\vec{z}(t) = \vec{\Phi}\left(\vec{z}^0, \vec{x}, \vec{h}, \vec{v}, t\right) , \tag{3.2}$$

$$\vec{y}(t) = \vec{F}(\vec{z}, t) . \tag{3.3}$$

Перше рівняння за початковим станом $\vec{z}^0$ і екзогенними змінними $\vec{x}(t), \vec{h}(t), \vec{v}(t)$ визначає вектор-функцію $\vec{z}(t)$, а друге – за набутим значенням станів $\vec{z}(t)$– ендогенні змінні на виході системи в $\vec{y}(t)$.

Таким чином, ланцюжок рівнянь об’єкта «вхід – стани – вихід» дозволяє визначити вихідні характеристики системи

$$\vec{y}(t) = \vec{F}\left[\vec{\Phi}\left(\vec{z}^0, \vec{x}, \vec{h}, \vec{v}, t\right)\right] . \tag{3.4}$$

Таким чином, математичну модель об’єкта (реальної системи) розуміють кінцеву підмножину змінних $\left\{\vec{x}(t), \vec{v}(t), \vec{h}(t)\right\}$ разом із математичними зв’язками між ними і вихідними характеристиками $\vec{y}(t)$.

Якщо математичний опис об’єкта моделювання не містить елементів випадковості або вони не враховуються, тобто якщо можна вважати, що в цьому випадку стохастичні дії зовнішнього середовища $\vec{v}(t)$ і стохастичні внутрішні параметри $\vec{h}(t)$ відсутні, то модель називається детермінованою в тому сенсі, що вихідні характеристики системи однозначно визначаються детермінованими вхідними діями:

$$\vec{y}(t) = f(\vec{x}, t). \tag{3.5}$$

Наведені математичні співвідношення (3.1)–(3.5) дозволяють описати широкий клас систем. Проте в практиці моделювання об’єктів на первинних етапах дослідження системи раціональніше використовувати типові математичні схеми. Не володіючи такою мірою узагальнення, як розглянуті моделі, типові математичні схеми мають переваги простоти і наочності, але при істотному звуженні можливостей вживання.

Перерахуємо основні принципи, на яких базується теорія моделювання і які покладені в основу технології моделювання:

***Принцип інформаційної достатності.*** Моделювання системи безглуздо, якщо є вичерпна інформація про її функціонування. Протилежністю цієї тези є таке: якщо немає ніякої інформації про систему, то здійснити моделювання такої системи не можливо. Проміжним варіантом є вимога наявності певного порогового рівня апріорних знань про систему (саме це є принципом інформаційної достатності), коли існують умови побудови моделі, адекватної досліджуваній системі.

***Принцип здійсненності.*** Здійсненність полягає в тому, що розроблювана модель повинна досягти реалізації мети дослідження з відміною від нуля ймовірністю за певний (кінцевий) час. Тобто має виконуватися співвідношення

$$P(t_0) \geq P_0,$$

де $P(t_0)$ – ймовірність досягнення мети моделювання за прийнятний час $t_0$; $P_0$ – задана (бажана) ймовірність досягнення мети моделювання.

***Принцип множинності моделей.*** У моделі, що розробляється, необхідно реалізовувати тільки ті властивості системи або явища, які мають істотний вплив на показник ефективності. З цього випливає, що використання отриманої моделі відображає тільки певні (враховані) сторони (характеристики) реального процесу. Тому для вичерпного дослідження модельованого процесу, можливо, буде потрібна побудова набору моделей, які б дозволили з різних сторін і з різним ступенем деталізації аналізувати характеристики реального процесу.

***Принцип агрегування.*** Будь-яка складна система може бути представлена набором деяких підсистем (агрегатів), а для їх математичного опису можна використовувати певні математичні схеми. Цей принцип дає можливість досить легко перебудовувати модель залежно від виникаючих проблем і завдань дослідження.

***Принцип параметризації.*** Часто зустрічаються системи, в структуру яких включені досить ізольовані компоненти (підсистеми). Якщо ці підсистеми характеризуються деяким параметром, то можливо замінити їх у моделі відповідними числовими значеннями (або графіками, таблицями чи формулами) та не описувати їх функціонування. Використання цього принципу скорочує обсяг і час моделювання, але слід враховувати, що при цьому, природно, знижується адекватність моделі.

## 3.3. Класифікація видів моделювання систем

Класифікація видів моделювання може бути проведена за різними підставами. Один із варіантів класифікації наведено на рис. 3.2.

Відповідно до класифікаційної ознаки повноти моделювання ділиться на

- повне,

- неповне,
- наближене.

При ***повному*** моделюванні моделі ідентичні об'єкту в часі і просторі.

Для ***неповного*** моделювання ця ідентичність не зберігається.

В основі ***наближеного*** моделювання лежить подібність, при якій деякі сторони реального об'єкта не моделюються зовсім. Відповідно до теорії подібності, абсолютна подібність можлива лише при заміні одного об'єкта іншим точно таким самим. Тому при моделюванні абсолютна подібність не має місця. Дослідники прагнуть до того, щоб модель добре відображала тільки досліджуваний аспект системи. Наприклад, для оцінки перешкодостійкості дискретних каналів передачі інформації функціональна та інформаційна моделі системи можуть не розроблятися. Для досягнення мети моделювання цілком достатня модель, що описується матрицею умовних ймовірностей переходів $i$-го символу алфавіту в $j$-й.

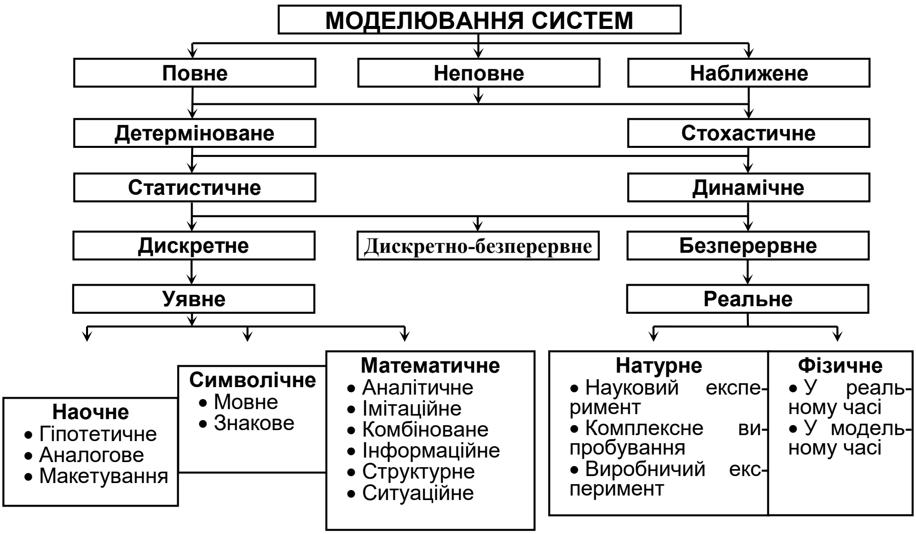


Рис. 3.2. Класифікація видів моделювання

Залежно від типу носія і сигнатури моделі розрізняють такі види моделювання: детерміноване і стохастичне, статичне і динамічне, дискретне, безперервне і дискретно-безперервне.

***Детерміноване*** моделювання відображає процеси, в яких передбачається відсутність випадкових впливів.

***Стохастичне*** моделювання враховує імовірнісні процеси та події.

***Статичне*** моделювання служить для опису стану об'єкта в фіксований момент часу, а динамічне – для дослідження об'єкта в часі. При цьому оперують аналоговими (безперервними), дискретними і змішаними моделями.

Залежно від форми реалізації носія і сигнатури моделювання класифікують на уявне і реальне.

***Уявне*** моделювання застосовується тоді, коли моделі не реалізовуються в заданому інтервалі часу або відсутні умови для їх фізичного створення (наприклад, ситуація мікросвіту). Уявне моделювання реальних систем реалізується у вигляді наочного, символічного і математичного. Для зображення функціональних, інформаційних і подієвих моделей цього виду моделювання розроблено значну кількість засобів і методів.

При ***наочному*** моделюванні на базі уявлень людини про реальні об'єкти створюються наочні моделі, що відображають явища і процеси, що протікають в об'єкті. Прикладом таких моделей є навчальні плакати, рисунки, схеми, діаграми.

В основу ***гіпотетичного*** моделювання закладається гіпотеза про закономірності перебігу процесу в реальному об'єкті, яка відображає рівень знань дослідника про об'єкт і базується на причинно-наслідкових зв'язках між входом і виходом досліджуваного об'єкта. Цей вид моделювання використовується, коли знань про об'єкт не достатньо для побудови формальних моделей. ***Аналогове*** моделювання ґрунтується на застосуванні аналогій різних рівнів. Для досить простих об'єктів найвищим рівнем є повна аналогія. З ускладненням системи використовуються аналогії таких рівнів, коли аналогова модель відображає кілька (або тільки одну) сторін функціонування об'єкта.

***Макетування*** застосовується, коли в реальному об'єкті процеси, що протікають, не піддаються фізичному моделюванню або можуть передувати проведенню інших видів моделювання. В основі побудови зображуваних макетів також покладені аналогії, що зазвичай базуються на причинно-наслідкових зв'язках між явищами і процесами в об'єкті.

***Символічне*** моделювання являє собою штучний процес створення логічного об'єкта, який заміщає реальний і виражає його основні властивості за допомогою певної системи знаків і символів.

В основі ***мовного*** моделювання покладений деякий тезаурус, що утворюється з набору понять досліджуваної предметної області, причому цей набір повинен бути фіксованим. Тезаурус розуміється як словник, що відображає зв'язки між словами чи іншими елементами цієї мови, призначений для пошуку слів за їх змістом.

Традиційний тезаурус складається з двох частин: списку слів і стійких словосполучень, згрупованих за смисловими (тематичними) рубриками; алфавітного словника ключових слів, які задають класи умовної еквівалентності, покажчика відносин між ключовими словами, де для кожного слова вказані відповідні рубрики. Така побудова дозволяє визначити семантичні (смислові) стосунки ієрархічного (рід/вид) і неієрархічного (синонімія, антонімія, асоціації) типу.

Між тезаурусом і звичайним словником є принципові відмінності. Тезаурус – словник, який очищено від неоднозначності, тобто в ньому кожному слову може відповідати лише єдине поняття, хоча в звичайному словнику одному слову може відповідати кілька понять.

Якщо ввести умовне позначення окремих понять, тобто позначки, а також певні операції між цими знаками, то можна реалізувати ***знакове*** моделювання і за допомогою знаків відображати набір понять – складати окремі ланцюжки зі слів і пропозицій. Використовуючи операції об'єднання, перетину і доповнення теорії множин, можна в окремих символах дати опис якогось реального об'єкта.

***Математичне моделювання*** – це процес встановлення відповідності даному реальному об'єкту деякого математичного об'єкта, що називається математичною моделлю. В принципі, для дослідження характеристик будь-якої системи математичними методами, включаючи і машинні, повинна бути обов'язково проведена формалізація цього процесу, тобто побудована математична модель. Вид математичної моделі залежить як від природи реального об'єкта, так і від завдань дослідження об'єкта, від необхідної достовірності і точності розв'язання задачі. Будь-яка математична модель, як і всяка інша, описує реальний об'єкт з деякою мірою наближення.

Для зображення математичних моделей можуть використовуватися різні форми запису. Основними є інваріантний, аналітичний, алгоритмічний і схематичний (графічний).

***Інваріантна*** форма – запис співвідношень моделі за допомогою традиційної математичної мови безвідносно до методу розв'язання рівнянь моделі. У цьому випадку модель може бути зображена як сукупність входів, виходів, змінних стану і глобальних рівнянь системи. *Аналітична* форма – запис моделі у вигляді результату рішення вихідних рівнянь моделі. Зазвичай моделі в аналітичній формі являють собою явні вирази вихідних параметрів як функцій входів і змінних стану.

Для ***аналітичного*** моделювання характерно те, що здебільшого моделюється тільки функціональний аспект системи. При цьому глобальні рівняння системи, що описують закон (алгоритм) її функціонування, записуються у вигляді деяких аналітичних співвідношень (алгебраїчних, інтегро-диференціальних, скінченнорізницевих і т.д.) або логічних умов. Аналітична модель досліджується кількома методами:

- аналітичним, коли прагнуть отримати в загальному вигляді явні залежності, що зв'язують шукані характеристики з початковими умовами, параметрами і змінними стану системи;
- числовим, коли, не вміючи розв'язувати рівняння в загальному вигляді, прагнуть отримати числові результати при конкретних початкових даних (нагадаємо, що такі моделі називаються цифровими);
- якісним, коли, не маючи розв'язку в явному вигляді, можна знайти деякі властивості розв'язку (наприклад, оцінити стійкість розв'язку).

В цей час поширені комп'ютерні методи дослідження характеристик процесу функціонування складних систем. Для реалізації математичної моделі на ЕОМ необхідно побудувати відповідний моделюючий алгоритм.

Алгоритмічна форма – запис співвідношень моделі й обраного числового методу розв'язання у формі алгоритму. Серед алгоритмічних моделей важливий клас становлять імітаційні моделі, призначені для імітації фізичних чи інформаційних процесів при різних зовнішніх впливах. Власне імітацію названих процесів називають імітаційним моделюванням.

При ***імітаційному*** моделюванні відтворюється алгоритм функціонування системи в часі – поведінка системи, причому імітуються елементарні явища, що становлять процес, зі збереженням їх логічної структури і послідовності перебігу, що дозволяє за вихідними даними отримати відомості про стани процесу в певні моменти часу, що дають можливість оцінити характеристики системи. Основною перевагою імітаційного моделювання порівняно з аналітичним є можливість розв'язання більш складних завдань. Імітаційні моделі дозволяють досить просто враховувати такі фактори, як наявність дискретних і безперервних елементів, нелінійні характеристики елементів системи, численні випадкові впливи та інші, які часто створюють труднощі при аналітичних дослідженнях. Сьогодні імітаційне моделювання – найбільш ефективний метод дослідження систем, а часто і єдиний практично доступний метод отримання інформації про поведінку системи, особливо на етапі її проектування.

В імітаційному моделюванні розрізняють:

- метод статистичних випробувань (Монте-Карло);
- метод статистичного моделювання.

Метод Монте-Карло – числовий метод, який застосовується для моделювання випадкових величин і функцій, імовірнісні характеристики яких збігаються з розв'язками аналітичних завдань. Полягає в багаторазовому відтворенні процесів, що є реалізаціями випадкових величин і функцій, з подальшою обробкою інформації методами математичної статистики.

Якщо цей прийом застосовується для машинної імітації з метою дослідження характеристик процесів функціонування систем, схильних до випадкових впливів, то такий метод називається методом статистичного моделювання.

Метод імітаційного моделювання застосовується для оцінки варіантів структури системи, ефективності різних алгоритмів керування системою, впливу зміни різних параметрів системи. Імітаційне моделювання може бути покладено в основу структурного, алгоритмічного і параметричного синтезу систем, коли потрібно створити систему із заданими характеристиками при певних обмеженнях.

***Комбіноване (аналітико-імітаційне)*** моделювання дозволяє об'єднати переваги аналітичного та імітаційного моделювання. При побудові комбінованих моделей проводиться попередня декомпозиція процесу функціонування об'єкта на складові підпроцеси, і для тих з них, де це можливо, використовуються аналітичні моделі, а для решти підпроцесів будуються імітаційні моделі. Такий підхід дає можливість охопити якісно нові класи систем, які не можуть бути досліджені з використанням аналітичного або імітаційного моделювання окремо.

***Інформаційне (кібернетичне)*** моделювання пов'язане з дослідженням моделей, в яких відсутня безпосередня подібність фізичних процесів, що відбуваються в моделях, реальним процесам. У цьому випадку прагнуть відобразити лише деяку функцію, розглядають реальний об'єкт як «чорний ящик», що має ряд входів і виходів, і моделюють деякі зв'язки між виходами і входами. Таким чином, в основі інформаційних (кібернетичних) моделей покладено відбиток деяких інформаційних процесів керування, що дозволяє оцінити поведінку реального об'єкта. Для побудови моделі в цьому випадку необхідно виділити досліджувану функцію реального об'єкта, спробувати формалізувати цю функцію у вигляді деяких операторів зв'язку між входом і виходом і відтворити дану функцію на імітаційній моделі, причому абсолютно іншою математичною мовою і, природно, іншою фізичною реалізацією процесу. Так, наприклад, експертні системи є моделями ОПР.

***Структурне*** моделювання системного аналізу базується на деяких специфічних особливостях структур певного виду, які використовуються як засіб дослідження систем або служать для розробки на їх основі специфічних підходів до моделювання із застосуванням інших методів формалізованого зображення систем (теоретико-множинних, лінгвістичних, кібернетичних і т.п.). Розвитком структурного моделювання є ***об'єктно-орієнтоване*** моделювання.

Структурне моделювання системного аналізу включає:

- методи мережевого моделювання;
- поєднання методів структуризації з лінгвістичними;
- структурний підхід у напрямку формалізації побудови і дослідження структур різного типу (ієрархічних, матричних, довільних графів) на основі теоретико-множинних уявлень і поняття номінальної шкали теорії вимірювань.

При цьому термін «структура моделі» може застосовуватися як до функцій, так і до елементів системи. Відповідні структури називаються функціональними і морфологічними. Об'єктно-орієнтоване моделювання об'єднує структури обох типів у ієрархію класів, що включають як елементи, так і функції.

***Ситуаційне*** моделювання спирається на модельну теорію мислення, в рамках якої можна описати основні механізми регулювання процесів прийняття рішень. У центрі модельної теорії мислення є уявлення про формування в структурах мозку інформаційної моделі об'єкта і зовнішнього світу. Ця інформація сприймається людиною на базі вже наявних у неї знань і досвіду. Доцільну поведінку людини будують шляхом формування цільової ситуації та уявного перетворення вихідної ситуації в цільову. Основою побудови моделі є опис об'єкта у вигляді сукупності елементів, пов'язаних між собою певними відношеннями, що відображають семантику предметної області. Модель об'єкта має багаторівневу структуру і являє собою той інформаційний контекст, на тлі якого протікають процеси керування. Чим багатша інформаційна модель об'єкта і вищі можливості маніпулювання нею, тим краща і різноманітніша якість прийнятих рішень при керуванні.

При ***реальному*** моделюванні використовується можливість дослідження характеристик або на реальному об'єкті цілком, або на його частини. Такі дослідження проводяться як на об'єктах, що працюють в нормальних режимах, так і при організації спеціальних режимів для оцінки характеристик, що цікавлять дослідника (при інших значеннях змінних і параметрів, в іншому масштабі часу і т.д.). Реальне моделювання є найбільш адекватним, але його можливості обмежені.

***Натурним*** моделюванням називають проведення дослідження на реальному об'єкті з наступною обробкою результатів експерименту на основі теорії подібності. Натурне моделювання поділяється на науковий експеримент, комплексні випробування та виробничий експеримент.

*Науковий експеримент* характеризується широким використанням засобів автоматизації, застосуванням досить різноманітних засобів обробки інформації, можливістю втручання людини в процес проведення експерименту. Один з різновидів експерименту – *комплексні випробування*, в процесі яких внаслідок повторення випробувань об'єктів у цілому (або великих частин системи) виявляються загальні закономірності про характеристики якості, надійності цих об'єктів. У цьому випадку моделювання здійснюється шляхом обробки та узагальнення відомостей про групу однорідних явищ. Поряд зі спеціально організованими випробуваннями можлива реалізація натурного моделювання шляхом узагальнення досвіду, накопиченого в ході виробничого процесу, тобто можна говорити про *виробничий експеримент*. Тут на базі теорії подібності обробляють статистичний матеріал з виробничого процесу і отримують його узагальнені характеристики. Необхідно пам'ятати про відміну експерименту від реального перебігу процесу. Вона полягає в тому, що в експерименті можуть з'явитися окремі критичні ситуації і визначитися межі стійкості процесу. Під час експерименту вводяться нові фактори впливу в процес функціонування об'єкта.

Іншим видом реального моделювання є ***фізичне***, відмінне від натурного тим, що дослідження проводиться в установках, які зберігають природу явищ і володіють фізичною подібністю. У процесі фізичного моделювання задаються деякі характеристики зовнішнього середовища і досліджується поведінка реального об'єкта, або його моделі при заданих або створюваних штучно впливах зовнішнього середовища.

## 3.4. Принципи і підходи до побудови математичних моделей

Математичне моделювання багато хто вважає швидше мистецтвом, ніж злагодженою і закінченою теорією. Тут дуже велика роль досвіду, інтуїції та інших інтелектуальних якостей людини. Тому неможливо написати досить формалізовану інструкцію, що визначає, як повинна будуватися модель тієї чи іншої системи. Проте відсутність точних правил не заважає досвідченим фахівцям будувати вдалі моделі. До теперішнього часу вже накопичено значний досвід, що дає підставу сформулювати деякі принципи і підходи до побудови моделей. При розгляді порізно кожен з них може здатися досить очевидним. Але сукупність взятих разом принципів і підходів далеко не тривіальна. Багато помилок і невдачі в практиці моделювання є прямим наслідком порушення цієї методології.

Принципи визначають ті загальні вимоги, яким повинна задовольняти правильно побудована модель. Розглянемо ці принципи.

1. **Адекватність**. Цей принцип передбачає відповідність моделі цілям дослідження за рівнем складності та організації, а також відповідність реальній системі щодо обраної множини властивостей. Доти, доки не вирішено питання, чи правильно відображає модель досліджувану систему, цінність моделі не значна.

2. **Відповідність моделі розв'язуваної задачі**. Модель повинна будуватися для розв'язання певного класу задач або конкретного завдання дослідження системи. Спроби створення універсальної моделі, націленої на рішення великої кількості різноманітних завдань, призводять до такого ускладнення, що вона виявляється практично непридатною. Досвід показує, що при розв'язанні кожної конкретної задачі потрібно мати свою модель, яка відображатиме ті аспекти системи, які є найбільш важливими в даній задачі. Цей принцип пов'язаний із принципом адекватності.

3. **Спрощення при збереженні істотних властивостей системи**. Модель повинна бути дещо простішою від прототипу – в цьому сенс моделювання. Чим складніша розглянута система, тим по можливості більш спрощеним має бути її опис, навмисно перебільшуючий типовий та ігноруючий менш істотні властивості. Цей принцип може бути названий принципом абстрагування від другорядних деталей.

4. **Відповідність між необхідною точністю результатів моделювання і складністю моделі**. Моделі за своєю природою завжди мають наближений характер. Виникає питання, яким має бути це наближення. З одного боку, щоб відобразити всі скільки-небудь істотні властивості, модель необхідно деталізувати. З іншого боку, будувати модель, близьку за складністю до реальної системи, очевидно, не має сенсу. Вона не повинна бути настільки складною, щоб існуюче рішення виявилося занадто складним. Компроміс між цими двома вимогами досягається нерідко шляхом проб і помилок. Практичними рекомендаціями щодо зменшення складності моделей є:

- зміна кількості змінних, що досягається або винятком несуттєвих змінних, або їх об'єднанням. Процес перетворення моделі в модель з меншою кільк істю змінних і обмежень називають агрегуванням. Наприклад, всі типи ЕОМ у моделі гетерогенних мереж можна об'єднати в чотири типи – ПЕОМ, робочі станції, великі ЕОМ (мейнфрейми), кластерні ЕОМ;
- зміна природи змінних параметрів. Змінні параметри розглядаються як постійні, дискретні – як безперервні і т.д. Так, умови поширення радіохвиль у моделі радіоканалу для простоти можна прийняти постійними;

➢ зміна функціональної залежності між змінними. Нелінійна залежність замінюється зазвичай лінійною, дискретна функція розподілу ймовірностей – безперервною;

➢ зміна обмежень (додавання, виключення або модифікація). При знятті обмежень виходить оптимістичне рішення, при введенні – песимістичне. Варіюючи обмеженнями, можна знайти можливі граничні значення ефективності. Такий прийом часто використовується для знаходження попередніх оцінок ефективності рішень на етапі постановки завдань;

➢ обмеження точності моделі. Точність результатів моделі не може бути вищою від точності вихідних даних.

5. **Баланс похибок різних видів**. Відповідно до принципу балансу необхідно домагатися, наприклад, балансу систематичної похибки моделювання за рахунок відхилення моделі від оригіналу і похибки вихідних даних, точності окремих елементів моделі, систематичної похибки моделювання і випадкової похибки при інтерпретації та усереднення результатів.

6. **Багатоваріантність реалізацій елементів моделі**. Різноманітність реалізацій одного і того ж елемента, що відрізняються за точністю (а отже, і за складністю), забезпечує регулювання співвідношення «точність/складність».

7. **Блочна будова**. При дотриманні принципу блочної будови полегшується розробка складних моделей і з'являється можливість використання накопиченого досвіду і готових блоків з мінімальними зв'язками між ними. Виділення блоків проводиться з урахуванням поділу моделі за етапами і режимами функціонування системи. Наприклад, при побудові моделі для системи радіорозвідки можна виділити модель роботи випромінювачів, модель виявлення випромінювачів, модель пеленгування і т.д.

Залежно від конкретної ситуації можливі такі підходи до побудови моделей:

➢ безпосередній аналіз функціонування системи;

➢ проведення обмеженого експерименту на самій системі;

➢ використання аналога;

➢ аналіз вихідних даних.

Є цілий ряд систем, які допускають проведення безпосередніх досліджень з виявлення істотних параметрів і відносин між ними. Потім або застосовуються відомі математичні моделі, або вони модифікуються або пропонується нова модель. Таким чином, наприклад, можна вести розробку моделі для направлення зв'язку в умовах мирного часу.

При проведенні експерименту виявляється значна частина істотних параметрів та їх вплив на ефективність системи. Таку мету ставлять, наприклад, усі командно-штабні гри і більшість навчань.

Якщо метод побудови моделі системи не ясний, але її структура очевидна, то можна скористатися схожістю з більш простою системою, модель для якої існує.

Побудову моделі можна розпочати на основі аналізу вихідних даних, які вже відомі або можуть бути отримані. Аналіз дозволяє сформулювати гіпотезу про структуру системи, яка потім апробується. Так з'являються перші моделі нового зразка іноземної техніки при наявності попередніх даних про їх технічні параметри.

Розробники моделей перебувають під дією двох взаємно суперечливих тенденцій:

- прагнення до повноти опису;
- прагнення до отримання необхідних результатів можливо більш простими засобами.

Досягнення компромісу ведеться зазвичай шляхом побудови серії моделей, що починаються з гранично простих і висхідних до високої складності (існує відоме правило: починай з простих моделей, а далі ускладнюй). Прості моделі допомагають глибше зрозуміти досліджувану проблему. Ускладнені моделі використовуються для аналізу впливу різних чинників на результати моделювання. Такий аналіз дозволяє виключати деякі чинники з розгляду.

Складні системи потребують розробки цілої ієрархії моделей, що розрізняються рівнем відображуваних операцій. Виділяють такі рівні, як вся система, підсистеми, керуючі об'єкти тощо.

## 3.5. Етапи побудови математичної моделі

Сутність побудови математичної моделі полягає в тому, що реальна система спрощується, схематизується і описується за допомогою того чи іншого математичного апарату. Можна виділити такі основні етапи побудови моделей.

1. ***Змістовний опис модельованого об'єкта.*** Об'єкти моделювання описуються з позицій системного підходу. Виходячи з мети дослідження, встановлюються сукупність елементів, взаємозв'язку між елементами, можливі стани кожного елемента, суттєві характеристики станів і відносини між ними. Наприклад, фіксується, що якщо значення одного параметра зростає, то значення іншого –

убуває і т.п. Питання, пов'язані з повнотою та одиничністю вибору характеристик, не розглядаються. Природно, в такому словесному описі можливі логічні протиріччя, невизначеності. Це вихідна природно-наукова концепція досліджуваного об'єкта. Таке попереднє, наближене уявлення системи називають концептуальною моделлю. Для того щоб змістовний опис служив хорошою основою для подальшої формалізації, потрібно докладно вивчити модельований об'єкт. Нерідко природне прагнення прискорити розробку моделі веде дослідника від цього етапу безпосередньо до вирішення формальних питань. У результаті побудована без достатнього змістовного базису модель виявляється непридатною до використання. На цьому етапі моделювання широко застосовуються якісні методи опису систем, знакові і мовні моделі.

2. ***Формалізація операцій.*** Формалізація зводиться в загальних рисах до такого. На основі змістовного опису визначається вихідна множина характеристик системи. Для виділення істотних характеристик необхідний хоча б наближений аналіз кожної з них. При проведенні аналізу спираються на постановку задачі і розуміння природи досліджуваної системи. Після виключення несуттєвих характеристик виділяють керовані та некеровані параметри і виробляють символізації. Потім визначається система обмежень на значення керованих параметрів. Якщо обмеження не мають принциповий характер, то ними нехтують.

Подальші дії пов'язані з формуванням цільової функції моделі. Згідно з відомими положеннями вибираються показники результату операції і визначається приблизний вигляд функції корисності на результатах. Якщо функція корисності близька до порогової (або монотонної), то оцінка ефективності рішень можлива безпосередньо за показниками результату операції. У цьому випадку необхідно вибрати спосіб згортки показників (спосіб переходу від множини показників до одного узагальненого показника) і провести саму згортку. Як згортка показників формуються критерій ефективності і цільова функція.

Якщо при якісному аналізі виду функції корисності виявиться, що її не можна вважати пороговою (монотонною), пряма оцінка ефективності рішень через показники результату операції не правомочна. Необхідно визначати функцію корисності і вже на її основі вести формування критерію ефективності та цільової функції.

В цілому заміна змістовного опису формальним – це ітеративний процес.

3. ***Перевірка адекватності моделі.*** Вимога адекватності перебуває у суперечності з вимогою простоти, і це потрібно враховувати при перевірці моделі на адекватність. Початковий варіант моделі попередньо перевіряється за такими основними аспектами:

- Чи всі суттєві параметри включені в модель?
- Чи немає в моделі несуттєвих параметрів?
- Чи правильно відображені функціональні зв'язки між параметрами?
- Чи правильно визначено обмеження на значення параметрів?

Для перевірки рекомендується залучати фахівців, які не брали участі в розробці моделі. Вони можуть більш об'єктивно розглянути модель і помітити її слабкі сторони, ніж її розробники. Така попередня перевірка моделі дозволяє виявити грубі помилки. Після цього розпочинають реалізацію моделі та проведення досліджень. Отримані результати моделювання піддаються аналізу на відповідність відомим властивостям досліджуваного об'єкта. Для встановлення відповідності створюваної моделі оригіналу використовуються такі шляхи:

- порівняння результатів моделювання з окремими експериментальними результатами, отриманими при однакових умовах;
- використання інших близьких моделей;
- зіставлення структури та функціонування моделі з прототипом.

Основним шляхом перевірки адекватності моделі досліджуваного об'єкта виступає практика. Однак вона потребує накопичення статистики, яка далеко не завжди буває достатньою для отримання надійних даних. Для багатьох моделей перші два прийнятні меншою мірою. У цьому випадку залишається один шлях: висновок про подібність моделі і прототипу робити на основі зіставлення їх структур і реалізованих функцій. Такі висновки не мають формального характеру, оскільки ґрунтуються на досвіді та інтуїції дослідника.

За результатами перевірки моделі на адекватність приймається рішення про можливість її практичного використання або про проведення коригування.

4. ***Коригування моделі.*** При коригуванні моделі можуть уточнюватися суттєві параметри, обмеження на значення керованих параметрів, показники результату операції, зв'язку показників результату операції з істотними параметрами, критерій ефективності. Після внесення змін до моделі знову виконується оцінка адекватності.

5. ***Оптимізація моделі***. Сутність оптимізації моделей полягає в їх спрощенні при заданому рівні адекватності. Основними показниками, за якими можлива оптимізація моделі, виступають час і витрати коштів для проведення досліджень на ній. В основі оптимізації лежить можливість перетворення моделей з однієї форми в іншу. Перетворення може виконуватися або з використанням математичних методів, або евристичним шляхом.

## 3.6. Технологія моделювання

Розширенню сфер застосування моделювання сприяли, звичайно, розвиток і застосування обчислювальних машин та систем. Із упевненістю можна сказати, що не існує сфер людської діяльності, де б не використовувалося моделювання. Це і функціонування атомних реакторів, і дослідження макроекономічних процесів, і виробництво машин та агрегатів, і розвиток біологічних систем, і аналіз наслідків використання природних ресурсів та екологічних катастроф.

Самі обчислювальні комплекси стали не тільки інструментом дослідження, але й об'єктами моделювання. Обчислювальні машини, комплекси, системи та мережі завдяки своїй складності й дорожнечі, природно, є об'єктами моделювання. При цьому моделювання доцільне як на етапах проектування обчислювальних комплексів, так і під час аналізу функціонування діючих систем в екстремальних умовах або при зміні їх складу, структури, способів керування або робочого навантаження.

Застосування моделювання на етапі проектування дозволяє аналізувати варіанти проектних рішень, визначати працездатність і продуктивність, виявляти дефіцитні й малозавантажені ресурси, обчислювати очікувані часи реакції та ухвалювати рішення щодо раціональної зміни складу й структури комплексу або за способом організації обчислювального процесу.

Доцільним є використання моделювання для діючих обчислювальних комплексів, оскільки можна дослідним шляхом перевірити адекватність моделі й оригіналу та більш точно визначити ті параметри системи й зовнішні впливи на неї, які служать вихідними даними для моделювання. Моделювання реальних обчислювальних комплексів дозволяє виявити його резерви та спрогнозувати якість функціонування за будь-яких змін, тому корисно мати моделі всіх систем, що розвиваються.

Здійснення моделювання передбачає конкретизацію його мети, створення моделі, проведення її дослідження й аналіз отриманих результатів. Навіть процес створення моделі складається з декількох етапів. Він починається з вивчення системи та зовнішніх впливів, а завершується розробкою або вибором математичної моделі і програмного забезпечення. Природно, що деякі класи математичних моделей можуть бути досліджені й без застосування комп'ютерної техніки.

Таким чином, моделювання припускає наявність таких укрупнених етапів:

- постановка мети моделювання;
- розробка концептуальної моделі;
- підготовка вихідних даних;

- розробка математичної моделі;
- вибір методу моделювання;
- вибір засобів моделювання;
- розробка програмного забезпечення;
- перевірка адекватності й коректування моделі;
- планування машинних експериментів;
- моделювання на обчислювальному комплексі;
- аналіз результатів моделювання.

Наведений перелік, звичайно, не є догмою. Різні дослідники відповідно до свого досвіду й уподобань змінюють зміст і кількість етапів, але логіка послідовності повинна зберігатися.

**Постановка мети моделювання.** Будь-яка система може являти собою певний набір моделей, що відрізняються одна від одної. Відмітності можуть міститися в ступені деталізації й обліку різних особливостей і режимів функціонування. Можуть також відображатися деякі грані сутності системи, можливе орієнтування на аналіз деяких наборів властивостей. Тому розробці моделі, природно, передує постановка (формулювання) мети моделювання.

Як підійти до розв’язання цієї задачі? Оскільки створення моделі зазвичай здійснюється при проектуванні або модернізації системи, то природно виникають задачі визначення її (системи) ефективності, що розв’язуються до моделювання. Цей етап закінчується ухваленням рішення про доцільність або навпаки – недоцільність проведення моделювання.

Необхідно пам’ятати, що подібність процесу, який протікає в моделі, до реального процесу не може бути метою. Це тільки умова правильного функціонування моделі, та й то не завжди.

Визначення мети моделювання – це розрахунок значень вибраного показника ефективності для різних варіантів реалізації проектованої або досліджуваної системи. Наприклад, при визначенні варіанта побудови комп’ютерної мережі, яка б мала мінімальну вартість при дотриманні вимог щодо продуктивності й надійності, метою моделювання може бути пошук параметрів мережі, що забезпечують мінімальне значення показника ефективності. Одним із цих параметрів може бути вартість.

Задача може формулюватися інакше. З декількох варіантів конфігурації комп’ютерної мережі необхідно вибрати найбільш надійний. У цьому випадку як показник ефективності можна взяти один із показників надійності (середнє напрацювання на відмову, імовірність безвідмовної роботи та ін.), а метою моделювання є порівняльна оцінка варіантів мережі за цим показником.

Одночасно із цим розглянута комп'ютерна мережа може використовуватися в системах навчання й керування набором студентів, у системі науково-технічних розробок та ін. Ця ж мережа може бути елементом систем енергопостачання, технічного обслуговування й постачання. У цьому випадку показник ефективності комп'ютерної мережі неможливо оцінити за допомогою одиниць СІ.

Тоді використовують показник техніко-економічної ефективності $E$, який ураховує не тільки витрати, але й деякі вимірювані вихідні характеристики системи

$$E = E(W), \tag{3.6}$$

де елементи множини характеристик $w_s \in W(s = 1, 2, \dots, n_w)$ – часткові показники якості системи (продуктивність, надійність, вартість, маса, габарити та ін.).

Якщо відомі аналітичні співвідношення (3.1) і (3.2), то показник ефективності $E$ обчислюється елементарно за сукупністю внутрішніх параметрів системи $H$ при певних зовнішніх впливах $V$, або навпаки – за заданим показником ефективності обчислюються необхідні параметри системи.

За відсутності аналітичних залежностей (3.1) і (3.2) використовують підходи, які називаються однокритеріальною або багатокритеріальною оцінками.

**Однокритеріальна оцінка.** Проводиться оцінка ефективності за одним частинним показником якості $w^0$, а на припустимі значення інших елементів множини $W$ накладаються обмеження

$$E = E(w^0),$$
$$w_{s_{\min}} \le w_s \le w_{s_{\max}} E(w^0), \tag{3.7}$$

де $w_{s_{\min}}$ і $w_{s_{\max}}$ – граничні значення $s$-ого частинного показника якості.

Особливістю такого підходу до оцінки ефективності є можливість одержання декількох варіантів системи для певного значення $w^0$ при різних значеннях інших частинних показників якості, що задовольняють нерівності (3.7).

**Багатокритеріальна оцінка.** У цьому випадку найпоширенішим підходом є подання невідомої функції (3.2) за допомогою нормованого адитивного критерію:

$$E = \sum_{s=1}^{S_w} g_s \, \gamma(w_s), \tag{3.8}$$

де $\gamma(w_s)$ – значення оцінки частинного показника якості, що забезпечують виключення розмірності й підібрані так, щоб $(w_s) \in [0,1]$; $g_s$ – вагові коефіцієнти, які враховують важливість $s$-ого показника та задовольняють умові

$$\sum_{s=1}^{s_w} g_s = 1, g_s > 0.$$

**Уточнення мети моделювання.** Таким чином, на основі аналізу взаємодії системи з навколишнім середовищем і попереднього вивчення внутрішніх структурних та функціональних особливостей визначається мета моделювання шляхом визначення виду критерію ефективності системи. Потім виключаються ті характеристики, які можна визначити без моделювання. Підсумкова множина зовнішніх впливів, які враховуються, повинна включати тільки сприятливі (корисні впливи) або перешкоджаючі (збурюючі впливи) функціонуванню системи. Але, природно, при моделюванні не вдається та й недоцільно включати все різноманіття збурюючих факторів. Тому це часто призводить до ідеалізації умов і, можливо, до викривлення результатів моделювання.

Якщо моделювання має на меті не тільки фіксації властивостей, але й оптимізації характеристик, то необхідно виявити ті параметри, які можна змінювати в процесі моделювання. Коли ставиться задача визначення залежності характеристик від деяких параметрів системи (3.1), то ці характеристики й параметри повинні бути точно зафіксовані.

Оцінка доцільності здійснення моделювання є завершальним моментом етапу визначення мети моделювання. Проведення моделювання потребує деяких витрат: чим складнішою є система, й чим вищими є вимоги до точності результатів, тим більші витрати. Якщо метою є обґрунтування працездатності системи, то економічний ефект від експлуатації цієї системи повинен перевищувати витрати на моделювання. При оптимізації системи витрати на моделювання мають окупатися за рахунок різниці між найкращим і найгіршим варіантами системи.

Якщо результат моделювання показує, що досліджуваний варіант системи є прийнятним (оптимальним), то не доцільно ухвалювати рішення щодо даремних витрат на моделювання. Важливо, що моделювання підтвердило правильність прийнятих проектних рішень.

**Розробка концептуальної моделі.** Концептуальна (змістовна) модель – це абстрактна модель, яка визначає склад і структуру системи, властивості її елементів і причинно-наслідкові зв’язки, властиві досліджуваній системі й істотні для досягнення мети моделювання.

Як бачимо, це всього-на-всього словесний опис природи, параметрів і умов взаємодії окремих компонентів системи. Це первинне представлення моделі, можливо, в уяві дослідника. Водночас концептуальна модель – це в деякому сенсі субстрат системи, що забезпечує досягнення мети моделювання.

Необхідність обґрунтування включення в модель певних важливих елементів і властивостей, а також виключення з моделі несуттєвих властивостей, потребує глибоких знань про саму систему. Це протиріччя (включення одних елементів і виключення інших) призводить до певних труднощів, оскільки виявлення впливу виключення того або іншого фактора на ступінь викривлення результатів вимагає створення вже двох моделей: з урахуванням і без урахування цього фактора. А оскільки часто кількість елементів буває надзвичайно великою, то й кількість моделей може збільшуватися лавиноподібно, призводячи до значного зростання витрат.

***Простота моделі та її адекватність досліджуваній системі*** – це той рідкий та з труднощами знайдений компроміс, який повинен бути досягнений при моделюванні.

Відповідальним за це є розробник. Саме він на основі своїх знань і досвіду приймає рішення про виключення з моделі деякого елемента або набору елементів без повної впевненості в тому, що ці дії не призведуть до одержання значної похибки при моделюванні. Тому часто говорять, що моделювання є не тільки наукою, а й мистецтвом.

**Розробка статичної складової концептуальної моделі.** Створення розробником концептуальної моделі можна формалізувати, тобто розбити на дві такі дії: розробку статичної й динамічної моделей. На етапі створення статичної моделі здійснюються стратифікація, деталізація, локалізація й структуризація.

**Стратифікація**. Обов'язковими властивостями системи (і обчислювальної в тому числі) є, з одного боку, її цілісність, а з іншого – здатність до розбивки системи на сукупність елементів. У свою чергу й модель – це сукупність частин, що забезпечують збереження цілісності системи.

Поділ системи на сукупність елементів (членованість системи та її елементів) може привести до побудови ієрархічної послідовності моделей. У підсумку система подається у вигляді набору моделей, що відображають її поведінку на різних рівнях декомпозиції (ці рівні називаються стратами). Кожний рівень ураховує притаманні йому властивості, змінні й залежності, а процес виділення рівнів називається стратифікацією. У модель, як правило, включають елементи одного рівня деталізації – $K$-ої страти. Іноді, у разі появи труднощів в описі елементів, у модель включають їх деталізоване подання з нижньої $(K-1)$-ої страти.

Побудова стратифікованої концептуальної моделі передбачає включення в неї параметрів, за допомогою яких можливе варіювання властивостей при моделюванні, які забезпечують визначення необхідних характеристик при конкретних зовнішніх впливах на заданому часовому інтервалі функціонування системи. Інші параметри без жалю виключаються з моделі.

**Деталізація**. Функціонування будь-якої системи – це реалізація певної кількості технологічних процесів перетворення речовини, енергії або інформації. Ці процеси складаються з послідовності елементарних операцій, а виконання кожної елементарної операції забезпечується певним ресурсом. Звідси випливає, що в моделі повинні бути наявними складові елементи, які реалізують виконання всіх технологічних процесів. Крім того, у модель можуть включатися елементи, що служать управлінню ресурсами й процесами, а також елементи, призначені для зберігання інформації, необхідної для керування.

Глибина деталізації здійснюється до такого рівня, щоб були залежності параметрів вихідних впливів для кожного елемента. Ці параметри повинні бути істотними для функціонування системи й визначення її вихідних характеристик від параметрів вхідних впливів.

Підвищення рівня деталізації опису системи приводить до одержання більш точної її моделі. Але це ускладнює процес моделювання й веде до зростання часових витрат на його здійснення.

З наведеного випливає правило: у модель повинні бути включені всі параметри, які забезпечують визначення характеристик, що цікавлять дослідника, системи на заданому часовому інтервалі її функціонування; інші параметри, за можливості, виключаються з моделі.

**Локалізація**. Це зображення зовнішнього середовища у вигляді генераторів зовнішніх впливів, що включаються до складу моделі як елементи (рис. 3.3). Їх можна підрозділити на:

– генератори робочого навантаження, що поставляють на вхід системи речовину, енергію або дані;

– генератори керуючих і збуруючих впливів.

Результати функціонування системи підрозділяють на:

– продукти перетворення;

– інформацію про стан системи;

– керуючі впливи на інші системи.

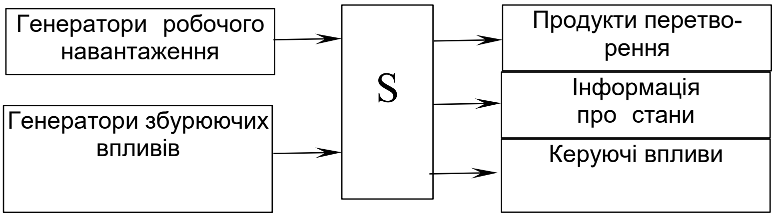


Рис. 3.3. Зображення зовнішнього середовища системи у вигляді генераторів зовнішніх впливів

**Структуризація**. Це перелік зв'язків між елементами моделі. Зв'язки бувають дійсні та інформаційні. Дійсні зв'язки показують шляхи можливого переміщення продуктів перетворення від елемента до елемента. За допомогою інформаційних зв'язків здійснюється передача між елементами керуючих впливів та інформації про стани. Ці зв'язки можуть і не бути в системі якимось матеріальним каналом.

У випадку, якщо система складається з однофункціональних елементів з одним вихідним дійсним зв'язком, інформаційні зв'язки, природно, відсутні, а керування функціонуванням проводиться самою структурою. Це є принципом структурного керування. Як приклад подібних систем зазвичай наводять логічні елементи обчислювальних комплексів.

Якщо система включає в себе елементи з декількома вихідними дійсними зв'язками і є керуючі засоби й інформаційні зв'язки, то вона функціонує відповідно до програмного або алгоритмічного керування. У цьому випадку керування потрібне для визначення самого елемента для того, яку операцію перетворення виконати й куди передати.

Концептуальна модель конкретно містить перелік і зміст усіх правил та алгоритмів керування робочим навантаженням, елементами й процесами.

**Розробка динамічної складової концептуальної моделі.** Усе перераховане вище спрямоване на розробку статичної моделі, тобто для відображення її складу і структури. А для опису динаміки функціонування системи необхідно враховувати алгоритми керування й параметри вхідних впливів та елементів.

Як уже згадувалося, у процесі функціонування системи реалізується технологічний процес перетворення речовини, енергії й інформації. В обчислювальних машинах реалізується мультипрограмний режим роботи, а технологічний процес у них реалізує деяку послідовність елементарних операцій, частина з яких виконується паралельно різними елементами (ресурсами).

У загальному випадку існує два підходи до реалізації процедур керування технологічним процесом. Це програмний і структурний підходи.

Технологічний процес, що задається програмою, реалізується деяким аналогом алгоритму керування. Алгоритм визначає, якими є ресурси системи, в якій послідовності вони знаходяться, які операції повинні виконуватися для досягнення мети функціонування.

Існує інтуїтивне уявлення про алгоритм як про формальну сукупність операцій і порядку їх виконання для розв'язання задач якого-небудь типу.

Алгоритм зазвичай розуміється як точне розпорядження, що визначає процес переробки вихідних даних у необхідний результат. Будь-який алгоритм характеризується масовістю, детермінованістю і результативністю.

У випадку використання структурного принципу керування для кожного елемента системи вибираються параметри, які змінюються в часі й, відповідно, відображають перебіг технологічного процесу. Множина цих параметрів відображає стан системи, а функціонування системи подається у вигляді послідовної зміни станів. Множину можливих станів системи називають простором станів (3.1).

На завершальному етапі створення концептуальної моделі необхідна перевірка її адекватності вихідній системі. Але оскільки самої моделі у фізичному сенсі поки не існує, таку перевірку повинен здійснити експерт, але, у жодному разі, не розробник моделі.

**Підготовка вихідних даних.** Потрібно пам'ятати, що розробник у процесі створення моделі взаємодіє з багатьма фахівцями, знайомими із самою системою. А оскільки вимоги до моделі можуть змінюватися в процесі вивчення системи, то розробнику моделі може знадобитися постійна консультація з фахівцями з досліджуваних питань. Головне, що при створенні моделі розробник має використовувати всю доступну інформацію, яка може бути отримана різними методами. Ці методи є такими:

– консультації з багатьма фахівцями з досліджуваного питання;

– спостереження за системою (якщо система, подібна до досліджуваної, уже існує, то зібрані дані будуть корисними при моделюванні).

Однак при використанні таких даних можуть виникнути деякі незручності:

- дані не являють собою те, що в дійсності необхідно моделювати;
- дані не мають підходящого типу або формату;
- у даних можуть бути помилки вимірювання, записи або округлення;
- дані можуть бути необ'єктивними для одержання певної вигоди;
- дані можуть бути подані в несумісних одиницях;

➢ результати, отримані під час моделювання подібних систем;

➢ досвід та інтуїція розробника – особливо, якщо система є подібною до тієї, що моделюється, що не існує на даний момент.

Таким чином, збір, підготовка, первинна обробка вихідних даних (якісних і кількісних) проводиться протягом декількох вихідних етапів процесу розробки моделі. Дані можуть постійно коректуватися й уточнюватися.

Кількісні параметри системи повинні мати конкретні значення, оскільки вони вирішально впливають на успіх моделювання. Точність і повнота вихідних даних багато в чому визначають вірогідність результатів моделювання.

Параметри моделі (вихідні дані) у загальному випадку можуть бути детермінованими або стохастичними. З іншого боку, вони часто виявляються нестаціонарними. І оскільки в більшості випадків ці дані повинні характеризувати (описувати) проектовану (не існуючу) систему або систему, що модернізується, то процедура збору вихідних даних перетворюється в дуже складну й дуже важливу проблему.

Більшість параметрів, за своєю природою, будучи випадковими величинами, при моделюванні часто подається у вигляді детермінованих середніх значень. Це припустимо в тих випадках, коли випадкові величини мають малий розкид, або коли припустимо враховувати в моделі тільки середні значення. Але це завжди веде до появи певної похибки моделювання, оскільки вплив випадкових факторів може призвести не тільки до розсіювання, але й до зсуву середніх значень результатів.

Іноді можливий і зворотний підхід, коли детерміновані характеристики заміняються випадковими величинами. Зазвичай це спрямовано на скорочення обсягів вихідних даних. Нехай, наприклад, моделюється обробка даних обчислювальною системою. У цьому випадку для багаторазових процесів реалізації програм із різними обсягами вихідних даних можна задавати всю сукупність кількостей цих даних. Також можна їх замінити випадковою величиною із заданим законом розподілу (методика реалізації випадкових величин показана в другій частині навчального посібника).

**Підбір закону розподілу.** Для систематизації значень випадкових параметрів необхідний збір статистичних даних і їх обробка для визначення можливості подання параметрів, у вигляді якого-небудь теоретичного закону розподілу.

Процедура підбору виду закону розподілу полягає в такому. За сукупністю числових значень параметра будується гістограма відносних частот – емпірична щільність розподілу. Гістограма апроксимується плавною кривою. Отримана крива послідовно порівнюється з кривими щільності розподілу різних теоретичних законів розподілу. Вибирається один із законів за найкращим збігом виду порівнюваних кривих. За емпіричним значенням обчислюють параметри цього розподілу. Потім виконують кількісну оцінку ступеня збігу емпіричного й теоретичного розподілу за тим або іншим критерієм згоди (наприклад, Пірсона (*хі*-квадрат), Колмогорова, Смирнова та ін.). Перелічені етапи досить детально розроблені й описані в математичній статистиці.

З іншого боку, вихідні дані можуть бути подані вже в обробленому вигляді. При цьому теж виникають проблеми: розподіл побудований за малою вибіркою, інформація подана в якісній формі, є тільки агреговані оцінки, дані застаріли або отримані в іншому місці та за інших умов, у вибірці відсутня частина даних.

**Апроксимація функцій і висування гіпотез**. Кожному елементу системи в будь-який момент часу можна поставити у відповідність функціональний зв'язок між параметрами вхідних впливів і його вихідними характеристиками. Ця функціональна залежність часто є очевидною, або іноді легко визначається в результаті аналізу природи функціонування системи.

Для деяких елементів вдається тільки одержати набір експериментальних даних про кількісні значення вихідних характеристик при різних значеннях параметрів. У цьому випадку висувається деяка гіпотеза про характер функціональної залежності, тобто здійснюється апроксимація цієї залежності певним математичним рівнянням. Для цього використовуються методи регресійного, кореляційного або дисперсійного аналізів.

Вид рівняння може задаватися розробником. Найбільш просто це здійснюється для двох змінних за результатами порівняння графіка з експериментальними точками й графіків розповсюджених апроксимуючих функцій (прямої, параболи, гіперболи, експоненти та ін.). Потім, для забезпечення найкращого наближення кривої до експериментальних даних, використовуючи регресійний аналіз, обчислюються значення констант вибраного рівняння. Зазвичай наближення оцінюються за критерієм найменших квадратів.

Якщо параметри відбивають нові елементи створюваної системи або нові умови її функціонування, то немає можливості зібрати фактичні дані. У цьому випадку залучаються фахівці, які добре уявляють обговорювану проблему і здатні висунути гіпотези про можливі значення таких параметрів. Ступінь суб'єктивності може бути зменшений, якщо скористатися методиками експертних оцінок.

Етап збору й обробки вихідних даних завершується класифікацією їх на зовнішні і внутрішні, постійні й змінні, безперервні та дискретні, лінійні й нелінійні, стаціонарні та нестаціонарні, детерміновані й стохастичні. Крім того, визначаються цілі зміни для змінних кількісних параметрів, а для дискретних величин – їх можливі значення.

**Розробка математичної моделі.** *Математична модель* – це сукупність математичних об'єктів і відношень, які відображають об'єкти й відношення деякої предметної області (області реального світу). Відомо, що тими самими математичними моделями можуть описуватися зовсім різні за природою процеси.

Основою для розробки математичної моделі служать концептуальна модель і кількісні вихідні дані. Переслідуються дві мети:

- одержати однозначний формалізований опис структури й процесу функціонування системи;
- представити процес функціонування з використанням однієї з можливих узагальнених формалізованих схем.

Цими найпоширенішими схемами є, наприклад, безперервні детерміновані системи, автомати, агрегативні системи, системи масового обслуговування та ін.

**Безперервні детерміновані системи**. Найпростішим видом моделей таких систем є лінійне співвідношення (пряма пропорційна залежність) між двома числовими змінними $y = kx$, де $k$ – коефіцієнт (числовий параметр), що відображає властивості моделі. Використання такої залежності дозволяє описувати багато процесів у реальних системах (це і закон Ньютона в механіці, і закон Ома в електротехніці).

У більш загальному випадку, якщо не враховується вплив випадкових факторів, а малі зміни вхідних впливів приводять до такого ж порядку малих змін вихідного впливу й станів системи, співвідношення (3.1) можна подати у вигляді векторного диференціального рівняння

$$\frac{dy}{dt} = \mathbf{F}[x(t), \vartheta(t), h(t), t], \tag{3.9}$$

де $\mathbf{F}$ – вектор-функція закону функціонування системи; $x, v, h, y$ – вектори вхідних, внутрішніх і вихідних впливів відповідно.

У випадку лінійності систем, коли змінні мають властивість однорідності й адитивності, вид рівнянь (3.9) спрощується, що дозволяє розв'язувати їх аналітичними або числовими (наближеними) методами.

**Автомати**. Автоматні моделі використовуються для аналізу функціонування й проектування окремих вузлів обчислювальних систем, для розробки програмного забезпечення, а також можуть бути використані при створенні моделей в інших галузях техніки, біології, медицини та ін.

Автомат подають як деяке обладнання (чорний ящик), на яке надходять зовнішні впливи (вхідні сигнали), що змінюють внутрішній стан автомата, причому реакція автомата на кожний вплив (вихідний сигнал автомата) залежить як від конкретного значення впливу, так і від стану автомата.

Множини дискретних впливів, реакцій і станів зображаються в теорії абстрактних автоматів за допомогою відповідних алфавітів: алфавіту (множини) вхідних символів $V = \{v\}$, алфавіту (множини) вихідних символів $W = \{w\}$, а також алфавіту (множини) станів автомата $S = \{s\}$. Кожний вхідний символ $v \in V$ відображає вхідний вплив, а вихідний символ $w \in W$ – реакцію автомата.

Припускається, що множини $V$ і $W$ є скінченними. Якщо множина $S$ є скінченою, то автомат називається скінченним, а якщо $S$ не скінченна, то автомат називається автоматом із нескінченною кількістю станів. Якщо автомат має всього лише один стан, то він називається тривіальним або автоматом без пам'яті.

У процесі функціонування в результаті чергового вхідного впливу автомат із пам'яттю змінює свій стан і виробляє вихідну реакцію в дискретні моменти часу.

Оскільки конкретні значення часу в моделі абстрактного автомата несуттєві, використовують дискретний автоматний час *t,* що набуває послідовно цілочислові значення $0, 1, 2, \ldots$, причому $t = 0$ відповідає початковому моменту, з якого розглядається функціонування автомата.

Мережі Петрі призначені для моделювання дискретних синхронних процесів. Їх основні відмінності від автоматної моделі – можливість відображати паралелізм, асинхронність та ієрархічність об'єктів, що моделюються.

**Агрегативні системи**. Це дуже поширений підхід, що дозволяє описувати функціонування безперервних і дискретних, детермінованих і стохастичних систем. Агрегат подається як основний елемент будь-якої системи (це агрегатне подання системи). Загальну схему агрегату зображено на рис. 3.4.

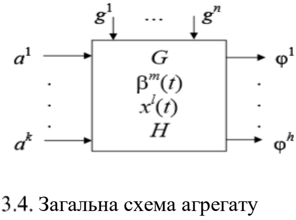


Рис. 3.4. Загальна схема агрегату

Властивості й функціонування будь-якого агрегату визначають такі фактори:

1. Вхідна інформація, оброблювана агрегатом. Ця інформація являє собою повідомлення вигляду ($k = 1, 2, \dots, k^*$), що надходять в моменти часу $t_j$ ($j = = 1, 2, \dots, j^*$). У загальному випадку вхідна інформація може бути функцією часу й подана $(k^* + 1)$-вимірним вектором

$$\bar{a}_j^k = \bar{a}_j^k\left(t_j, a_j^1, a_j^2, \dots, a_j^{k^*}\right).$$

Кожне з повідомлень $\bar{a}_j^k$ залежно від виду агрегату може характеризуватися законом розподілу моментів часу вступу, типом, пріоритетом, довжиною, параметром обслуговування та ін.

2. Вхідна інформація, що керує функціонуванням агрегату. Ця інформація подана повідомленнями у вигляді $g^n (n = 1,2, \dots, n^*)$, що надходять у моменти часу $\tau_j$ ($i = 1, 2, \dots, i^*$), і визначає зміни алгоритму $G$. Ці повідомлення також є функціями часу й можуть бути задані $(n^* + 1)$-вимірним вектором

$$\bar{g}_i^n = \bar{a}_i^n\left(t_i, g_i^1, g_i^2, \dots, g_i^{n^*}\right).$$

Повідомлення, що керують функціонуванням агрегату, можуть характеризуватися законом розподілу моментів вступу, тривалістю, типом і іншими факторами.

3. Координати агрегату, тобто сукупність деяких величин $x^l (l = 1,2, \dots, l^*)$, які визначають стан агрегату в кожний момент часу $t$. Координати агрегату також можна навести у вигляді $(l^* + 1)$-вимірного вектора

$$\bar{x}_i^l = \bar{x}_i^l\left(t_i, g_i^1, g_i^2, \dots, g_i^{l^*}\right).$$

Залежно від виду агрегату координати також можуть бути розбиті на ряд груп залежно від їхнього типу.

4. Вихідна інформація агрегату, що видається ним в інші агрегати або зовнішнє середовище. Ця інформація подається повідомленнями вигляду $\varphi_r^h (h = 1,2, \dots, h^*)$, що видаються агрегатом у моменти часу $t_r (r = 1,2, \dots, r^*)$. Вона також є функцією часу й може бути подана $(h^* + 1)$-вимірним вектором

$$\bar{\varphi}_r^h = \bar{x}_r^l\left(t_r, \varphi_r^1, \varphi_r^2, \dots, \varphi_r^{h^*}\right).$$

Повідомлення вихідної інформації також можуть характеризуватися типом, пріоритетом, довжиною й іншими факторами.

5. Параметри агрегату, що являють собою сукупність величин $\beta^m(t)$, які є вихідною інформацією агрегату й визначають значення його вектора $\bar{x}^l$.

6. Алгоритм $G$, реалізований агрегатом, який залежно від $t_j$, $\bar{a}_j^k$ і $\bar{x}_i^l(t)$ визначає $\bar{\varphi}_r^h$ і нові значення параметрів $\beta^m(t)$, (алгоритм виходів).

7. Алгоритм $H$, що дозволяє за відомими параметрами $\beta^m(t)$, а також значеннями величин $t_j$, $\bar{a}_j^k$ визначити координати агрегату $\bar{x}_i^l(t)$ для будь-якого моменту часу $t$ (алгоритм переходів).

Таким чином, агрегат можна визначити як об'єкт

$$A = A\{\bar{a}^k, \bar{\varphi}^h, \bar{x}^l, G[\bar{g}^n], H[\bar{\beta}^m], t\},$$

який складається з непустих множин

$$\{\bar{a}^k\}, \{\bar{\varphi}^h\}, \{\bar{x}^l\}, \{\bar{g}^n\}, \{\bar{\beta}^m\},$$

і алгоритмів $G$ і $H$, визначених відповідно на множинах

$$\{\bar{a}^k\}, \{\bar{x}^l\}, \{\bar{g}^n\}, \text{ і } \{\bar{a}^k\}, \{\bar{\beta}^m\},$$

де $\{\bar{a}^k\}$ – множина вхідних сигналів; $\{\bar{g}^n\}$ – множина керуючих сигналів; $\{\bar{x}^l\}$ – множина станів агрегату; $\{\bar{\beta}^m\}$ – множина його параметрів; $\{\bar{\varphi}^h\}$, – множина вихідних сигналів.

Викладене описове визначення агрегату є загальним. На практиці часто доводиться мати справу з агрегатами більш простого типу. Їхнє функціонування й властивості описуються лише частиною перелічених факторів, тобто вони будуть окремими випадками описаного агрегату.

Відомі роботи, спрямовані на оптимізацію конкретних систем агрегатним поданням їх елементів.

**Системи масового обслуговування**. В основу побудови моделей систем масового обслуговування (СМО) покладена формалізована схема, що передбачає наявність обслуговуючих приладів (каналів обслуговування), вхідного потоку заявок на обслуговування, можливі черги із заявок, що очікують початку обслуговування, і вихідного потоку обслужених заявок або заявок, які одержали відмову. Розроблена й широко використовується в повсякденній практиці для створення й дослідження математичних моделей відповідна теорія – теорія масового обслуговування. Це дозволяє в багатьох випадках широко використовувати СМО для побудови моделей.

Таким чином, для побудови математичної моделі може використовуватися одна з описаних вище формалізованих схем. Вибір тієї або іншої схеми здійснюється в результаті аналізу концептуальної моделі та вихідних даних

## Контрольні запитання

1. Що є об'єктом моделювання?
2. Що таке модель системи?
3. Назвіть ознаки, за якими класифікують моделі.
4. Дайте визначення математичної моделі систем.
5. Що є моделюванням систем?
6. Що є предметом теорії моделювання?
7. Що є законом функціонування системи і в якому вигляді він може бути заданий?
8. Дайте визначення алгоритму функціонування системи.
9. Як визначають мету моделювання?
10. Назвіть основні принципи, за якими здійснюється побудова математичних моделей.
11. Дайте характеристику основних підходів до побудови математичних моделей.
12. Назвіть основні етапи побудови математичної моделі.
13. Дайте характеристику основним етапам моделювання систем.

# 4. МОДЕЛЮВАННЯ СИСТЕМ У ВИГЛЯДІ «ЧОРНОГО ЯЩИКА»

## 4.1. Модель «Чорного ящика»

У теорії досліджень систем і на практиці часто достатньо мати лише частину інформації об'єкта, що досліджується. Наприклад, коли ми не знаємо поточного числового значення точного часу (проблема – незнання точного часу, мета – не спізнитися будь-куди), потрібно просто подивитися на годинник і не треба думати про його внутрішню роботу і джерело енергії роботи.

У попередньому прикладі призначення годинників (метою їхнього існування) є показати точний час у будь-який момент і тим самим впливати на зовнішнє для нього середовище.

Якщо дотримуватися першого визначення системи, то вона є засобом і, таким чином, існують можливості впливати на це з зовнішнього середовища (уточнювати хід, забезпечити енергію, дивитися і т. д.).

Склад системи при цьому не відомий (або не є цікавим зовнішньому середовищу), але цього достатньо, щоб вирішити це завдання.

Іншими словами, важливо визначити, що необхідно на вході до системи та що має бути на виході з неї, і не важливо, що знаходиться всередині системи. Таким чином, наведену послідовність називають моделлю «Чорного ящика».

Поняття «Чорний ящик» було запропоновано У.Р. Ешбі. У кібернетиці воно дозволяє вивчати поведінку системи, тобто її реакції на різні подразники, абстрагуючись від їх внутрішньої структури. Таким чином, система вивчається не як набір взаємопов'язаних елементів, а як щось ціле взаємодіюче на своїх входах і виходах. Метод «Чорного ящика» можна застосовувати у різних ситуаціях.

Цей метод слід використовувати, якщо внутрішні процеси системи не доступні. Метод «Чорного ящика» використовується при дослідженні систем, всі елементи і зв'язки яких в принципі доступні, але вони або численні і складні, що призводить до величезних витрат часу і грошей при безпосередньому вивченні; або таке вивчення експертизи є не дійсним з будь-якої причини.

Дослідження за методом «чорного ящика» полягає у попередньому спостереженні взаємодії системи з зовнішнім середовищем та створенні списку впливу на вході та виході, серед яких виявляють суттєві впливи. Потім здійснюється вибір входів і виходів для вивчення, з урахуванням існуючих засобів впливу на систему та засобів спостереження за її поведінкою.

На наступному етапі виконується вплив на входи системи та реєстрації виходів. У процесі вивчення спостерігач і «Чорний ящик» формують систему зі зворотним зв'язком, а перщі результати дослідження – множина пар станів входу та виходу, аналіз яких дає їм можливість встановити причинно-наслідковий зв'язок.

Сьогодні є два види «чорних ящиків». До першого виду відносять будь-який «чорний квадрат», який можна розглядати як автомат, що називається скінченним або нескінченним. Поведінка цих «чорних ящиків» відома.

До другого виду належать такі «чорні ящики», поведінка яких може спостерігатися лише в експерименті. У цьому випадку явною або неявною формою виражається гіпотеза про передбачуваність поведінки «Чорного ящика» в імовірнісному сенсі. Без попередньої гіпотези не можливе будь-яке узагальнення або, як кажуть, неможливо зробити індуктивний висновок, що базується на експериментах з «чорним квадратом».

Таким чином, «Чорний ящик» – це система, в якій вхідні та вихідні дані, внутрішнє улаштування роботи та процесів, що відбуваються в ній, не відомі. Можна вивчати систему тільки на її входах і виходах, однак, це дослідження не забезпечує повне уявлення про внутрішню роботу системи, тому що таку саму поведінку можуть мати різні системи.

Потрібно зазначити, що основною причиною безлічі входів і виходів моделі «Чорний ящик» є те, що будь-яка реальна система, як і будь-який об'єкт, взаємодіє з об'єктами зовнішнього середовища необмежену кількість разів і по-різному.

За ступенем інформованості дослідника про об'єкт є розподіл об'єктів на три типи «ящиків»:

- **«білий ящик»** – про об'єкт знає все;
- **«сірий ящик»** – відома структура об'єкта та не відомі кількісні значення параметрів;
- **«чорний ящик»** – про об'єкт нічого не відомо.

Чорний ящик умовно зображують як на рис. 4.1.

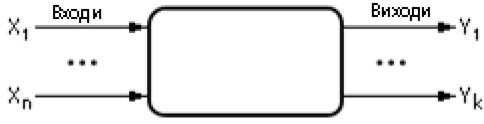


Рис. 4.1. Позначення чорного ящика на схемах

## 4.2. Лінійні регресійні моделі

Для того щоб дослідження були доцільними, часто є зручним евести спробний об'єкт у формі ящика, який має входи і виходи, без урахування деталей його внутрішню структури. Звичайно, перетворення в ящику (на об'єкті) відбуваються (сигнали проходять через зв'язки та елементи, змінюють свою форму і т. д.), але з цього погляду вони відбуваються приховано від спостерігача.

Значення на входах і виходах «чорного ящика» можна спостерігати і виміряти. Вміст ящика не відомий.

Завдання полягає в тому, щоб, знаючи багато значень на входах і виходах, побудувати модель, тобто визначити функцію ящика, згідно з якою вхід перетворюється на вихід. Це завдання має назву **задача регресійного аналізу**.

Залежно від входів, які доступні досліднику для керування або тільки моніторингу, можемо говорити про активний або пасивний експеримент з ящиком.

Наприклад, перед нами є завдання, яке полягає в тому, щоб визначити, як виготовлення продукції залежить від обсягу споживаної електроенергії. Результати спостережень відображені в графі (рис. 4.2.). Всього в графі $n$ експериментальних значень, які відповідають за $n$ спостережень.

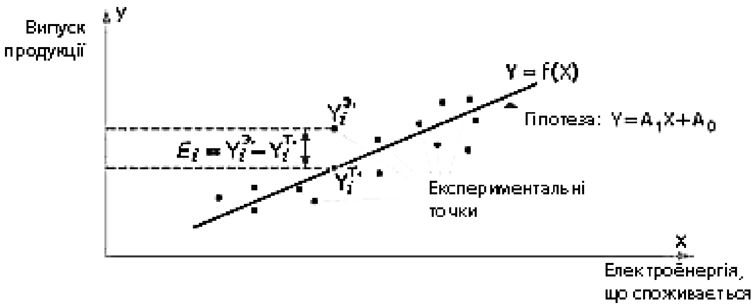


Рис. 4.2. Графічне зображення результатів моніторингу

Для початку, давайте припустимо, що ми маємо справу з чорним ящиком з одним входом і виходом. Уявимо собі для спрощення, що відношення між входом і виходом є лінійні або майже лінійні. Тоді ця модель буде називатися **лінійною одновимірною регресійною моделлю**.

**1) Дослідник робить гіпотезу про структуру поштової скриньки**

Враховуючи експериментально отримані дані, припустимо, що вони підпорядковуються лінійній гіпотезі, тобто вихід $Y$ залежить від входу $X$ лінійно, тобто гіпотеза має вигляд: $Y = A_1X + A_0$ (рис. 4.2).

**2) Визначення невідомих коефіцієнтів $A_0$ та $A_1$ моделі**

Для кожної з $n$ зафіксованих точок, які здобуті за допомогою експерименту розрахувати помилку ($E_i$) між експериментальним значенням ($Y_i^{\text{Експ}}$) та теоретичним значенням ($Y_i^{\text{Теор}}$), що лежить на гіпотетичній прямій $A_1X + A_0$ (див. рис. 4.2):

$$E_i = \left(Y_i^{\text{Експ}} - Y_i^{\text{Теор}}\right), i = 1, \dots, n;\quad E_i = Y_i - A_0 - A_1 \cdot X_i, i = 1, \dots, n.$$

Похибки $E_i$ для всіх $n$ точок потрібно скласти. Кожну з похибок підносять до квадрата та складають їх значення у сумарну похибку вже до єдиного знака, щоб додатні похибки не компенсували у сумі від'ємні:

$$E_i^2 = (Y_i - A_0 - A_1 \cdot X_i)^2, i = 1, \dots, n;\ F(A_0, A_1) = \sum_i^n E_i^2.$$

Мета методу – мінімізація сумарної похибки $F$ за рахунок підбору коефіцієнтів $A_0$, $A_1$. Іншими словами, це означає, що вам потрібно знайти коефіцієнти $A_0$, $A_1$ лінійної функції $Y = A_1X + A_0$, щоб її графік підходив якомога точніше одночасно до усіх точок. Таким чином, це називається **Методом найменших квадратів**.

$$F(A_0, A_1) = \sum_i^n E_i^2 = \sum_i^n (Y_i - A_0 - A_1 \cdot X_i)^2 \Rightarrow \min_{A_0, A_1}.$$

Сумарна похибка $F$ – функція двох змінних$A_0$ і $A_1$, тобто $F(A_0, A_1)$, змінюючи які, можна впливати на загальне значення сумарної похибки (рис. 4.3).

Щоб звести до мінімуму сумарну похибку, знайдемо частинні похідні функцій $F$ для кожної змінної та прирівняємо до нуля (умова екстремуму):

$$\frac{\partial F}{\partial A_0} = -2\sum_{i=1}^n (Y_i - A_0 - A_1X_i) = 0;\quad \frac{\partial F}{\partial A_1} = -2\sum_{i=1}^n (Y_i - A_0 - A_1X_i)X_i = 0.$$

Після розкриття дужок отримуємо систему двох лінійних рівнянь:

$$\sum_{i=1}^n A_0 + A_1\sum_{i=1}^n X_i = \sum_{i=1}^n Y_i;\ A_0\sum_{i=1}^n X_i + A_1\sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_iY_i.$$

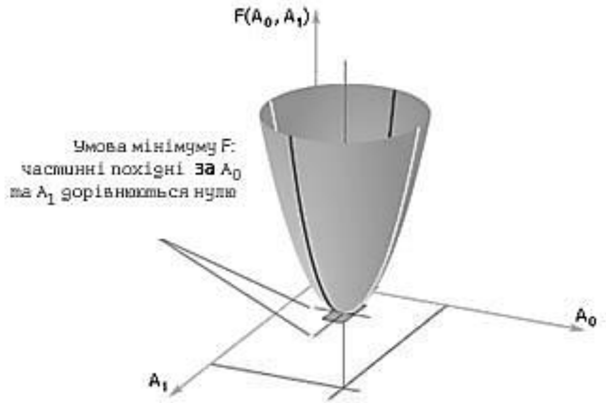


Рис. 4.3. Приблизний вигляд функції похибки

Щоб знайти коефіцієнти $A_0$ і $A_1$ методом Крамера, зобразимо систему у вигляді матриці:

$$\begin{Vmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{Vmatrix} \cdot \begin{Vmatrix} A_0 \\ A_1 \end{Vmatrix} = \begin{Vmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{Vmatrix}.$$

Розв’язок має вигляд:

$$A_0 = \left( \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} X_i Y_i \sum_{i=1}^{n} X_i \right) \Bigg/ \left( \sum_{i=1}^{n} X_i^2 - \left( \sum_{i=1}^{n} X_i \right)^2 \right);$$

$$A_1 = (\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i)/(\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2).$$

Обчислюємо значення $A_0$ і $A_1$.

### 3) Перевірка

Щоб визначити, чи приймається гіпотеза, чи ні, потрібно, по-перше, розрахувати похибку між точками залежності, що визначена експериментально та тією, що визначена теоретично, та сумарну похибку:

$$E_i = \left( Y_i^{\text{Експ}} - Y_i^{\text{Теор}} \right), i = 1, \dots, n; \quad F(A_0, A_1) = \sum_i^n E_i^2.$$

І, по-друге, необхідно знайти значення $\sigma$ за формулою $\sigma = \sqrt{F/n}$ , де $F$ – сумарна похибка, $n$ – це загальна кількість експериментальних значень.

Якщо в смугу, обмежену лініями $Y^{\text{Теор}} - S$ та $Y^{\text{Теор}} + S$ (рис. 4.4), потрапляє 68,26 % і більше експериментальних точок $Y_i^{\text{Експ}}$, то висунута нами гіпотеза приймається. А якщо ні, то вибирають більш складну гіпотезу, або перевіряють вихідні дані. Якщо потрібна більша впевненість у результаті, то використовують додаткову умову: у смугу, обмежену лініями $Y^{\text{Теор}} - S$ та $Y^{\text{Теор}} + S$, має потрапити 95,44 % і більше експериментальних точок $Y_i^{\text{Експ}}$.

Відстань $S$ пов'язана з $\sigma$ таким співвідношенням:

$$S = \sigma/\sin(\beta) = \sigma/\sin(90 - \text{arctg}(A_1)) = \sigma/\text{arctg}(A_1).$$

Умову прийняття гіпотези виведено з нормального закону розподілу випадкових помилок (рис. 4.5). $P$ – ймовірність розподілу нормальної похибки.

### 4.3. Лінійна множинна модель

Припустимо, що функціональна структура ящика знову має лінійну залежність, але кількість вхідних сигналів, що діють одночасно на об'єкт, дорівнює $m$:

$$Y = A_0 + A_1X_1 + \cdots + A_mX_m.$$

Оскільки що мається на увазі, що ми маємо експериментальні дані про всі входи й виходах чорного ящика, то можна обчислити помилку між експериментальним ($Y_i^{\text{Експ}}$) і теоретичним ($Y_i^{\text{Теор}}$) значеннями $Y$ для кожної $i$-той крапки (нехай, як і кількісь, експериментальних точок дорівнює $n$):

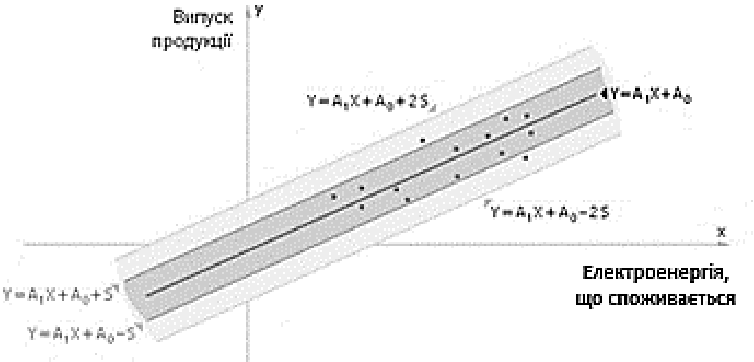


Рис. 4.4. Дослідження допустимості прийняття гіпотези

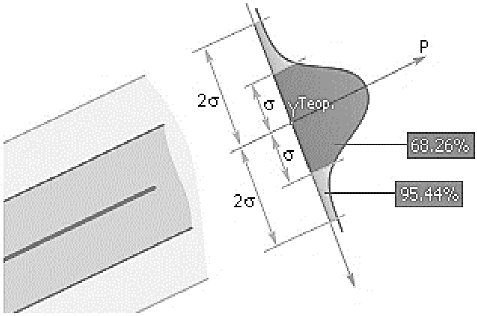


Рис. 4.5. Ілюстрація закону з нормальним розподілом похибок

$$E_i = \left(Y_i^{eksp} - Y_i^{teor}\right), i = 1 \cdots, n;$$

$$E_i = Y_i - A_0 - A_1 X_{1i} - \cdots - A_m X_{mi}, i = 1 \cdots, n.$$

Мінімізуємо сумарну похибку $F$:

$$F(A_0, A_1, \cdots, A_m) = \sum_{i=1}^{n} E_i^2 \Rightarrow \min_{A_0, A_1, \cdots, A_m}.$$

Похибка $F$ залежить від вибору параметрів $A_0, A_1, \cdots, A_m$. Для знаходження екстремуму прирівняємо всі частинні похідні $F$ за невідомими $A_0, A_1, \cdots, A_m$ до нуля:

$$\frac{\partial F}{\partial A_j} = 0, j = \overline{0, m}.$$

Одержимо систему з $m + 1$ рівняння з $m + 1$ невідомими, яку слід розв'язати, щоб визначити коефіцієнти лінійної множинної моделі $A_0, A_1, \cdots, A_m$. Для знаходження коефіцієнтів методом Крамера зобразимо систему в матричному вигляді:

$$\begin{Vmatrix} n & \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{2i} & \cdots & \sum_{i=1}^{n} X_{mi} \\ \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{1i} X_{1i} & \sum_{i=1}^{n} X_{2i} X_{1i} & \cdots & \sum_{i=1}^{n} X_{mi} X_{1i} \\ \sum_{i=1}^{n} X_{2i} & \sum_{i=1}^{n} X_{1i} X_{2i} & \sum_{i=1}^{n} X_{2i} X_{2i} & \cdots & \sum_{i=1}^{n} X_{mi} X_{2i} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^{n} X_{mi} & \sum_{i=1}^{n} X_{1i} X_{mi} & \sum_{i=1}^{n} X_{2i} X_{mi} & \cdots & \sum_{i=1}^{n} X_{mi} X_{mi} \end{Vmatrix} \cdot \begin{Vmatrix} A_0 \\ A_1 \\ A_2 \\ \cdots \\ A_m \end{Vmatrix} =$$

$$= \left\| \begin{matrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} Y_i X_{1i} \\ \sum_{i=1}^{n} Y_i X_{2i} \\ \cdots \\ \sum_{i=1}^{n} Y_i X_{mi} \end{matrix} \right\|.$$

Обчислюємо коефіцієнти $A_0, A_1, \cdots, A_m$.

Далі, за аналогією з одновимірною моделлю, для кожної точки обчислюється помилка $E_i$; потім знаходять сумарну помилку $F$ і значення $\sigma$ і $S$ з метою визначити, чи ухвалюється висунута гіпотеза про лінійність багатовимірного чорного ящика, чи ні.

За допомогою підстановок і перепозначень до лінійної множинної моделі приводяться багато нелінійних моделей.

## 4.4. Нелінійні регресійні моделі

**Поліноміальна множинна регресійна модель.** Якщо чорний ящик має, наприклад, два входи, а залежність виходу від входів нагадує квадратичну, то доцільно вибрати таку гіпотезу:

$$Y = A_0 + A_1 \cdot X_1 + A_2 \cdot X_2 + A_3 \cdot X_1 \cdot X_2 + A_4 \cdot X_1 \cdot X_1 + A_5 \cdot X_2 \cdot X_2.$$

Позначимо: $Z_1 = X_1 \cdot X_2; Z_2 = X_1 \cdot X_1; Z_3 = X_2 \cdot X_2$; і підставимо ці вирази в попередню формулу:

$$Y = A_0 + A_1 \cdot X_1 + A_2 \cdot X_2 + A_3 \cdot Z_1 + A_4 \cdot Z_2 + A_5 \cdot Z_3.$$

Таким чином, це завдання зведене до лінійної множинної моделі.

### *4.4.1. Мультиплікативна регресійна модель*

$$Y = A_0 \cdot X_1 \cdot A_1 \cdot X_2 \cdot A_2 \cdot \ldots \cdot X_m \cdot A_m.$$

Прологарифмуємо ліву й праву частини цього рівняння:

$$\ln(Y) = \ln(A_0) + A_1 \cdot \ln(X_1) + A_2 \cdot \ln(X_2) + \cdots + \ln(X_m).$$

Позначимо:

$$W = \ln(Y), B_0 = \ln(A_0), Z_1 = \ln(X_1), Z_2 = \ln(X_2), \ldots, Z_m = \ln(X_m).$$

Одержимо:

$$W = B_0 + A_1 \cdot Z_1 + A_2 \cdot Z_2 + \cdots + A_m \cdot Z_m.$$

Тобто знову здійснено перехід до *лінійної множинної моделі*.

### *4.4.2. Обернена регресійна модель*

$$Y = k/(A_0 + A_1 \cdot X_1 + \cdots + A_m \cdot X_m).$$

Замінимо: $W = 1/Y\,, a_i = A_i/k$. І перейдемо до лінійної множинної моделі:

$$W = 1/Y\,, a_i = A_i/k.$$

### *4.4.3. Експоненційна модель*

$$Y = \mathrm{e} \cdot B_0 + B_1 \cdot X_1 + B \cdot X_2 + \cdots + B_m \cdot X_m.$$

Прологарифмуємо ліву й праву частини рівняння:

$$\ln(Y) = B_0 + B_1 \cdot X_1 + B \cdot X_2 + \cdots + B_m \cdot X_m.$$

Виконаємо заміну $W = \ln(Y)$ і одержимо:

$$W = B_0 + B_1 \cdot X_1 + B \cdot X_2 + \cdots + B_m \cdot X_m.$$

Далі користуємося виразом для лінійної множинної моделі.

## 4.5. Динамічні системи

Вище ми розглядали статичні моделі, тобто випадок, коли один експеримент не залежить від іншого. Можна сказати, що система не мала пам'ять. Тобто, в який би момент часу ми не вимірювали значення вихідної величини, при однаковому значенні вхідного сигналу результат був той самий. Якщо щоразу значення на виході, при тому самому вхідному значенні, різне, тобто залежить від того, в якій послідовності подавалися вхідні значення, то ми маємо справу з динамічною системою.

Динамічні системи, на відміну від статичних, пам'ятають свій минулий стан, тобто мають пам'ять. Тому в записі моделі динамічних систем присутня похідна, що зв'язує минулий стан системи із сьогоденням. Чим більшу пам'ять має система, тим більше станів з минулого впливають на сьогодення, тим більший ступінь старшої похідної використовується в записі моделі. Розглянемо моделювання динамічних систем.

Постановка завдання: на вході й виході чорного ящика є залежності параметрів $X$ і $Y$ від часу $t$. Завдання полягає в тому, щоб адекватно визначити чорний ящик.

Графіки залежностей $X(t)$ і $Y(t)$ можуть бути самими різними, наприклад, такими, як показано на рис. 4.6.

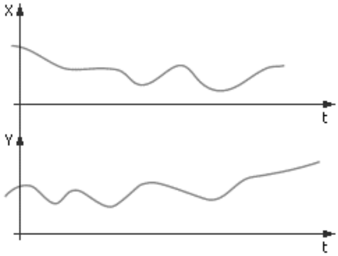


Рис. 4.6. Часова залежність – вхідний і вихідний сигнали

Оскільки моделювання систем передбачає числові розрахунки на комп'ютері, аналоговий сигнал переводять у дискретний вигляд. Для цього з певною частотою вихідний сигнал дискретизують, як показано на рис. 4.7.

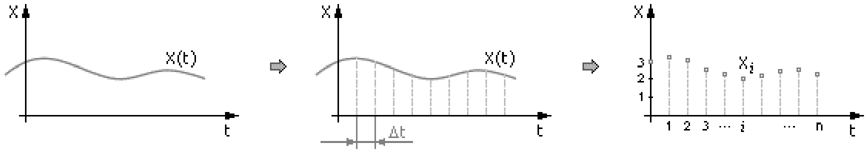


Рис. 4.7. Дискретизований часовий сигнал

За цими даними будують таблицю відкликів (табл. 4.1, де $\Delta t = 0{,}1$).

Таблиця 4.1 – Таблична вистава часового сигналу

| $I$ | 0 | 1 | 2 | 3 | … | $i$ | … | $n$ |
|---|---|---|---|---|---|---|---|---|
| $T$ | 0 | 0,1 | 0,2 | 0,3 | … | $\Delta t \cdot i$ | … | $\Delta t \cdot n$ |
| $x_i$ | 3 | 3,2 | 3,1 | 2,6 | … | $x_i$ | … | $x_n$ |

Сукупність значень змінної в таблиці, упорядкованих у часі, часто називають динамічним рядом. Природно, частина інформації при такій операції втрачається. Чим менша відстань між відділеннями, чим більша частота дискретизації, тим менші втрати інформації. Частоту дискретизації ухвалюють такою, щоб не втратити високочастотні складові в сигналі, окремі піки.

Будь-яка динамічна система характеризується рядом параметрів. Зазвичай (найчастіше) параметрами називають коефіцієнти при похідних (першої, другої і т.д.) у записі моделі. Чим більший ступінь старшої похідної, яка присутня в записі моделі, тим більший порядок динамічної системи, тим глибша її пам'ять, і тим більше коефіцієнтів (параметрів) необхідно визначити, щоб ідентифікувати систему.

Для того щоб визначити параметри динамічної системи, спочатку потрібно оцінити її порядок: він збігається зі ступенем найбільшої з похідних $Y$ стосовно $t$. Допустимо, що на вхід системи, який перебував до цього у нульових початкових умовах, подали одиничний сигнал $Y(t)$, як показано на рис. 4.8.

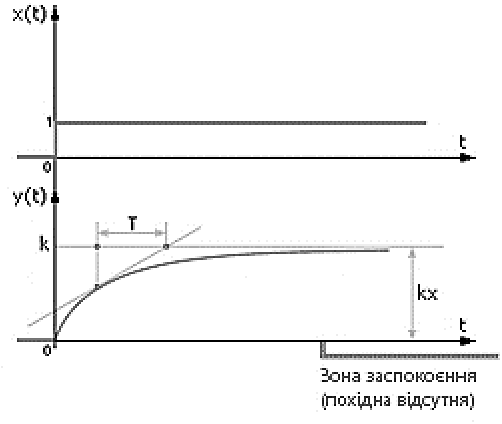


Рис. 4.8. Вхідний і вихідний сигнали, типові для системи першого порядку

Пояснимо зміст графіка. При нульових початкових умовах, якщо вхідний сигнал відсутній, вихідний сигнал дорівнює нулю, і говорять, що система перебуває в спокої. Якщо подати на вхід одиничний (пробний) сигнал і втримувати його на вході досить довго, то система на виході спробує підпорядкуватися йому, почне відхилятися від нульового стану. Очікується, що система на виході повинна дійти до значення $kx$, тобто збільшувати сигнал $x$ в $k$ разів ($k$ – коефіцієнт підсилення вхідного сигналу). Але, як видно, відбувається це не відразу, а з деякою затримкою, сигнал на виході наростає поступово, інерційно. Наскільки інерційно реагує система, залежить від параметра $T$. Система досягненя значення $kx$ на виході й буде тримати цей сигнал, поки тримається на вході одиничний сигнал. Перехід від нуля до $kx$ відбувається в часі. Перехід – процес динамічний, тобто в сигналі присутня зміна, що описується похідною, і вихід виявляється меншим від входу на деяку величину $f$:

$$y = kx - f(dy/dt).$$

Коли система досягне на виході значення, рівного $kx$, то змін не буде, значення похідної стане рівним нулю $y = kx$.

Якщо на виході буде спостерігатися експонентний сигнал, то система буде називатися системою першого порядку (або ланкою першого порядку). Для її опису досить однієї похідної (а в розв'язку моделі буде присутній один інтеграл):

$$T(dy/dt) + y = kx.$$

У такої системи два параметри – $T$ і $k$.

Помітимо, що один інтеграл у лінійних динамічних систем завжди «породжує» одну експоненту, подвійний інтеграл – суму двох експонент і так далі.

Щоб визначити, чи є крива експонентною, у кожній її точці проводиться дотична до перетинання з лінією рівня, що встановився (на рис. 4.8 це лінія $y(t) = k$); у випадку, якщо крива є експонентом, величина $T$ у будь-якій точці буде сталою.

Визначити $T$, використовуючи графік, можна ще так. Проведіть лінію, паралельну осі $t$ на рівні 0,95 $k$. Із точки, де ця лінія перетне експоненту, вилучите перпендикуляр на вісь $t$. Відрізок від 0 до точки перетину перпендикуляра з віссю $t$ буде рівний $3T$.

$T$ характеризує інерційність системи (пам'ять). При малій величині $T$ система слабко залежить від передісторії й вхід миттєво змушує змінитися на вихід. При великому значенні $T$ система повільно реагує на вхідний сигнал, а при дуже великому значенні $T$ система видає незмінний вихідний сигнал, практично не реагуючи на вхідні впливи.

Коефіцієнт $k$ характеризує здатність системи до посилення (при $k < 1$ – до ослаблення) рівня вхідного сигналу. Щоб визначити коефіцієнт $k$ на графіку, досить дочекатися заспокоєння сигналу на виході системи й обчислити відношення рівня вихідного сигналу до рівня вхідного. Математично це означає, що всі, що складаються, містять похідні, дорівнюють нулю (система заспокоїлася, руху немає), а доданок, що залишився, $Y = kX$ визначає значення $k$.

**Ланка першого порядку.** Ланка першого порядку має два параметри: інерційність $T$ і коефіцієнт підсилення $k = Y(t = \infty)/X$.

Чим більше похідних ураховується в записі моделі, тим з ланкою більшого порядку ми маємо справу, тим більше коефіцієнтів при похідних слід визначити.

Уведемо поняття передавальної функції як моделі динамічної системи. За визначенням передавальна функція – це відношення виходу до входу:

$$W = Y/X.$$

Передавальна функція ланки першого порядку має вигляд:

$$W = k/\left(T_p + 1\right),$$

де « $p$» – символ диференціювання, тотожно рівний «$d/dt$» Тоді, використовуючи визначення передавальної функції, маємо:

$$Y/X = k/\left(T_p + 1\right).$$

Далі одержимо:

$$\left(T_p + 1\right)Y = k \cdot X$$

або

$$mT(dY/dt) + Y = k \cdot X,$$

або

$$T(\Delta Y/\Delta t) + Y = k \cdot X.$$

У різницевому вигляді рівняння можна записати як

$$T(Y_{i+1} - Y_i) + Y_i \cdot \Delta t = k \cdot X_i \cdot \Delta t.$$

Або, виразивши сьогодення через минуле,

$$Y_{i+1} = A \cdot X_i + B \cdot Y_i,$$

де $A = k \cdot \Delta t/T$ і $B = 1 - \Delta t/T$ – вагові коефіцієнти. $A$ вказує на вагу компонента $X$, що визначає вплив зовнішнього світу на систему, $B$ указує на вагу компонента $Y$, що визначає пам’ять системи, вплив на її поведінку історії.

Зокрема , якщо $B = 0$, то $Y_{i+1} = A \cdot X_i$, і ми маємо справу з безінерційною системою $Y = k \cdot X$, що миттєво реагує на вхідний сигнал, і збільшує його в $k$ разів.

Якщо коефіцієнт $B = 0{,}5$, тобто $1 - \Delta t/T = 0{,}5$ або $\Delta t/T = 0{,}5$ , то одержуємо, що коефіцієнт $A = k \cdot \Delta t/T = 0{,}5$ і, отже, $Y_{i+1} = 0{,}5 \cdot X_i + 0{,}5 \cdot\ Y_i$. При постійному (одиничному) вхідному сигналі $X$ буде отриманий графік, як на рис. 4.9.

Експонента, зображена на графіку, при великому $n$ (у межі $n = \infty$) прямує до значення вхідного (одиничного) сигналу $X$, помноженого на коефіцієнт підсилення $k$, що підтверджується розрахунками:

$$Y_{n+1} = 0{,}5 \cdot k \cdot X_n + 0{,}5 \cdot\ Y_n = 0{,}5 \cdot k \cdot X_n + 0{,}5 \cdot (0{,}5 \cdot k \cdot X_{n-1} + 0{,}5 \cdot\ Y_{n-1}) = \\ \cdots == (0{,}5^1 + 0{,}5^2 + \cdots + 0{,}5^{n+1}) \cdot k \cdot X_0 + 0.5^{n+1} \cdot\ Y_0 = 1 \cdot k \cdot X_0.$$

Нагадаємо, що вираз $(0{,}5^1 + 0{,}5^2 + \cdots + 0{,}5^{n+1})$ є геометричною прогресією, сума якої при $n = \infty$ рівна 1. А вираз, що стоїть при $Y_0$, $0{,}5^{n+1}$ обертається в 0 при $n = \infty$.

Якщо ще підсилити вплив минулого ($B = 1$), то система почне інтегрувати саму себе (вихід поданий на вхід системи), додаючи увесь час вхідний сигнал, що відповідає експонентному необмеженому зростанню вихідного сигналу: $Y_{i+1} = A \cdot X_i + Y_i$. За змістом це відповідає додатному зворотному зв'язку. При $B = -1$ маємо модель: $Y_{i+1} = A \cdot X_i - Y_i$ , за змістом відповідну від'ємному зворотному зв'язку. При визначенні моделі потрібно знайти невідомі коефіцієнти $k$ і $T$.

**Ланка другого порядку (коливальна ланка).** Такі ланки описуються диференціальним рівнянням виду:

$$a_0 \frac{d^2y}{dt^2} + a_1 \frac{dy}{dt} + a_2 y = bU.$$

Якщо на вхід ланки подати одиничну функцію Хевісайда від часу $l[t]$, при нульових початкових умовах системи, то реакція на виході буде називатися **перехідною функцією (або перехідною характеристикою)**, яку часто позначають як $h(t)$. Сигнал $l[t]$ – це, у деякому розумінні, еталонний іспитовий сигнал. Існують і інші еталонні іспитові сигнали. Наприклад, нескінченний імпульс нульової довжини (дельта-функція Дирака), гармонічний сигнал, періодичні прямокутні імпульси.

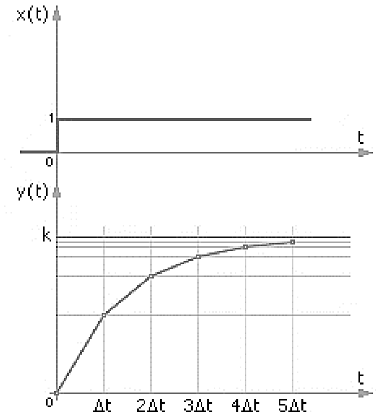


Рис. 4.9. Реакція ланки першого порядку

Перетворимо за Лапласом це рівняння:

$$a_0 \cdot p^2 \cdot Y(p) + a_1 \cdot p \cdot Y(p) + a_1 \cdot Y(p) = b \cdot U(p),$$

або, інакше:

$$(a_0 \cdot p^2 + a_1 \cdot p + a_1)Y(p) = b \cdot U(p).$$

Визначимо передавальну функцію ланки:

$$W(p) = \frac{Y(p)}{U(p)} = \frac{b}{a_0 p^2 + a_1 p + a_2} = \frac{b/a_2}{\frac{a_0}{a_2}p^2 + \frac{a_1}{a_2}p + 1} = \frac{k}{T^2 p^2 + 2\xi T p + 1}$$

Якщо записати рівняння без вхідного впливу (нульові вхідні впливи $U = 0$) і скоротити $Y$, тобто : $\mathrm{T}^2 p^2 + 2\xi \mathrm{tp} + 1 = 0$, то таке рівняння буде називатися характеристичним, оскільки характеризує винятково внутрішні властивості ланки. Зверніть увагу, що в записі ланки дотримуються три параметри:

$$T^2 = \frac{a_0}{a_2}, \qquad 2\xi T = \frac{a_1}{a_2}, \qquad k = \frac{b}{a_2},$$

де $T$ – стала часу (у секундах); $\xi$ – коефіцієнт загасання (безрозмірна величина); $k$ – передавальний коефіцієнт.

Залежно від величини $\xi$ ланки другого порядку класифікуються за видами:

$\xi = 0$ – консервативна ланка другого порядку;

$0 < \xi < 1$ – коливальна ланка другого порядку;

$\xi > 1$ – аперіодична ланка другого порядку.

**Аперіодична ланка 2-го порядку $\xi \geq 1$.** Характеристичне рівняння ланки таке:

$$T^2p^2 + 2\xi Tp + 1 = 0.$$

І воно має дійсний від'ємний корінь:

$$p_{1,2} = \frac{-\xi \pm \sqrt{\xi^2 - 1}}{T}.$$

Цю ланку можна зобразити у вигляді послідовно з'єднаних ланок з різними сталими часу:

$$W(p) = \frac{k}{(T_1p + 1)(T_2p + 1)}.$$

Тоді при $T_1 > T_2$ перехідна характеристика ланки має вигляд:

$$h(t) = k\left(1 - \frac{T_1}{T_1 - T_2}e^{-\frac{t}{T_2}}\right).$$

Тобто в розв'язку присутні загасаючі експоненти. Типову поведінку ланки з такими параметрами показано на рис. 4.10.

## 4.6. Модель у вигляді фільтра Калмана

Калманом було доведено теорему про те, що будь-який динамічний сигнал може бути поданий у вигляді рівняння:

$$Y_i = A_1X_i + A_2X_{i-1} + \cdots + B_1Y_{i-1} + B_2Y_{i-2} + \cdots + C.$$

Суть фільтра Калмана полягає в тому, що вихід системи в $i$-й момент часу визначається вхідним сигналом, його передісторією й передісторією самого стану системи.

Чим більше є членів ряду, тобто чим більше змінних $Y$ ураховується в записі моделі, тим глибша пам'ять системи. Зазначимо, що наявність члена $Y_{i-1}$ у моделі динамічної системи відповідає наявності першої похідної, $Y_{i-2}$ – другої похідної і т.д.

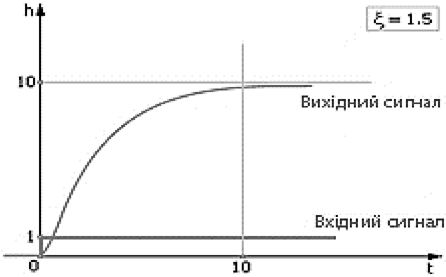


Рис. 4.10. Реакція аперіодичної ланки на одиничний вхідний сигнал

Допустимо, відомі такі експериментальні дані: стану сигналів $X_i$ і $Y_i$ в $n$ часових точках (табл. 4.2).

Оскільки для кожної експериментальної точки $X_i$ потрібно вказати її сусідів, що задаються поруч, то зручно відліки подати в розширеній таблиці, яка використовується для розрахунків (табл. 4.3).

Таблиця 4.2 – Таблиця експериментальних даних

| № | $X_i$ | $Y_i$ |
|---|---|---|
| 1 | $X_1$ | $Y_1$ |
| 2 | $X_2$ | $Y_2$ |
| … | … | … |
| $n-1$ | $X_{n-1}$ | $Y_{n-1}$ |
| $n$ | $X_n$ | $Y_n$ |

Знаходимо помилку між значенням експериментально знятої точки та її теоретичним значенням (гіпотезою):

$$E_m = Y_m - A_1 X_m - A_2 X_{m-1} - \cdots - B_1 Y_{m-1} - B_2 Y_{m-2} - \cdots - C.$$

Сумарна помилка $F$ (сума береться за всіма експериментальними точками) повинна бути мінімізована щодо обумовлених змінних $A_1, A_2, \dots, B_1\ B_2, \dots, C$:

$$F = \sum_{i=1}^{m} E_i^2 \Rightarrow \min_{A_1, A_2, \dots, B_1\ B_2, \dots, C}.$$

Після взяття частинних похідних від $F$ за $A_1, A_2, \dots, B_1\ B_2, \dots, C$, прирівнювання їх до нуля й складання системи рівнянь випливає *лінійна множинна регресійна модель*, з якої визначаються невідомі коефіцієнти $A_1, A_2, \dots, B_1\ B_2, \dots, C$ моделі.

Таблиця 4.3 – Таблиця експериментальних даних і проміжних розрахунків

| | $X_i$ | $X_{i-1}$ | … | $Y_i$ | $Y_{i-1}$ | $Y_{i-2}$ |
|---|---|---|---|---|---|---|
| $m$ | $X_m$ | $X_{m-1}$ | … | $Y_m$ | $Y_{m-1}$ | $Y_{m-2}$ |
| $m+1$ | $X_{m+1}$ | $X_m$ | … | $Y_{m+1}$ | $Y_m$ | $Y_{m-1}$ |
| $m+2$ | $X_{m+2}$ | $X_{m+1}$ | … | $Y_{m+2}$ | $Y_{m+1}$ | $Y_m$ |
| … | … | … | … | … | … | … |

Оскільки коефіцієнти моделі визначені, побудуємо реалізацію (рис. 4.11), що імітує поведінку системи, описаної фільтром Калмана.

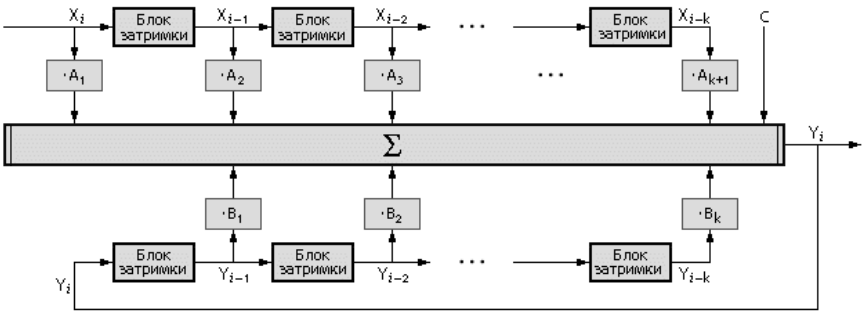


Рис. 4.11. Варіант технічної реалізації фільтра Калмана

«Блок затримки» у наведеній реалізації необхідний для того, щоб зрушити сигнал на такт і одержати сусідній відлік для наступної змінної ряду моделі. Залежно від середовища реалізації «блок затримки» можна організувати різними способами.

На рис. 4.12 наведено схему налагодження (автоматичного знаходження коефіцієнтів).

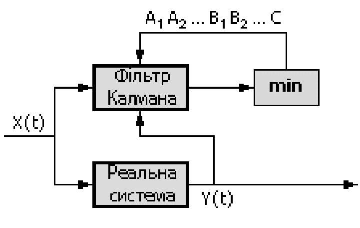


Рис. 4.12. Схема автоматичного налагодження коефіцієнтів моделі «на ходу»

На рис. 4.13 наведено схему перевірки фільтра Калмана.

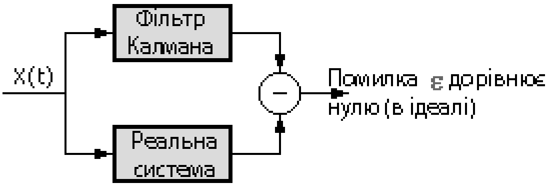


Рис. 4.13. Схема перевірки роботи моделі фільтра Калмана

## 4.7. Модель динамічної системи у вигляді подання Фур'є

Цей спосіб моделювання динамічних систем ґрунтується на тому, що в будь-якому сигналі наявні гармонічні складові. Залежно від частоти складові називаються гармоніками (перша, друга і т. д.). Сума гармонік із відповідними вагами становить модель сигналу.

Нехай, наприклад, у деякому сигналі наявна сума трьох гармонік : $3\cos(t) + 2\cos(3t) + 0{,}5\cos(5t)$. Це значить, що в сигналі наявна перша гармоніка з амплітудою 3, третя гармоніка – з амплітудою 2, п'ята гармоніка – з амплітудою 0,5. Сам сумарний сигнал виглядає так, як показано на рис. 4.14.

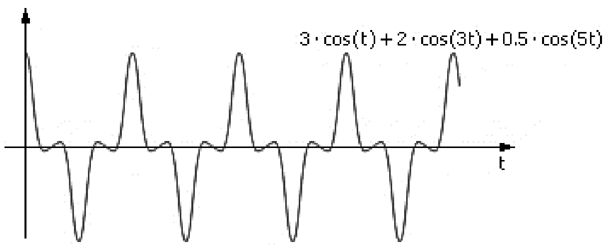


Рис. 4.14. Приклад гармонічного сигналу

Спектр цього сигналу показаний на рис. 4.15. Зрозуміло, що в нашому прикладі більшу вагу (амплітуду) у сигналі має (більше за інших наведена) перша гармоніка, а найменшу вагу – п'ята.

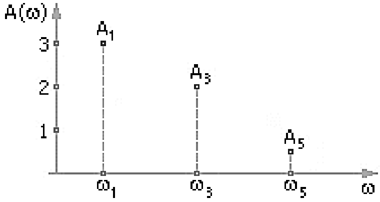


Рис. 4.15. Приклад спектра гармонічного сигналу

Будь-який сигнал, яким би складним він не був, може бути поданий сумою гармонік. Більш простий сигнал подається у вигляді меншої кількості гармонік, більш складний – більшого. Швидко змінний сигнал, що містить різкі піки, має у своєму складі гармоніки високих порядків. Чим більше гармонік подано в моделі сигналу, тим точніше, у загальному випадку, модель відображає реальний сигнал.

Нехай заданий якийсь сигнал $X(t)$ (рис. 4.16).

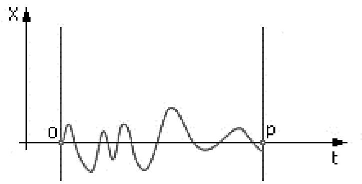


Рис. 4.16. Часовий сигнал на вході перетворення Фур'є

Визначимося з часом розгляду сигналу: якщо сигнал *періодичний*, то час розгляду дорівнює періоду $p$ сигналу; якщо сигнал *неперіодичний*, то періодом сигналу вважається увесь час його розгляду.

$$A_0 = \frac{2}{p}\int_0^p X(t)dt,$$

$$A_i = \frac{2}{p}\int_0^p X(t)\cos\left(\frac{2\pi i t}{p}\right)dt\,,$$

$$B_i = \frac{2}{p}\int_0^p X(t)\sin\left(\frac{2\pi i t}{p}\right)dt.$$

$A_i$ і $B_i$ – це ваги відповідних гармонік, що наявні у сигналі; $i$ – номер гармоніки. Формули їх розрахунку називаються прямим перетворенням Фур'є.

Значення $2\pi \cdot i/p = \omega_i$ – це частота $i$-тої гармоніки. Відзначимо також, що частота $i$-тої гармоніки пов'язана із частотою першої гармоніки простим співвідношенням: $\omega_i = i \cdot \omega_1$.

Зазначимо важливу особливість цього способу подання: замість усього сигналу з усіма його подробицями достатньо зберігати вектор чисел, що зображають вагові коефіцієнти складових його гармонік: $(A_0, A_1, A_2, \dots, B_1, B_2, \dots)$. Тобто ці числа повністю характеризують вихідний сигнал, оскільки за ними сигнал можна повністю відновити за формулою зворотного перетворення Фур'є:

$$x(t) = \frac{A_0}{2} + A_1 \cos\left(\frac{2\pi t}{p}\right) + B_1 \cos\left(\frac{2\pi t}{p}\right) + \dots + A_i \cos\left(\frac{2\pi i t}{p}\right) + B_i \cos\left(\frac{2\pi i t}{p}\right) + \cdots$$

Саме ці числа використовуються також при обробці сигналу в моделі динамічної системи. Зображення цих чисел на графіку, залежно від номера гармоніки (частоти), називається *спектром сигналу* (рис. 4.17). Спектр показує, наскільки є наявною у сигналі відповідна складова. *Спектр* – це частотна характеристика сигналу.

У цьому випадку сигнал поданий у частотній області. Завжди за формулами прямого перетворення Фур'є можна перейти із часової області в частотну, а за формулами зворотного перетворення Фур'є перейти із частотної області в часову. В якій області (частотній або часовій) працювати із сигналом в окремий момент, вирішують із міркувань зручності, наочності й економії обчислень. Помітимо, що ємні, з погляду обчислень, операції інтегрування й диференціювання сигналу в часовій області заміняються на операції алгебраїчного додавання й множення в частотній області, що з обчислювальної точки зору реалізується набагато точніше і швидше.

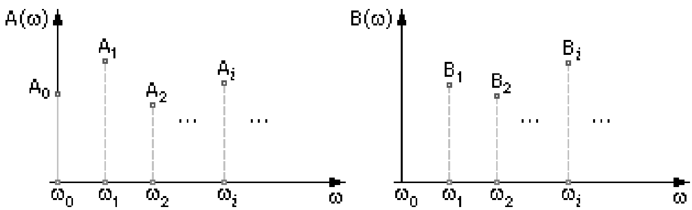


Рис. 4.17. Сигнал, представлений у частотній області на виході перетворення Фур'є – спектр сигналу

Система чисел $A_i$ і $B_i$ є повною характеристикою сигналу. Такою ж повною характеристикою сигналу є система чисел $S$ і φ, які також утворюють спектр (рис. 4.18). У даному випадку $S$ – це амплітудно-частотна характеристика (АЧХ), φ – фазочастотна характеристика (ФЧХ).

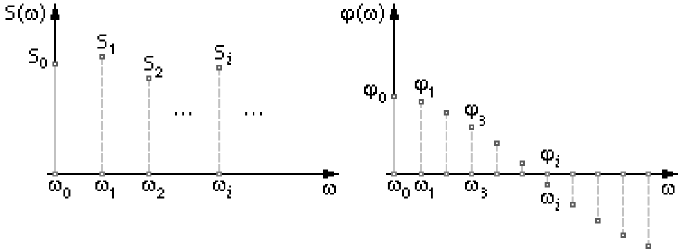


Рис. 4.18. Сигнал, представлений у частотній області, амплітудно-частотна й фазочастотна характеристика сигналу

Системи «$A$ і $B$» та «$S$ і φ» є повністю рівнозначними. Перехід із системи «$A$ і $B$» у систему «$S$ і φ» проводиться за такими формулами: $S_i = \mathrm{sqrt}(A_i^2 + B_i^2)$ – абсолютна амплітуда сигналу; $\varphi_i = \mathrm{arctg}(B_i/A_i)$ – фаза сигналу; при додаванні гармонік потрібно враховувати зсув фаз (зсув фаз проілюстрований на рис. 4.21).

У випадку із системою «$S$ і φ» зворотне перетворення Фур'є має вигляд:

$$x(t) = \sum_{i=1}^{k} S_i \cdot \sin(\omega_i t + \varphi_i).$$

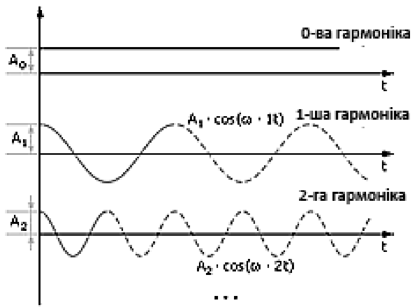


Рис. 4.19. Геометрична ілюстрація параметрів A і ω для косинусової складової гармонічного сигналу

Рис. 4.19 і 4.20 роз'яснюють зміст коефіцієнтів $A$ і $B$ різних гармонік. Ці коефіцієнти – амплітуди синусів та косинусів відповідних частот (гармонік). У часовій області графічно вони відповідають розмаху гармонічних коливань; у частотній – висоті спектральної смужки на відповідній частоті.

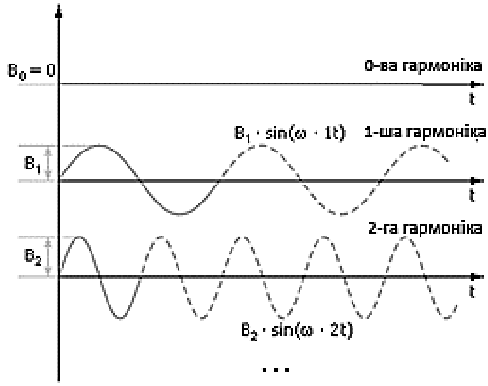


Рис. 4.20. Геометрична ілюстрація параметрів B і ω для синусної складової гармонічного сигналу

Зміст чисел $S_i$ і $\varphi_i$ роз'яснений на рис. 4.21.

Розглянемо модель об'єкта. Нехай є вхідний динамічний сигнал $X(t)$ та об'єкт $F$, який перетворює цей сигнал у вихідний $Y(t)$(рис. 4.22). Якщо об'єкт описується диференціальними рівняннями, то таким перетворенням є інтегрування вхідного сигналу й обчислення $Y(t)$. Інтегрування, як було раніше показано, – операція, що потребує значних обчислювальних ресурсів та має значну похибку при реалізації на цифрових машинах.

Якщо перейти від опису вхідного сигналу в часовій області до опису в частотній області, а від диференціальних рівнянь перейти до частотної характеристики об'єкта, – тобто фактично замінити сигнал на частотну модель сигналу, а об'єкт – на частотну модель об'єкта, – то з обчислювальної точки зору процес перетворення сигналу спроститься. Звичайно, отриманий результат теж буде частотною моделлю вихідного сигналу, яку для одержання остаточної відповіді доведеться сконвертувати в часову область $Y(t)$. Процес такої конвертації із частотної області в часову й обернено називається *перетворенням Фур'є* (є й інші перетворення). Для тих об'єктів, для яких відома їхня модель у частотній області, такий підхід досить просто реалізується на комп'ютері й дозволяє досягти будь-якої наперед заданої точності.

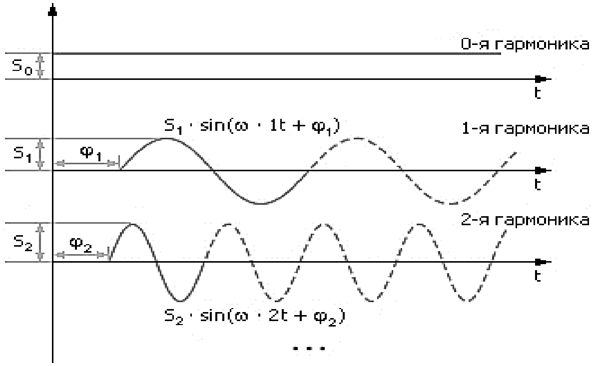


Рис. 4.21. Геометрична ілюстрація параметрів S і φ для складової гармонічного сигналу

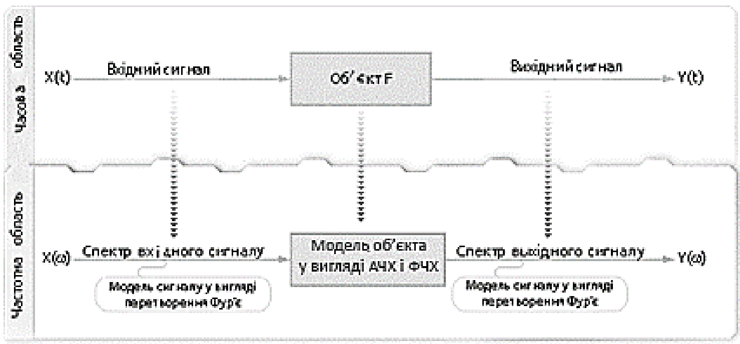


Рис. 4.22. Схема моделювання динамічного об'єкта при переході з часової області подання в частотну

Модель об'єкта в частотному вигляді називається передавальною функцією, або АЧХ (амплітудно-частотною характеристикою). Об'єкти, для яких відомі АЧХ, зазвичай називають типовими ланками (підсилювальна ланка, аперіодична, коливальна і т.д.). Нехай для прикладу характеристика об'єкта в частотній області така (рис. 4.23).

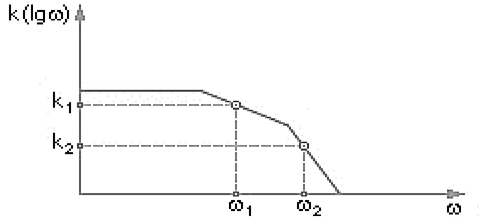


Рис. 4.23. Амплітудно-частотна характеристика

Амплітудно-частотна характеристика (АЧХ) показує, наскільки пропускається об'єктом на вихід відповідна гармоніка. Значення $k_i$ характеризує коефіцієнт підсилення гармонічного сигналу на певній частоті $\omega_i$.

Моделювання проходження сигналу через об'єкт у цьому вигляді полягає в множенні коефіцієнта $A_i$ гармоніки із частотою $\omega_i$ вхідного сигналу $X(t)$ на коефіцієнт підсилення $k_i$ при тій же гармоніці із частотою $\omega_i$ в АЧХ: $A_i^* = A_i(\omega_i)k_i(\omega_i)$. Для коефіцієнта $B$ перетворення аналогічне. У результаті виходить коефіцієнт $A_i^*$ вихідної гармоніки цієї частоти $\omega_i$. Процедура виконується для всіх частот, поданих у вхідному сигналі й АЧХ. Після одержання спектра вихідного сигналу можна відновити сигнал як тимчасову залежність за допомогою *формули зворотного перетворення Фур'є.*

У такий спосіб моделювання проходження сигналу через динамічний об'єкт звелося до операції множення двох змінних, точніше, до операції поелементного множення вектора одних змінних на вектор інших.

Схема перетворення показана на рис. 4.24.

Якби часовий сигнал проходив через ланку, яка в часовій області подана у вигляді диференціального рівняння, то довелося б його інтегрувати, що, звичайно, призводить до похибок результату. У частотній області достатньо перемножити значення коефіцієнтів ряду Фур'є сигналу й ланки при однакових частотах. Очевидно, що перевагою методу є заміна диференціальних рівнянь моделі на алгебраїчні. Зрозуміло, що цей підхід може бути використаний тільки для об'єктів, у яких відомий вид передавальної функції. Для невідомих випадків АЧХ може бути отримана числовим розкладанням у ряд.

У процесі моделювання набору об'єктів для перетворення сигналу (наприклад, протяжних трактів радіоелектронних обладнань) іноді доводиться застосовувати пряме й зворотне перетворення Фур'є неодноразово. На практиці послідовні блоки часто називають каскадами.

Метод, який ми розглянули, є одним із найбільш швидкодіючих. Це пов'язано із заміною операцій інтегрування й диференціювання, що зустрічаються в моделях динамічних ланок, на операції додавання й множення при переході в частотну область. Така процедура забезпечує точність і швидкодію моделі.

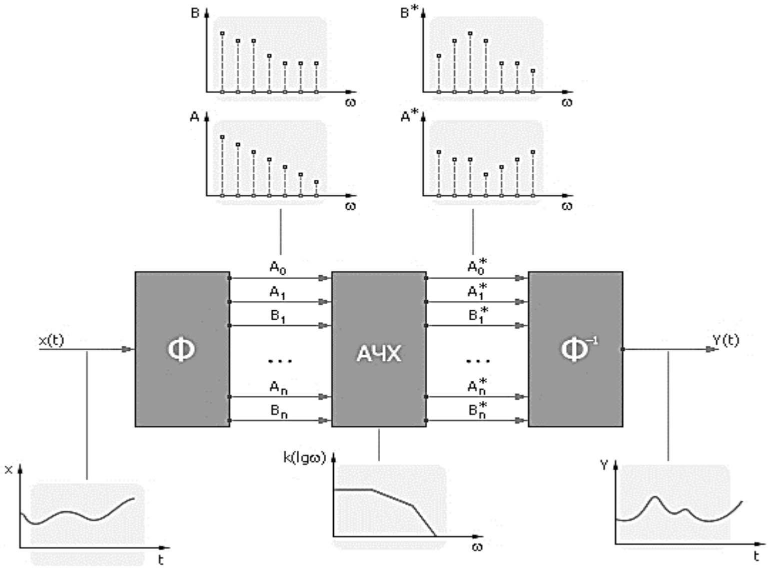


Рис. 4.24. Процедура перетворення сигналу при використанні методу Фур'є

Для методу важливо, з якою частотою дискретизується сигнал при розкладанні в ряд Фур'є. Якщо частота дискретизації мала, тобто відліки в сигналі випливають рідко, з великими інтервалами, то частина сигналу залишається втраченою, оскільки між відліками може виявитися різко зрослий і опалий пік, інформація про який зникне. Тобто говорять, що мала частота дискретизації зрізає високі частоти в сигналі. (Пік – це і є високочастотна складова, яка може бути втрачена).

За теоремою Нейквіста, щоб не втратити відповідну гармоніку, потрібно дискретизувати сигнал із частотою щонайменше у 2 рази більшою, ніж найвища частота із поданих в аналоговому сигналі:

$$2F_{\max} \le F_d,$$

де $F_d = 1/\tau_d$ – частота дискретизації; $F_{\max}$ – максимальна частота, наявна в сигналі.

## Контрольні запитання

1. Дайте визначення «чорному ящику» та назвіть типи «ящиків» за ступенем інформованості дослідника про об’єкт, що розглядається?
2. Яке завдання вирішує задача регресійного аналізу?
3. Дайте загальну характеристику лінійним регресійним моделям.
4. Дайте загальну характеристику нелінійним регресійним моделям.
5. У чому різниця моделювання динамічних систем від статичних систем?
6. Дайте загальну характеристику моделей у вигляді фільтра Калмана.
7. В якому випадку можливо створити модель динамічної системи у вигляді подання Фур'є?

# 5. СТАТИСТИЧНЕ МОДЕЛЮВАННЯ СИСТЕМ

## 5.1. Загальні відомості про статистичне моделювання

***Статистичне моделювання*** (СМ) – це вид комп’ютерного моделювання, що дозволяє одержувати статистичні дані про процеси в модельованій системі *S*. Статистичне моделювання – базовий метод моделювання, який полягає в тому, що модель випробовується множиною випадкових сигналів із заданою щільністю ймовірності. Метою такого моделювання є статистичне визначення вихідних результатів.

*Статистична модель випадкового процесу* – це алгоритм, за допомогою якого імітують роботу складної системи, що зазнала випадкового збурюванням; імітують взаємодію елементів системи, що мають імовірнісний характер. Тоді статистичне моделювання можна визначити як спосіб вивчення складних процесів і систем, що зазнали випадкового збурювання, за допомогою імітаційних моделей.

Методика статистичного моделювання, яку часто називають методом Монте-Карло, складається з таких етапів:

- моделювання на ЕОМ псевдовипадкових послідовностей із заданою кореляцією й законом розподілу ймовірностей (метод Монте-Карло), що імітують на ЕОМ випадкові значення параметрів при кожному випробуванні;
- використання отриманих числових послідовностей в імітаційних математичних моделях;
- статистична обробка результатів моделювання.

Проілюструємо суть методу Монте-Карло відносно простим прикладом. Нехай потрібно оцінити надійність системи (рис. 5.1).

Система виконує свою функцію, якщо працюють ланцюжки блоків: 1, 2, 5,7; 1, 3, 5, 7; 1, 4, 6, 7. Якісь блоки можуть відмовити. Кожний блок характеризується часом безвідмовної роботи $\tau_i, i = \overline{1,7}$. Нехай задані щільності розподілу $p_i(\tau_i), i = \overline{1,7}$. Яка надійність системи загалом?

Розглянемо випадкову величину

$$T = \min\{\tau_1, \max[\min(\tau_4, \tau_6), \min[\max(\tau_2, \tau_3), ]], \tau_7\}, \quad (5.1)$$

де $T$ – час безвідмовної роботи системи. В одному досліді розігруються значення всіх $\tau_i, i = \overline{1,7}$, відповідно до $p_i(\tau_i), i = \overline{1,7}$. Використовуючи отримані реалізації $\tau_i, i = \overline{1,7}$, за (5.1) обчислюємо реалізацію $T$. Один дослід дає одну реалізацію (одне вибіркове значення) $T$. Проводимо $M$ дослідів (випробування), одержуємо «статистичний» матеріал (вибірку). Беремо середнє арифметичне часу безвідмовної роботи системи $T$ з $p$ як оцінку надійності системи. За необхідності можна побудувати закон розподілу ймовірностей випадкової величини $T$ у вигляді відповідної гістограми.

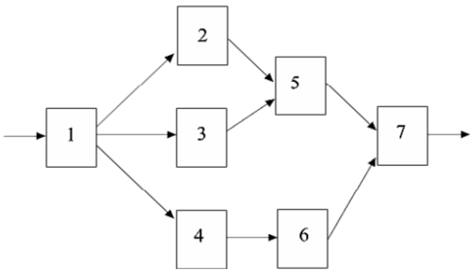


Рис. 5.1. Структура системи, що досліджується

Таким чином, випробування реальної системи замінені на випробування математичної моделі. Кожне випробування супроводжується розрахунками, тому імітаційне моделювання й називають числовим експериментом на ЕОМ із математичною моделлю (модель виступає як об’єкт дослідження).

Термін «метод Монте-Карло» (запропонований Дж. Фон Нейманом і С. М. Уламу в 40-х рр. XX ст.) належить до моделювання процесів із використанням генератора випадкових чисел. Термін «Монте-Карло» (місто, широко відоме своїми казино) виник від того факту, що «кількість шансів» (методи моделювання Монте-Карло) було використано з метою знаходження інтегралів від складних рівнянь при розробці перших ядерних бомб (інтеграли квантової механіки). За допомогою формування великих вибірок випадкових чисел із, наприклад, декількох розподілів, інтеграли цих (складних) розподілів можуть бути апроксимовані зі (згенерованих) даних.

Спочатку метод Монте-Карло використовувався, передусім, для розв’язання задач нейтронної фізики, де традиційні числові методи виявилися малопридатними. Далі його вплив поширився на широкий клас задач статистичної фізики, дуже різних за своїм змістом. До розділів науки, де дедалі більшою мірою використовується метод Монте-Карло, слід віднести задачі теорії масового обслуговування, задачі теорії ігор і математичної економіки, задачі теорії передачі повідомлень за наявності перешкод та ряд інших.

Метод Монте-Карло впливав і продовжує впливати на розвиток методу обчислювальної математики (наприклад, розвиток методів числового інтегрування) і при розв’язанні багатьох задач успішно поєднується з іншими обчислювальними методами та доповнює їх. Його застосування виправдане, насамперед, у тих задачах, які допускають теоретико-імовірнісний опис. Це пояснюється як природністю одержання відповіді з певною заданою ймовірністю в задачах із імовірним змістом, так і істотним спрощенням процедури розв’язання. У тих випадках, коли наявний теоретико-імовірнісний опис задачі, використання методу Монте-Карло може суттєво спростити згадані її розв’язки.

## 5.2. Загальна схема методу Монте-Карло

Метод Монте-Карло – це числовий метод дослідження математичних моделей складних систем, оснований на моделюванні випадкових елементів і подальшому статистичному аналізі результатів моделювання.

Припустимо, що нам потрібно обчислити певну невідому величину $m$, і ми прагнемо зробити це, розглядаючи випадкову величину $\xi$, таку, в якій математичне спдівання $E\{\xi\} = m$. Нехай при цьому дисперсія даної випадкової величини $D\{\xi\} = b$.

Розглянемо $N$ випадкових незалежних величин $\xi_1, \xi_2, \dots, \xi_n$, розподіли яких збігаються з розподілом розглянутої випадкової величини $\xi$. Відповідно до центральної граничної теореми теорії ймовірностей розподіл суми $p(n) = \xi_1, \xi_2, \dots, \xi_n$ буде приблизно нормальним із середнім, таким, що дорівнює $\mu = Nm$, і дисперсією $\sigma^2 = Nb^2$. Згідно з «правилом трьох сигм», якими б не були $m$ і $\sigma$,

$$P\{\mu - 3\sigma < \rho_n < \mu + 3\sigma\} = 0{,}997,$$

тобто

$$P\left\{m - \frac{3b}{\sqrt{N}} < \frac{\rho_n}{N} < m + \frac{3b}{\sqrt{N}}\right\} \approx 0{,}997.$$

Останнє співвідношення можна переписати у вигляді

$$\left\{\frac{1}{N}\sum_{i=1}^{N}\xi_i - m < \frac{\rho_n}{N} < \frac{3b}{\sqrt{N}}\right\} \approx 0{,}997.$$

Отримана формула дає метод розрахунку $m$ і оцінку похибки цього методу.

Суть застосування методу Монте-Карло полягає у визначенні результатів на основі статистики, одержуваної до моменту прийняття певного розв'язку. Тому імовірність результатів, одержуваних при використанні методу Монте-Карло, вирішально визначається якістю генератора випадкових чисел.

Для одержання випадкових чисел на ЕОМ використовуються способи генерування, які зазвичай основані на багаторазовому повторенні певної операції. Отриманій у такий спосіб послідовності більше відповідає назва псевдовипадкових чисел, оскільки генерована послідовність є періодичною й, починаючи з певного моменту, числа почнуть повторюватися.

Основним недоліком методу Монте-Карло є імовірнісний характер результату, тобто відсутність строгих оцінок похибки, що прямують до нуля при $n \to \infty$.

Розглянемо використання методу Монте-Карло на прикладі обчислення інтеграла, значення якого аналітичним способом знайти не вдається.

Потрібно знайти значення інтеграла

$$y = \int_{x_1}^{x_2} f(x)dx.$$

На рис. 5.2 поданий графік функції $f(x)$. Обчислити значення інтеграла цієї функції – значить, знайти площу під цим графіком.

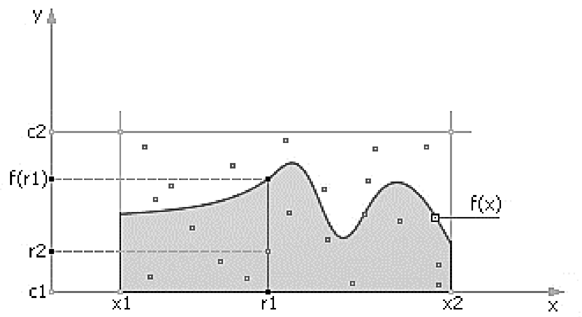


Рис. 5.2. Визначення значення інтеграла методом Монте-Карло

Обмежуємо криву зверху, праворуч і ліворуч. Випадковим чином розподіляємо точки в прямокутнику пошуку. Позначимо через $N_1$ кількість точок, прийнятих для випробувань, тобто тих, що потрапили в прямокутник, та через $N_2$ – кількість точок під кривою, тобто тих, які потрапили в зафарбовану площу під функцією. Тоді природно припустити, що кількість точок, які потрапили під криву відносно загальної кількості точок пропорційна площі під кривою (величині інтеграла) відносно площі випробуваного прямокутника. Математично це можна виразити так:

$$\frac{N_2}{N_1} = \frac{y}{(x_2 - x_1)(c_2 - c_1)}.$$

Ці міркування, звичайно, статистичні і тим більш правильніші, чим більшу кількість випробуваних точок ми візьмемо.

Значення $r_1$ і $r_2$ на рис. 5.2 є рівномірно розподіленими випадковими числами з інтервалів $(x_1;\ x_2)$ та $(c_1;\ c_2)$ відповідно.

Метод Монте-Карло є надзвичайно ефективним, простим, але необхідний «якісний» генератор випадкових чисел. Друга проблема застосування методу полягає у визначенні об’єму вибірки, тобто кількості точок, необхідних для забезпечення розв’язання із заданою точністю. Експерименти свідчать: щоб збільшити точність у 10 разів, об’єм вибірки потрібно збільшити в 100 разів; тобто точність приблизно пропорційна квадратному кореню з об’єму вибірки:

$$\text{точність} \cong \sqrt{\text{об’єм вибірки}}.$$

Розглянемо схему використання методу Монте-Карло при дослідженні систем із випадковими параметрами.

Побудувавши модель системи з випадковими параметрами, на її вхід подають вхідні сигнали від генератора випадкових чисел (ГВЧ), як показано на рис. 5.3. Цей генератор побудований так, що він видає *рівномірно розподілені* випадкові числа $r_{\text{рр}}$ із інтервалу [0; 1]. Оскільки одні події можуть бути більш імовірними, інші – менш імовірними, то рівномірно розподілені випадкові числа від генератора подають на перетворювач закону випадкових чисел (ПЗВЧ), який перетворює їх у *заданий* користувачем закон розподілу ймовірності, наприклад, у нормальний або експоненційний закон. Ці перетворювані випадкові числа $x$ подають на вхід моделі. Модель відпрацьовує вхідний сигнал $x$ за певним законом $y = \varphi(x)$ та отримує вихідний сигнал $y$, який також є випадковим.

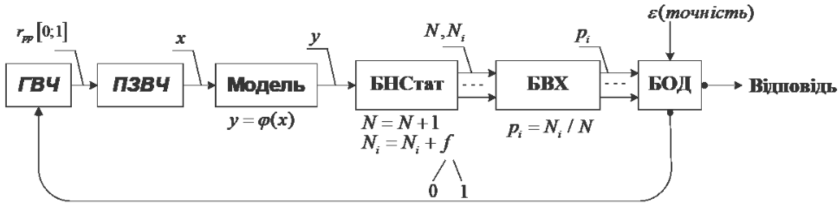


Рис. 5.3. Загальна структура методу статистичного моделювання

У блоці накопичення статистики (БНСтат) установлені фільтри й лічильники. Фільтр (певна логічна умова) визначає за значенням $y$, чи відбулася в конкретному досліді певна подія (виконалася умова, $f = 1$), чи ні (умова не виконалася, $f = 0$). Якщо подія відбулася, то лічильник події збільшується на одиницю. Якщо подія не реалізувалася, то значення лічильника не змінюється. Якщо потрібно стежити за декількома різними типами подій, то для статистичного моделювання знадобиться декілька фільтрів і лічильників $N_i$. Завжди ведеться облік кількості експериментів – $N$.

Далі відношення $N_i$ до $N$, що розраховується в блоці обчислення статистичних характеристик (БОСХ) за методом Монте-Карло, дає оцінку ймовірності $P_i$ появи події $i$, тобто вказує на частоту його випадання в серії із $N$ дослідів. Це дозволяє зробити висновки про статистичні властивості модельованого об’єкта.

Наприклад, подія $A$ відбулася в результаті проведених 200 експериментів 50 разів. Це означає, що згідно з методом Монте-Карло, імовірність здійснення події дорівнює: $P_A = 50/200 = 0{,}25$. Імовірність того, що подія не відбудеться, дорівнює, відповідно, $1 - 0{,}25 = 0{,}75$.

***Слід звернути увагу на те***, що коли говорять про ймовірність, отриману експериментально, то її називають частістю, а слово «ймовірність» уживають, коли прагнуть підкреслити, що мова йде про теоретичне поняття.

За великої кількості дослідів $N$ частота появи події, отримана експериментальним шляхом, наближається до значення теоретичної ймовірності появи події.

У блоці оцінки достовірності (БОД) аналізують ступінь достовірності статистичних експериментальних даних, знятих із моделі (беручи до уваги точність результату ε, задану користувачем), і визначають необхідну для цього кількість статистичних випробувань. Якщо коливання значень частоти появи подій відносно теоретичної ймовірності менші за задану точність, то експериментальну частоту приймають як відповідь, інакше генерація випадкових вхідних впливів продовжиться, і процес моделювання повториться. За малої кількості випробувань результат може виявитися недостовірним. Але чим більше випробувань, тим точніша відповідь, відповідно до центральної граничної теореми.

Зазначимо, що оцінювання ведуть за гіршою із частот. Це забезпечує достовірний результат відразу за всіма характеристиками моделі, які знімаються.

**Приклад 1.** Розв'яжемо просту задачу. Яка ймовірність випадання монети орлом догори при падінні її з висоти випадково?

Почнемо підкидати монетку й фіксувати результати кожного кидка (див. табл. 5.1).

Таблиця 5.1 – Результати випробувань кидання монети

| Кількість дослідів $N$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Значення лічильника випадання орла $N_{\text{о}}$ | 0 | 0 | 1 | 1 | 2 | 3 | 4 | … | … | … | … | … | … | … |
| Значення лічильника випадання решки $N_{\text{р}}$ | 1 | 2 | 2 | 3 | 3 | 3 | 3 | … | … | … | … | … | … | … |
| Частість випадання орла $P_{\text{о}} = N_{\text{о}}/N$ | 0 | 0 | 0,33 | 0,25 | 0,4 | 0,5 | 0,57 | … | … | … | … | … | … | … |
| Частість випадання решки $P_{\text{р}} = N_{\text{р}}/N$ | 1 | 1 | 0,66 | 0,75 | 0,6 | 0,5 | 0,43 | … | … | … | … | … | … | … |

Будемо підраховувати частість випадання орла як відношення кількості випадків випадання орла до загальної кількості спостережень. Подивіться в табл. 5.1, випадки для $N = 1,\ N = 2, N = 3$ – спочатку значення частості не можна назвати достовірними. Спробуємо побудувати графік залежності $P_{\text{о}}$ від $N$ і подивимося, як змінюється частість випадання орла залежно від кількості проведених дослідів. Зрозуміло, що при різних експериментах будуть виходити різні таблиці, а отже, різні графіки. На рис. 5.4 показано один із варіантів такого графіка.

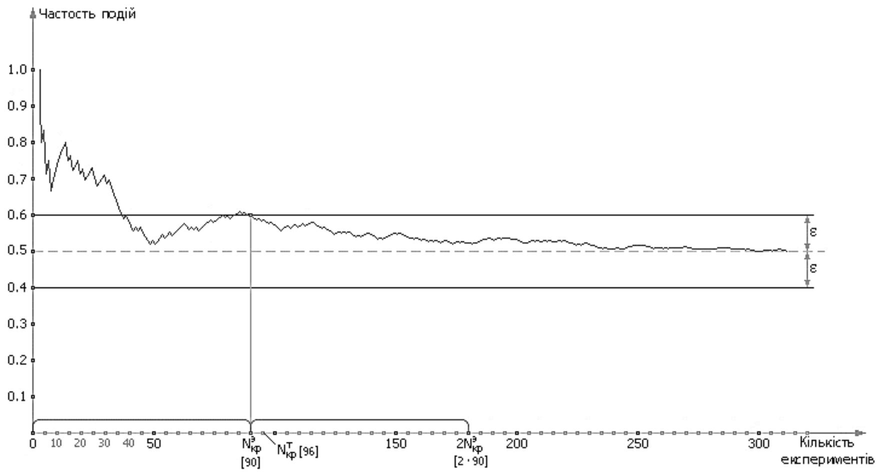


Рис. 5.4. Експериментальна залежність частості появи випадкової події

З наведених експериментальних залежностей видно, що при малих значеннях $N$, наприклад, $N = 1, N = 2, N = 3$, відповіді взагалі довіряти не можна. Наприклад, $P_o = 0$ при $N = 1$, тобто ймовірність випадання орла при одному кидку дорівнює нулю! Хоча всім добре відомо, що це не так. Тобто поки що ми одержали дуже неточну відповідь. Однак слід подивитися на графік: у процесі нагромадження інформації відповідь повільно, але вірно наближається до правильної (відповідь виділена пунктирною лінією). На щастя, у цьому конкретному випадку правильна відповідь нам відома: в ідеалі, імовірність випадання орла дорівнює 0,5 (в інших, більш складних задачах, відповідь нам, звичайно, буде невідома). Припустимо, що відповідь нам потрібно знати з точністю $\varepsilon = 0{,}1$. Проведемо дві паралельні лінії, що стоять від правильної відповіді 0,5 на відстані 0,1. Ширина коридору, який утворився, буде дорівнювати 0,2. Як тільки крива $P_o(N)$ увійде в цей коридор так, що вже ніколи його не покине, можна буде зупинитися й подивитися, для якого значення $N$ це відбулося. Це і є експериментально обчислене критичне значення необхідної кількості дослідів $N_{kr}^{e}$ для визначення відповіді з точністю $\varepsilon = 0{,}1$; ε-*окіл* у наших міркуваннях відіграє роль своєрідної трубки точності. Слід звернути увагу на те, що відповіді $P_o(91)$, $P_o(92)$ і т. д. уже не змінюють сильно своїх значень: принаймні, у них не змінюється перша цифра після коми, якій ми зобов'язані довіряти за умовами задачі.

Причиною такої поведінки кривої є дія центральної граничної теореми. Поки тут ми сформулюємо її в найпростішому варіанті: «Сума випадкових величин є величиною невипадковою». Ми використали середню величину $P_0$, яка несе в собі інформацію про суму дослідів, і тому поступово ця величина стає дедалі більш достовірною.

Якщо виконати ще раз цей дослід спочатку, то, звичайно, його результатом буде інший вид випадкової кривої, і відповідь буде іншою, хоча й приблизно такою ж. Проведемо цілу серію таких експериментів (рис. 5.5). Така серія називається ансамблем реалізацій. Якій же відповіді в підсумку слід вірити? Адже вони, хоч і є близькими, та відрізняються. На практиці роблять по-різному. Перший варіант – обчислити середнє значення відповідей за кілька реалізацій (табл. 5.2).

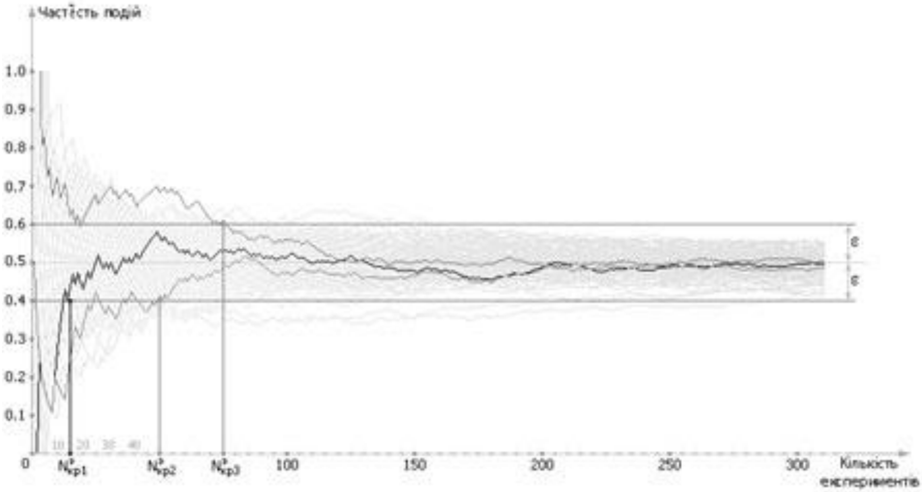


Рис. 5.5. Експериментально знятий ансамбль випадкових залежностей частості появи випадкової події

Ми поставили декілька експериментів і щороку визначали, скільки необхідно було зробити дослідів, тобто $N_{kr}^{e}$. Було виконано 10 експериментів, результати яких були зведені в табл. 5.2. За результатами десяти експериментів було обчислене середнє значення $N_{kr}^{e}$.

Таким чином, провівши 10 реалізацій різної довжини, ми визначили, що в середньому достатньо було зробити 1 реалізацію завдовжки в 94 кидки монети.

Ще один важливий результат реалізацій – 100 ліній. Відзначили на ньому абсцису $N = 94$ вертикальною рискою. Існує якийсь відсоток ліній, які не встигнули перетнути *$\varepsilon$-окіл*, тобто ($P^{\text{eksp}} - \varepsilon \leq P^{\text{teor}} \leq P^{\text{eksp}} + \varepsilon$), та увійти в коридор точності до моменту $N = 94$. Слід звернути увагу на те, що таких ліній 5. Це значить, що 95 зі 100, тобто 95 ліній імовірно увійшли в позначений інтервал.

Таблиця 5.2 – Експериментальні дані необхідної кількості кидків монети для досягнення точності $\boldsymbol{\varepsilon = 0,1}$

| Дослід | $N_{kr}^{e}$ |
|---|---|
| 1 | 288 |
| 2 | 95 |
| 3 | 50 |
| 4 | 29 |
| 5 | 113 |
| 6 | 210 |
| 7 | 30 |
| 8 | 42 |
| 9 | 39 |
| 10 | 48 |
| Середнє $N_{kr}^{e}$ | 94 |

Таким чином, провівши 100 реалізацій, ми досягли приблизно 95 %-тої довіри до отриманої експериментально величини ймовірності випадання орла, визначивши її з точністю 0,1. Для порівняння отриманого результату обчислимо теоретичне значення $N_{kr}^{t}$ теоретично. Однак для цього доведеться увести поняття надійної ймовірності $Q_F$, яка показує, наскільки ми готові вірити відповіді. Наприклад, при $Q_F = 0{,}95$ ми готові вірити відповіді в 95 % випадків зі 100. Формула теоретичного розрахунку кількості експериментів, як нами буде показано далі, має вигляд

$$N_{kr}^{t} = k(Q_F)P(1 - P)/\varepsilon^2 \ ,$$

де $k(Q_F)$ – коефіцієнт Лапласа; $P$ – імовірність випадання орла; $\varepsilon$ – точність (довірчий інтервал). У таблиці 5.3 показані значення теоретичної величини кількості необхідних дослідів при різних $Q_F$ (для точності $\varepsilon = 0{,}1$ та ймовірності $P = 0{,}5$).

Як бачимо, отримана нами оцінка довжини реалізації, що дорівнює 94 дослідам, дуже близька до теоретичної, яка дорівнює 96. Деяка розбіжність пояснюється тим, що, очевидно, десяти реалізацій не достатньо для точного обчислення

$N_{kr}^{e}$. Якщо ми вирішимо, що нам потрібний результат, якому слід довіряти більше, то змінимо значення надійної ймовірності. Наприклад, теорія свідчить, що якщо дослідів буде 167, то всього 1–2 лінії з ансамблю не увійдуть у запропоновану трубку точності. Але слід мати на увазі те, що кількість експериментів зі зростанням точності й достовірності росте дуже швидко.

Таблиця 5.3 – Теоретичний розрахунок необхідної кількості кидків монети для досягнення точності $\varepsilon = 0{,}1$

| Надійна ймовірність $Q_F$ | Коефіцієнт Лапласа $k(Q_F)$ | Необхідна кількість дослідів $N_{kr}^{t} = k(Q_F)P(1-P)/\varepsilon^2$ |
|---|---|---|
| 0,90 | 2,72 | 68 |
| 0,95 | 3,84 | 96 |
| 0,99 | 6,66 | 167 |

Якщо придивитися до ансамблю випадкових реалізацій, то можна виявити, що збіжність частості до значення теоретичної ймовірності відбувається за кривою, яка відповідає зворотній квадратичній залежності від кількості експериментів (рис. 5.6).

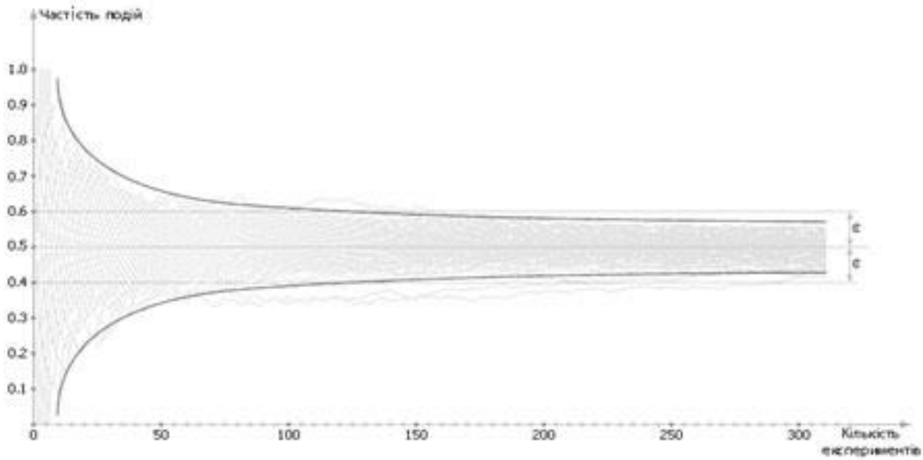


Рис. 5.6. Ілюстрація швидкості сходження експериментально одержуваної частості до теоретичної ймовірності

Це дійсно так виходить і теоретично. Якщо змінювати точність $\varepsilon$, що задається, та досліджувати кількість експериментів, необхідних для забезпечення кожної з них, то вийде табл. 5.4.

Таблиця 5.4 – Теоретична залежність кількості експериментів, необхідних для забезпечення заданої точності при $Q_F = 0{,}95$

| Точність $\varepsilon$ | Критична кількість експериментів $N_{kr}^T$ |
|---|---|
| 0,1 | 96 |
| 0,01 | 9600 |
| 0,001 | 960000 |

Побудуємо за табл. 5.4 графік залежності $N_{kr}^T(\varepsilon)$ (рис. 5.7).

Отже, розглянуті графіки підтверджують наведену вище оцінку:

$$\text{точність} \cong \sqrt{\phantom{x}}\ \text{об}'\text{єм вибірки}.$$

Зазначимо, що оцінок точності може бути кілька.

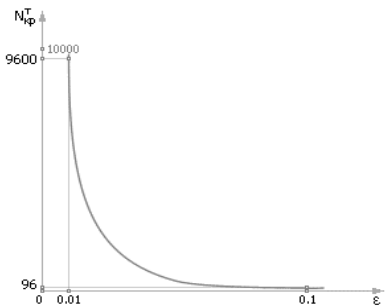


Рис. 5.7. Залежність кількості експериментів, необхідних для досягнення заданої точності ε при фіксованому $Q_F = 0{,}95$

**Приклад 2**. Знаходження площі фігури методом Монте-Карло. Визначимо методом Монте-Карло площу п’ятикутника з координатами кутів (0, 0), (0, 10), (5, 20), (10, 10), (7, 0).

Намалюємо у двовимірних координатах заданий п'ятикутник, вписавши його в прямокутник, чия площа, як неважко здогадатися, становить (10 – 0) · (20 – 0) = 200 (рис. 5.8).

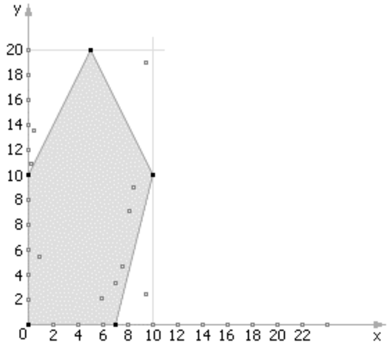


Рис. 5.8. Ілюстрація до розв'язання задачі про площу фігури

Таблиця 5.5 – Розв'язок задачі методом Монте-Карло

| Номер точки | $R$ | $G$ | $X$ | $Y$ | Точка $(X;\ Y)$ потрапила в прямокутник? | Точка $(X;\ Y)$ потрапила в п'ятикутник? |
|---|---|---|---|---|---|---|
| 1 | 0,8109 | 0,3557 | 8,109 | 7,114 | Так | Так |
| 2 | 0,0333 | 0,5370 | 0,333 | 10,740 | Так | Ні |
| 3 | 0,1958 | 0,2748 | 1,958 | 5,496 | Так | Так |
| 4 | 0,6982 | 0,1652 | 6,982 | 3,304 | Так | Так |
| 5 | 0,9499 | 0,1090 | 9,499 | 2,180 | Так | Ні |
| 6 | 0,7644 | 0,2194 | 7,644 | 4,388 | Так | Так |
| 7 | 0,8395 | 0,4510 | 8,395 | 9,020 | Так | Так |
| 8 | 0,0415 | 0,6855 | 0,415 | 13,710 | Так | Ні |
| 9 | 0,5997 | 0,1140 | 5,997 | 2,280 | Так | Так |
| 10 | 0,9595 | 0,9595 | 9,595 | 19,190 | Так | Ні |
| | | | | Усього: | 10 | 6 |

Використаємо таблицю випадкових чисел для генерації пар чисел $R, G$, рівномірно розподілених в інтервалі від 0 до 1. Число $R$ буде імітувати координату $X(0 \leq X \leq 10)$, отже, $X = 10 \cdot R$. Число $G$ буде імітувати координату $(0 \leq Y \leq$ $\leq 20)$, отже, $Y = 20 \cdot G$. Згенеруємо по 10 чисел $R$ і $G$ і відобразимо 10 точок $(X;\ Y)$ на рис. 5.8 і в табл. 5.5.

Статистична гіпотеза полягає в тому, що кількість точок, які потрапили в контур фігури, пропорційні площі фігури: $6\!:10\ =\ S\!:200$. Тобто за формулою методу Монте-Карло одержуємо, що площа $S$ п’ятикутника дорівнює: 200 · 6/10 = 120.

Простежимо, як змінювалася величина $S$ від досліду до досліду (табл. 5.6).

Оскільки у відповіді все ще змінюється значення другого розряду, то можлива неточність становить поки що більше 10 %. Точність розрахунків може бути збільшена зі зростанням кількості випробувань (рис. 5.9).

Таблиця 5.6 – Оцінка точності відповіді

| Кількість випробувань $N$ | Оцінка ймовірності потрапляння випадкової точки у випробувану область | Оцінка площі $S$ методом Монте-Карло |
|---|---|---|
| 1 | 1/1 = 1,00 | 200 |
| 2 | 1/2 = 0,50 | 100 |
| 3 | 2/3 = 0,67 | 133 |
| 4 | 3/4 = 0,75 | 150 |
| 5 | 3/5 = 0,60 | 120 |
| 6 | 4/6 = 0,67 | 133 |
| 7 | 5/7 = 0,71 | 143 |
| 8 | 5/8 = 0,63 | 125 |
| 9 | 6/9 = 0,67 | 133 |
| 10 | 6/10 = 0,60 | 120 |

Точне значення площі п’ятикутника

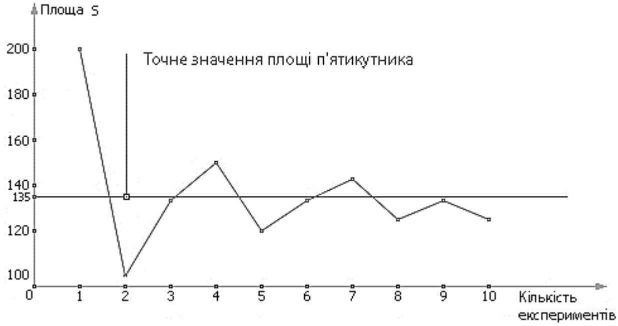


Рис. 5.9. Ілюстрація процесу збіжності, зумовленої експериментально, відповіді до теоретичного результату

## 5.3. Генератори випадкових чисел

В основі методу Монте-Карло лежить генерація випадкових чисел, які повинні бути рівномірно розподілені в інтервалі $(0; 1)$.

Якщо генератор видає числа, зміщені в якусь частину інтервалу (одні числа випадають частіше за інші), то результат розв’язання задачі, розв’язуваної статистичним методом, може виявитися неправильним. Тому задача використання якісного генератора дійсно випадкових і дійсно рівномірно розподілених чисел стоїть дуже гостро.

Математичне очікування $m_r$ і дисперсія $D_r$ такої послідовності, що складається з $n$ випадкових чисел $r_i$, повинні бути такими (якщо це дійсно рівномірно розподілені випадкові числа в інтервалі від 0 до 1):

$$m_r = \frac{\sum_{i=1}^{n} r_i}{n} = 0{,}5,\ \ D_r = \frac{\sum_{i=1}^{n} (r_i - m_r)^2}{n} = \frac{1}{12}.$$

Якщо користувачеві буде потрібно, щоб випадкове число $x$ перебувало в інтервалі $(a;\ b)$, відмінному від $(0;\ 1)$, то необхідно буде скористатися формулою $x = a + (b - a)r$, де $r$ – випадкове число з інтервалу $(0;\ 1)$. Законність цього перетворення наведено на рис. 5.10.

Тепер $x$ – випадкове число, рівномірно розподілене в діапазоні від $a$ до $b$.

Як еталон генератора випадкових чисел (ГВЧ) прийнятий такий генератор, який породжує послідовність випадкових чисел рівномірним законом розподілу в інтервалі (0; 1). За один оберт цій генератор повертає одне випадкове число. Якщо спостерігати такий ГВЧ досить тривалий час, то виявиться, що, наприклад, у кожний з десяти інтервалів (0; 0,1), (0,1; 0,2), (0,2; 0,3), …, (0,9; 1) потрапить практично однакова кількість випадкових чисел, тобто вони будуть розподілені рівномірно по всьому інтервалу (0; 1). Якщо зобразити на графіку $k = 10$ інтервалів і частоти $N_i$ потраплянь у них, то вийде експериментальна крива щільності розподілу випадкових чисел (рис. 5.10).

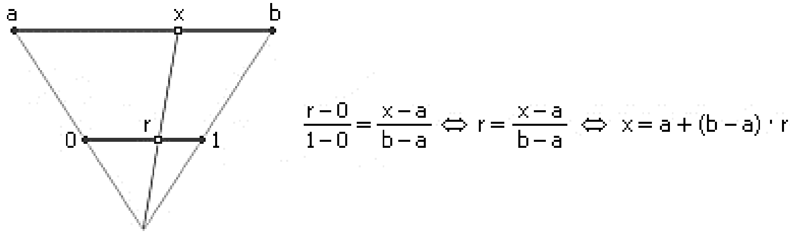


Рис. 5.10. Схема переведення числа з інтервалу $(0;\ 1)$ в інтервал $(a;\ b)$

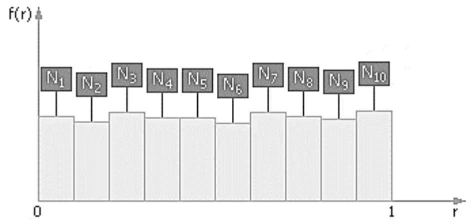


Рис. 5.11. Частотна діаграма випадання випадкових чисел, породжуваних реальним генератором

Зазначимо, що в ідеальному випадку в кожний інтервал потрапляє однакова кількість точок: $N_i = N/k$, де $N$ – загальна кількість точок, $k$ – кількість інтервалів, $i = 1,\ \dots,\ k$.

Таким чином, генерація довільного випадкового числа складається із двох етапів:

- генерація нормалізованого випадкового числа (тобто рівномірно розподіленого від 0 до 1);
- перетворення нормалізованих випадкових чисел $r_i$ у випадкові числа $x_i$, розподілені за необхідним користувачеві (довільним) законом розподілу або в необхідному інтервалі.

## 5.4. Перевірка якості роботи генератора

Від якості роботи ГВЧ залежить якість роботи всієї системи й точність результатів. Тому випадкова послідовність, породжувана ГВЧ, має задовольняти цілому ряду критеріїв.

Здійснювані перевірки (цього) двох типів:
- перевірки на рівномірність розподілу;
- перевірки на статистичну незалежність.

### *5.4.1. Перевірки на рівномірність розподілу*

Генератор випадкових чисел повинен видавати близькі до наступних значення статистичних параметрів, характерних для рівномірного випадкового закону

$$m_r = \frac{\sum_{i=1}^{n} r_i}{n} = 0{,}5 \text{ – математичне очікування;}$$

$$D_r = \frac{\sum_{i=1}^{n} (r_i - m_r)^2}{n} = \frac{1}{12} \text{ – дисперсія;}$$

$$\sigma = \sqrt{D_r} \approx 0{,}2887 \text{ – середньоквадратичне відхилення.}$$

Частотний тест. Частотний тест дозволяє з'ясувати, скільки чисел потрапило в інтервал $(m_r - \sigma_r;\ m_r + \sigma_r)$, тобто $(0{,}5 - 0{,}2887; 0{,}5 + 0{,}2887)$ або, в остаточному підсумку, $(0{,}2113;\ 0{,}7887)$. Оскільки $0{,}7887 - 0{,}2113 = 0{,}5774$, припустимо, що в якісному ГВЧ у цей інтервал повинно потрапляти близько $57{,}7\ \%$ з усіх випадкових чисел, що випали (рис. 5.12).

Також необхідно враховувати, що кількість чисел, які потрапили в інтервал (0; 0,5), повинна приблизно дорівнювати кількості чисел, які потрапили в інтервал (0,5; 1,0).

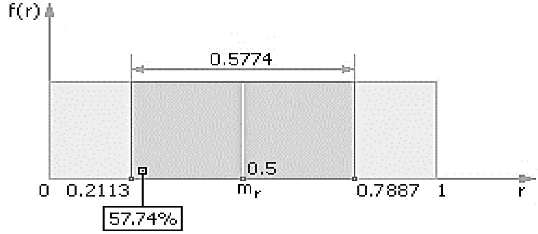


Рис. 5.12. Частотна діаграма ідеального ГВЧ у випадку перевірки його на частотний тест

Перевірка за критерієм «$xi$-квадрат». Критерій «$xi$-квадрат» ($\chi^2$ – критерій) – це один із найвідоміших статистичних критеріїв; він є основним методом, використовуваним у комбінації з іншими критеріями. Критерій «$xi$-квадрат» був запропонований в 1900 р. Карлом Пірсоном. Його дослідження розглядається як фундамент сучасної математичної статистики.

Для нашого випадку перевірка за критерієм «$xi$-квадрат» дозволить дізнатися, чи задовольняє він вимозі рівномірного розподілу, чи ні. Створений нами *реальний* ГВЧ близький до еталона ГВЧ.

Оскільки закон розподілу еталонного ГВЧ рівномірний, то (теоретична) імовірність $P_i$ потрапляння чисел в $i$-й інтервал (усього цих інтервалів $k$) дорівнює $P_i = 1/k$. Таким чином, у кожний з $k$ інтервалах потрапить *рівно* по $P_i \cdot N$ чисел ($N$– загальна кількість згенерованих чисел).

Реальний ГВЧ буде видавати числа, розподілені (причому, необов'язково рівномірно!) по $k$ інтервалах, і в кожний інтервал потрапить по $n_i$ чисел (у сумі $n_1 + n_2 + \cdots + n_k = N$). Як же нам визначити, наскільки ГВЧ, який випробовується, якісний і близький до еталонного? Цілком логічно розглянути квадрати різниць між отриманою кількістю чисел $n_i$ та «еталонним» $P_i N$. Складемо їх, а в результаті одержимо

$$\chi^2_{\text{eksp}} = (n_1 - P_1 N)^2 + (n_2 - P_2 N)^2 + \cdots + (n_k - P_k N)^2.$$

Із цієї формули випливає, що чим менша різниця в кожному з доданків, (а значить і чим менше значення $\chi^2_{\text{eksp}}$), тим сильніше закон розподілу випадкових чисел, що генеруються реальним ГВЧ, тяжіє до рівномірного.

У попередньому виразі кожному з доданків приписується однакова вага (дорівнює 1), що насправді може не відповідати дійсності; тому для статистики «$xi$-квадрат» необхідно провести нормування кожного $i$-го доданка, поділивши його на $P_i \cdot N$:

$$\chi^2_{\text{eksp}} = \frac{(n_1 - P_1 N)^2}{P_1 N} + \frac{(n_2 - P_2 N)^2}{P_2 N} + \cdots + \frac{(n_k - P_k N)^2}{P_k N}.$$

Нарешті, запишемо отриманий вираз більш компактно й спростимо його

$$\chi^2_{\text{eksp}} = \sum_{i=1}^{k} \frac{(n_i - P_i N)^2}{P_i N} = \frac{1}{N} \sum_{i=1}^{k} \left( \frac{n_i^2}{P_i} \right) - N.$$

Ми одержали значення критерію «хі-квадрат» для *експериментальних* даних.

Деякі відсоткові точки $\chi^2$-розподілу

$$\nu + sqrt(2\nu)\, x_p + 2/3 \cdot x_{2p} - 2/3 + O(1/sqrt(\nu)).$$

У табл. 5.7 наведені *теоретичні* значення «$xi$-квадрат» ($\chi^2_{\text{teor}}$), де $\nu = N - 1$ – це кількість степенів вільності, $P$ – це надійна ймовірність, що задається користувачем, який указує, наскільки ГВЧ повинен задовольняти вимогам рівномірного розподілу, або *P – це ймовірність того, що експериментальне значення* $\chi^2_{\text{eksp}}$ *буде меншим за табульоване (теоретичне)* $\chi^2_{\text{teor}}$, *або дорівнює йому. Прийнятним вважають P від 10 до 90 %.*

Якщо $\chi^2_{\text{eksp}}$ набагато більше за $\chi^2_{\text{teor}}$ (тобто $P$ – велике), то генератор не задовольняє вимозі рівномірного розподілу, оскільки спостережувані значення $n_i$ занадто далеко відходять від теоретичних $P_i N$ і не можуть розглядатися як випадкові. Інакше кажучи, установлюється такий великий довірчий інтервал, що обмеження на числа стають дуже нежорсткими, а вимоги до чисел – слабкими. При цьому буде спостерігатися дуже велика абсолютна похибка.

Мати $\chi^2_{\text{eksp}}$ маленьким загалом також не дуже добре, хоча це й здається, на перший погляд, чудово з погляду рівномірності. Дійсно, візьмемо ряд чисел 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; … – вони ідеальні з погляду рівномірності, і $\chi^2_{\text{eksp}}$ буде практично нульовим, але навряд чи ми їх визнаємо випадковими.

Якщо $\chi^2_{\text{eksp}}$ набагато менше за $\chi^2_{\text{teor}}$ (тобто $P$ – мале), то генератор не задовольняє вимозі випадкового рівномірного розподілу, оскільки спостережувані значення $n_i$ занадто близькі до теоретичних $P_i \cdot N$ і не можуть розглядатися як випадкові.

А от якщо $\chi^2_{\text{eksp}}$ лежить у певному діапазоні між двома значеннями $\chi^2_{\text{teor}}$, які відповідають, наприклад, $P = 25$ % і $P = 50$ %, то можна вважати, що значення випадкових чисел, породжувані датчиком, цілком є випадковими.

При цьому додатково слід мати на увазі, що всі значення $P_i \cdot N$ повинні бути досить великими, наприклад більше 5 (з’ясовано емпіричним шляхом). Тільки тоді (при досить великій статистичній вибірці) умови проведення експерименту можна вважати задовільними.

Таблиця 5.7 – Теоретичні значення «$xi$-квадрат»

| | $P$ = 1 % | $P$ = 5 % | $P$ = 25 % | $P$ = 50 % | $P$ = 75 % | $P$ = 95 % | $P$ = 99 % |
|---|---|---|---|---|---|---|---|
| $v$ = 1 | 0,0001 | 0,00393 | 0,1015 | 0,4549 | 1,323 | 3,841 | 6,635 |
| $v$ = 2 | 0,02010 | 0,1026 | 0,5754 | 1,386 | 2,773 | 5,991 | 9,210 |
| $v$ = 3 | 0,1148 | 0,3518 | 1,213 | 2,366 | 4,108 | 7,815 | 11,34 |
| $v$ = 4 | 0,2971 | 0,7107 | 1,923 | 3,357 | 5,385 | 9,488 | 13,28 |
| $v$ = 5 | 0,5543 | 1,1455 | 2,675 | 4,351 | 6,626 | 11,07 | 15,09 |
| $v$ = 6 | 0,8721 | 1,635 | 3,455 | 5,348 | 7,841 | 12,59 | 16,81 |
| $v$ = 7 | 1,239 | 2,167 | 4,255 | 6,346 | 9,037 | 14,07 | 18,48 |
| $v$ = 8 | 1,646 | 2,733 | 5,071 | 7,344 | 10,22 | 15,51 | 20,09 |
| $v$ = 9 | 2,088 | 3,325 | 5,899 | 8,343 | 11,39 | 16,92 | 21,67 |
| $v$ = 10 | 2,558 | 3,940 | 6,737 | 9,342 | 12,55 | 18,31 | 23,21 |
| $v$ = 11 | 3,053 | 4,55 | 7,584 | 10,34 | 13,7 | 19,68 | 24,72 |
| $v$ = 12 | 3,71 | 5,226 | 8,438 | 11,34 | 14,85 | 21,03 | 26,22 |
| $v$ = 15 | 5,229 | 7,261 | 11,04 | 14,34 | 18,25 | 25,00 | 30,58 |
| $v$ = 20 | 8,260 | 10,85 | 15,45 | 19,34 | 23,83 | 31,41 | 37,57 |
| $v$ = 30 | 14,95 | 18,49 | 24,48 | 29,34 | 34,80 | 43,77 | 50,89 |
| $v$ = 50 | 29,71 | 34,76 | 42,94 | 49,33 | 56,33 | 67,50 | 76,15 |
| $v$ > 30 | | | | | | | |
| $x_p$ = | –2,33 | –1,64 | –0,674 | 0,00 | 0,674 | 1,64 | 2,33 |

Отже, процедура перевірки має такий вигляд.

1. Діапазон від 0 до 1 розбивається на $k$ рівних інтервалів.
2. Запускається ГВЧ $N$ разів ($N$ повинне бути велике, наприклад, $N/k > 5$).
3. Визначається кількість випадкових чисел, які потрапили в кожний інтервал: $n_i, i = 1, \dots, k$.
4. Обчислюється експериментальне значення $\chi^2_{\text{eksp}}$ за такою формулою:

$$\chi^2_{\text{eksp}} = \sum_{i=1}^{k} \frac{(n_i - P_i N)^2}{P_i N} = \frac{1}{N} \sum_{i=1}^{k} \left( \frac{n_i^2}{P_i} \right) - N,$$

де $P_i = 1/k$ – теоретична ймовірність потрапляння чисел у $k$-й інтервал.

5. Шляхом порівняння експериментально отриманого значення $\chi^2_{\text{eksp}}$ з теоретичним $\chi^2_{\text{teor}}$ (див. табл. 5.7) робиться висновок про придатність генератора для використання. Для цього:

– входимо в табл. 5.7 (рядок = кількість експериментів – 1);
– зрівнюємо обчислене $\chi^2_{\text{eksp}}$ з $\chi^2_{\text{teor}}$, які зустрічаються в рядку.

При цьому можливі три випадки.

Перший випадок: $\chi^2_{\text{eksp}}$ багато більше за будь-яке $\chi^2_{\text{teor}}$; у рядку – гіпотеза про випадковість рівномірного генератора не виконується (розкид чисел занадто великий, щоб бути випадковим).

Другий випадок: $\chi^2_{\text{eksp}}$ багато менше за будь-яке $\chi^2_{\text{teor}}$; у рядку – гіпотеза про випадковість рівномірного генератора не виконується (розкид чисел занадто малий, щоб бути випадковим).

Третій випадок: $\chi^2_{\text{eksp}}$ лежить між значеннями $\chi^2_{\text{teor}}$ двох стовпців, які стоять поряд, – гіпотеза про випадковість рівномірного генератора виконується з імовірністю $P$ ( тобто в $P$ випадках зі 100).

Зазначимо, що чим ближче виходить $P$ до значення 50 %, тим краще.

### *5.4.2. Перевірки на статистичну незалежність*

Здійснюється, як правило, перевірка двох типів:

- перевірка на частоту появи цифри в послідовності;
- перевірка появи серій з однакових цифр.

*Перевірка на частоту появи цифри в послідовності*. Розглянемо приклад. Випадкове число 0,2463389991 складається із цифр 2463389991, а число 0,5467766618 із цифр 5467766618. З’єднуючи послідовності цифр, маємо: 24633899915467766618.

Зрозуміло, що теоретична ймовірність $P_i$ випадання $i$-тої цифри (від 0 до 9) дорівнює 0,1.

Далі слід обчислити частоту появи кожної цифри, яка випала, в експериментальній послідовності. Наприклад, цифра 1 випала 2 рази з 20, а цифра 6 випала 5 разів із 20.

Далі рахують оцінку й ухвалюють рішення щодо критерію «хі-квадрат».

*Перевірка появи серій з однакових цифр*. Позначимо через $n_L$ кількість серій однакових підряд цифр довжини $L$. Перевіряти потрібно всі $L$ від 1 до $m$, де $m$ – це задана користувачем кількість однакових цифр: що максимально зустрічається в серії.

У прикладі «24633899915467766618» виявлено 2 серії завдовжки в 2 (33 і 77), тобто $n_2 = 2$ і 2 серії завдовжки в 3 (999 і 666), тобто $n_3 = 2$.

Імовірність появи серії завдовжки в $L$ дорівнює:

$$P_L = 9 \cdot 10^{-L} \text{ (теоретична).}$$

Тобто ймовірність появи серії завдовжки в один символ дорівнює: $P_1$ = 0,9 (теоретична). Імовірність появи серії завдовжки у два символи дорівнює: $P_2$ = 0,09 (теоретична). Імовірність появи серії завдовжки в три символи дорівнює: $P_3$ = 0,009 (теоретична).

Наприклад, імовірність появи серії завдовжки в один символ дорівнює $P_L$ = 0,9, оскільки може зустрітися лише один символ із 10, а всього символів – 9 (нуль не береться до уваги). Імовірність того, що підряд зустрінеться два однакові символи «XX», дорівнює 0,1 · 0,1 · 9, тобто ймовірність 0,1 того, що в першій позиції з’явиться символ «X», множиться на ймовірність 0,1 того, що в другій позиції з’явиться такий же символ «X» і множиться на кількість таких комбінацій 9.

Частість появи серій підраховується за раніше розібраною нами формулою «хі-квадрат» з використанням значень $P_L$.

*Примітка*: генератор може бути перевірений багаторазово, однак перевірки не мають властивості повноти й не гарантують, що генератор видасть випадкові числа. Наприклад, генератор, який видає послідовність 12345678912345…, при перевірках буде вважатися ідеальним, що, насправді, не зовсім так.

## 5.5. Моделювання випадкової події й повної групи несумісних подій

Почнемо з найпростішого. Використаємо наше вміння генерувати випадкові числа для імітації випадання випадкових подій.

Випадкова подія передбачає, що в деякої події є декілька результатів і те, який із наслідків, з’явиться в черговий раз, визначається тільки його ймовірністю. Тобто результат вибирається випадково з урахуванням його ймовірності.

Наприклад, припустимо, що нам відома ймовірність випуску бракованих виробів $P_b$= 0,1. Змоделювати випадання цієї події можна, розігравши рівномірно розподілене випадкове число з діапазону від 0 до 1 і встановивши, в який із двох інтервалів (від 0 до 0,1 або від 0,1 до 1) воно потрапило (рис. 5.13). Якщо число потрапляє в діапазон (0; 0,1), то випущений брак, тобто подія відбулася, якщо навпаки – подія не відбулася (випущений кондиційний виріб). При значній кількості експериментів частота потрапляння чисел в інтервал від 0 до 0,1 буде наближатися до ймовірності $P_b$ = 0,1, а частота влучення чисел в інтервал від 0,1 до 1 буде наближатися до $P_k$ = 0,9.

Зазначимо, що не важливо, як ви розташуєте на відрізку [0; 1] інтервал $P_b$ – на початку або наприкінці, оскільки метод Монте-Карло враховує тільки частоту потрапляння випадкових точок в інтервал, а вона залежить тільки від величини інтервалу й не залежить від його місця розташування.

Події називаються *несумісними*, якщо ймовірності появи цих подій одночасно дорівнюють 0. Звідси випливає, що сумарна ймовірність групи несумісних подій дорівнює 1.

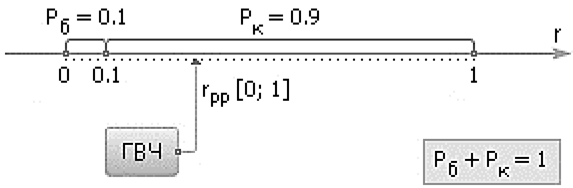


Рис. 5.13. Схема використання генератора випадкових чисел для імітації випадкової події

Позначимо через $a_1, a_2, \dots, a_n$ події, а через $P_1, P_2, \dots, P_n$ – ймовірності появи окремих подій.

Оскільки події не сумісні, то сума ймовірностей їх випадання дорівнює 1:

$$P_1 + P_2 + \cdots + P_n = 1.$$

Знову використаємо для імітації випадання однієї з подій генератор випадкових чисел, значення яких також завжди перебуває в діапазоні від 0 до 1. Відкладемо на одиничному інтервалі [0; 1] відрізки $P_1, P_2, \dots, P_n$. Зрозуміло, що в сумі відрізки становлять точно одиничний інтервал. Точка, яка відповідає числу, що випало з ГВЧ на цьому інтервалі, вкаже на один із відрізків. Відповідно до великих відрізків випадкові числа потраплять частіше (імовірність появи цих подій більша), до менших відрізків – рідше (рис. 5.14).

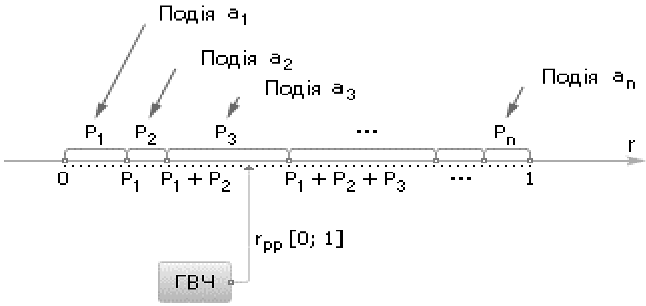


Рис. 5.14. Схема генерації несумісних випадкових подій за допомогою генератора випадкових чисел

Алгоритм роботи визначається за допомогою фільтра, побудованого у вигляді послідовності умовних операцій (IF), у який з інтервалів – від 0 до $P_1$, від $P_1$ до $(P_1 + P_2)$, від $(P_1 + P_2)$ до $(P_1 + P_2 + P_3)$ і т. д. потрапило число, згенероване генератором випадкових чисел. Якщо число потрапило в якийсь із інтервалів (що відбувається завжди й обов'язково), то це відповідає появі пов'язаної з ним події.

## 5.6. Моделювання випадкової величини із заданим законом розподілу

Більшу інформативність порівняно з такими статистичними характеристиками, як математичне сподівання, дисперсія, для інженера має закон розподілу ймовірності випадкової величини $X$. Уявимо, що $X$ набуває випадкових значень з деякого діапазону. Наприклад, $X$ – діаметр деталі, що виточується. Діаметр може відхилятися від запланованого ідеального значення під впливом різних факторів, які не можна врахувати, тому він є випадковою слабко передбачуваною величиною. Але в результаті тривалого спостереження за деталями, що випускаються, можна відзначити, скільки деталей з 1000 мали *діаметр* $X_1$ (*позначимо* $N_{X_1}$), скільки деталей мали *діаметр* $X_2$ (*позначимо* $N_{X_2}$) і т. д. У підсумку можна побудувати гістограму частості діаметрів, відкладаючи для $X_1$ величину $N_{X_1}/1000$, для $X_2$ – величину $N_{X_2}/1000$ і т. д. (Слід звернути увагу на те, що якщо бути точним, $N_{X_1}$– це кількість деталей, діаметр яких не просто дорівнює $X_1$, а перебуває в діапазоні від $X_1 - \Delta/2$ до $X_1 + \Delta/2$, де $\Delta = X_1 - X_2$). Важливо, що сума всіх частостей буде дорівнювати 1 (сумарна площа гістограми не змінна). Якщо $X$ змінюється безперервно, дослідів проведено дуже багато, то в межах $N \to \infty$ гістограма перетворюється в графік розподілу ймовірності випадкової величини. На рис. 5.16, *а* показаний приклад гістограми дискретного розподілу, а на рис. 5.15, *б* – варіант безперервного розподілу випадкової величини.

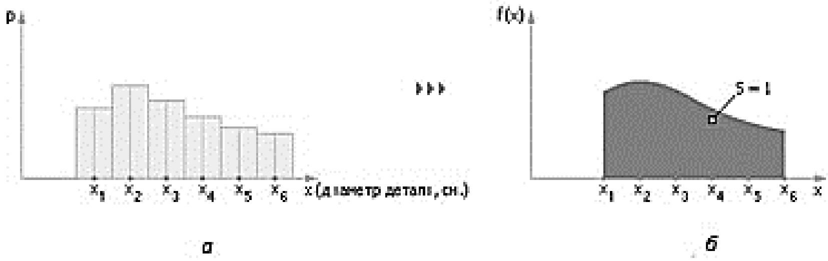

Рис. 5.15. Порівняння дискретного й безперервного законів розподілу випадкової величини

У нашому прикладі закон розподілу ймовірності випадкової величини показує наскільки ймовірне те або інше значення діаметра деталей, що випускаються. Випадковою величиною є діаметр деталі.

У виробництві й техніці часто такі закони розподілу задані за умовою задачі. Наша задача наразі полягає в тому, щоб навчитися імітувати появу конкретних випадкових подій згідно з імовірностями такого розподілу.

***Метод ступеневої апроксимації.*** Оскільки закони розподілу ймовірності подій можуть бути різної форми, а не тільки рівноймовірними, то необхідно вміти перетворювати рівномірний ГВЧ у генератор випадкових чисел із заданим довільним законом розподілу. На рис. 5.3 викладене вище відповідає двом першим блокам методу статистичного моделювання. Для цього безперервний закон розподілу ймовірності події дискретизуємо, перетворимо в дискретний.

Позначимо: $h_i$ – висота $i$-го стовпця; $f(x)$ – розподіл імовірності (показує наскільки ймовірною є певна подія $x$). Значення $h_i$ операцією нормування необхідно перевести в одиниці ймовірності появи значень $x$ з інтервалу

$$x_i < x \ \leq x_{i+1} \ P_i = h_i/(h_1 + h_2 + \cdots + h_i + \cdots + \ h_n).$$

Операція нормування забезпечує суму ймовірностей усіх $n$ подій, яка дорівнює 1:

$$\sum_{i=1}^{n} P_i = 1.$$

На рис. 5.16 графічно показані перехід від довільного безперервного закону розподілу до дискретного (рис. 5.16, а), відображення одержуваних імовірностей на інтервал $r_{pp}[0;1]$ і генерація випадкових подій із використанням еталонного рівномірно розподіленого ГВЧ (рис. 5.16, б).

Зазначимо, що всередині інтервалу $x_i < x \leq x_{i+1}$ значення $x$ тепер не є розрізненим, воно однакове. Метод узагальнення початкової постановки задачі, переходячи від безперервного закону розподілу до дискретного. Тому слідураховувати кількість розбивок $n$ з умов точності подання.

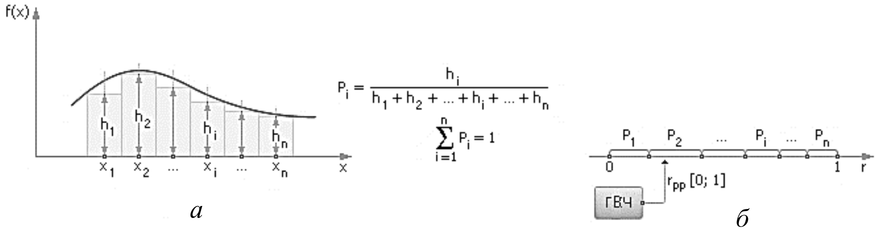


Рис. 5.16. Ілюстрація методу ступеневої апроксимації

Алгоритм, що реалізує описаний метод, генерує випадкове число, рівномірно розподілене від 0 до 1. Потім, порівнюючи довжину відрізків, розташованих на інтервалі від 0 до 1, що являють собою імовірності $P$ випадання тих або інших випадкових величин $X$, алгоритм визначає в циклі, яка з випадкових подій $i$ у результаті цього випадає.

***Метод зрізання.*** Метод використовується у випадку, коли функція задана аналітично (у вигляді формули). Графік функції вписують у прямокутник (рис. 5.17). На осі $X$ та $Y$ подають випадкове рівномірно розподілене число із ГВЧ. Якщо точка в перетині цих двох координат лежить нижче ніж крива щільності ймовірності, то подія $X$ відбулася, і навпаки – не відбулася.

Недоліком цього методу є те, що ті точки, які виявилися вищими за криву розподілу щільності ймовірності, відкидаються як непотрібні, і час, витрачений на їх обчислення, виявляється даремним. Метод застосований тільки для аналітичних функцій щільності ймовірності.

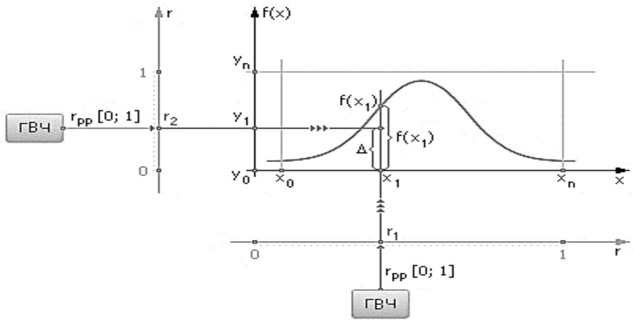


Рис. 5.17. Ілюстрація методу зрізання

Алгоритм, що реалізує метод зрізання, у циклі генерує два випадкові числа з діапазону від 0 до 1. Числа масштабуються у шкалу $X$ і $Y$ та перевіряється потрапляння точки зі згенерованими координатами під графік заданої функції $Y = f(X)$. Якщо точка перебуває під графіком функції, то подія $X$ відбулася з імовірністю $Y$, в іншому разі точка відкидається.

***Метод узяття оберненої функції.*** Припустимо, що нам заданий інтегральний закон розподілу ймовірності $F(x)$, де $f(x)$ – функція щільності ймовірності й

$$F(x) = \int_{-\infty}^{x} f(x)dx.$$

Тоді достатньо розіграти випадкове число, рівномірно розподілене в інтервалі від 0 до 1. Оскільки функція $F$ теж змінюється в цьому інтервалі, то випадкову подію $x$ можна визначити узяттям оберненої функції за графіком або аналітично: $x = F^{-1}(r)$, де $r$ – число, що генерується еталонним ГВЧ в інтервалі від 0 до 1, $x_1$ – генерована в підсумку випадкова величина. Графічно суть методу зображено на рис. 5.18, на якому показані графіки щільності ймовірності й інтегральної щільності ймовірності від $x$.

Цим методом особливо зручно користуватися у випадку, коли інтегральний закон розподілу ймовірності заданий аналітично й можливе аналітичне взяття оберненої функції від нього, як це й показано на такому прикладі.

Приймемо до розгляду експонентний закон розподілу ймовірності випадкових подій $f(x) = \lambda \cdot e^{-\lambda x}$. Тоді інтегральний закон розподілу щільності ймовірності має вигляд $F(x) = 1 - e^{-\lambda x}$.

Оскільки $r$ і $F$ у цьому методі передбачаються аналогічними й розташовані в одному інтервалі, то, заміняючи $F$ на випадкове число $r$, маємо: $r = 1 - e^{-\lambda x}$.

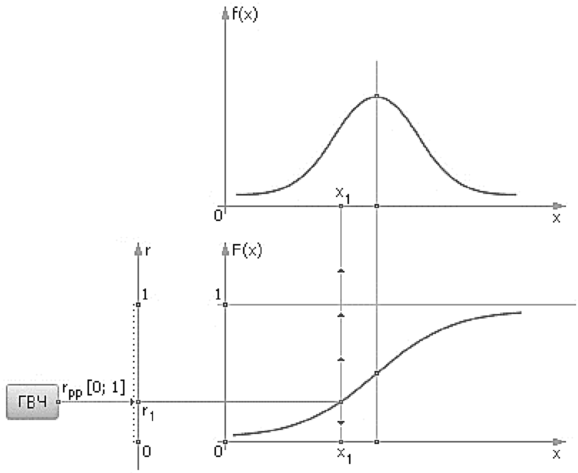


Рис. 5.18. Ілюстрація методу оберненої функції для генерації випадкових подій $x$, значення яких розподілені безперервно

Виражаючи шукану величину $x$ із цього виразу (тобто, перетворюючи функцію exp()), одержуємо $x = -1/\lambda \cdot \ln(1-r)$.

Оскільки в статичному сенсі $(1-r)$ і $r$ – це одне й те саме, то $x = -1/\lambda \cdot \ln(r)$.

## 5.7. Моделювання нормально розподілених випадкових величин

Нормальний закон розподілу зустрічається в природі досить часто, тому для нього розроблені окремі ефективні методи моделювання. Формула розподілу ймовірності значень випадкової величини $x$ за нормальним законом має вигляд

$$y = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}},$$

де $x$ – випадкова величина; $y(x)$ – імовірність прийняття випадковою величиною значення $x$; $m_x$ – математичне очікування; $\sigma_x$ – середнє квадратичне відхилення.

Бачимо, нормальний розподіл має два параметри: математичне очікування $m_x$ і середньоквадратичне відхилення $\sigma_x$ величини $x$ від цього математичного очікування.

*Нормалізованим нормальним розподілом* називається такий нормальний розподіл, у якого $m_x = 0$ і $\sigma_x = 1$. З нормалізованого розподілу можна одержати будь-який інший нормальний розподіл із заданими $m_x$ і $\sigma_x$ за формулою $z = m_x + x \cdot \sigma_x$.

Розглядаючи останню формулу, слід згадати формули комп'ютерної графіки: операція масштабування виражається в математичній моделі через множення (це відповідає зміні розкиду величини, розтягуванню геометричного образу), операція зсуву виражається через додавання (це відповідає зміні значення найбільш імовірної величини, зсуву геометричного образу).

Функція нормального розподілу має вигляд піка. На рис. 5.19 показаний нормалізований нормальний розподіл.

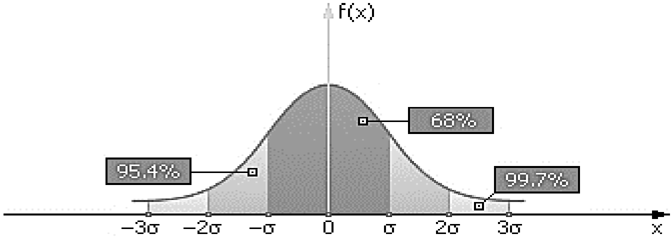


Рис. 5.19. Графічний вигляд нормального закону розподілу випадкової величини x із параметрами $m_x = 0$ і $\sigma_x = 1$

Графік на рис. 5.19 показує, що в області $-\sigma < x < \sigma$ на графіку зосереджено 68 % площі розподілу, в області $-2\sigma < x < 2\sigma$ на графіку зосереджено 95,4 % площі розподілу, в області $-3\sigma < x < 3\sigma$ на графіку зосереджено 99,7 % площі розподілу («правило трьох сигм»).

***Властивості нормального розподілу.*** Зміна параметра нормального розподілу $m_x$ призводить до зсуву кривої по осі $x$ (рис. 5.20).

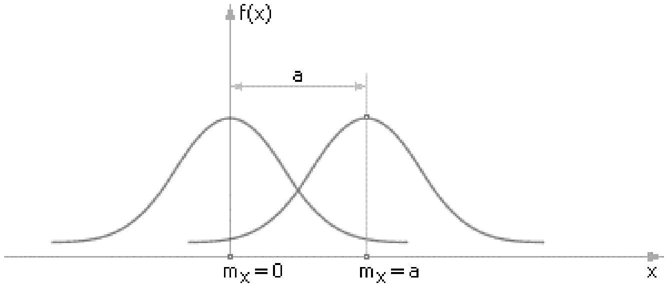


Рис. 5.20. Вплив параметра «математичне очікування» з погляду закону нормального розподілу

Зміна параметра нормального розподілу $\sigma_x$ приводить до масштабування форми по осі $x$ (нагадуємо, у будь-якому випадку завжди площа під кривою щільності ймовірності не змінна й дорівнює 1).

Чим більш не випадковим є процес, тим менше його середньоквадратичне відхилення і тим вищий пік на графіку. Зміна параметра нормального розподілу $\sigma_x$ приводить до масштабування форми (рис. 5.21) по осі $x$ (нагадуємо, у будь-якому випадку завжди площа під кривою щільності ймовірності не змінна й дорівнює 1).

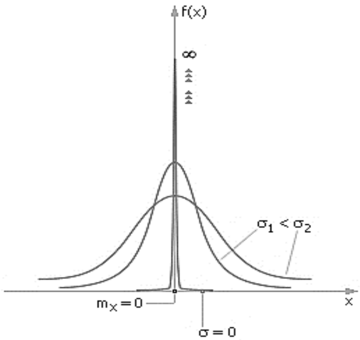


Рис. 5.21. Вплив параметра «середньоквадратичне відхилення» з погляду закону нормального розподілу

Дійсно, розкид випадковості відносно математичного очікування стає дедалі більш мінімальним. У межах детермінований процес має вигляд, показаний на рис. 5.22.

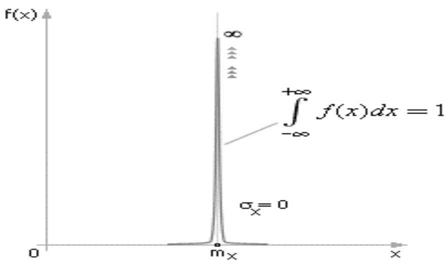


Рис. 5.22. Вигляд закону нормального розподілу ймовірності при переході його до детермінованого випадку

Випадкова подія стає детермінованою: $x = m_x \pm 0$ (розкиду немає).

Вивчати детерміновані процеси простіше. Чим більша величина $\sigma_x$, тим менш закономірною є поведінка досліджуваного об'єкта, оскільки можливі будь-які значення параметрів, що його характеризують, розкид величин відносно середньої очікуваної збільшується. Прогнозування й керування поведінкою об'єкта в цьому випадку ускладнюється.

Розглянемо вид інтегральної кривої щільності розподілу випадкової величини, розподіленої за нормальним законом. Вигляд її наведений на рис. 5.23, де $F$ – інтегральна функція Лапласа. Зміст інтегральної функції – імовірність того, що випадкова величина набуде значення з діапазону від $-\infty$ до $x$. Наприклад, запис $F(170) = 0{,}5$ для нашого прикладу означає: імовірність того, що випадково вибрана з аудиторії людина буде з ростом не вище 170 см, становить 0,5 (тобто кожна друга).

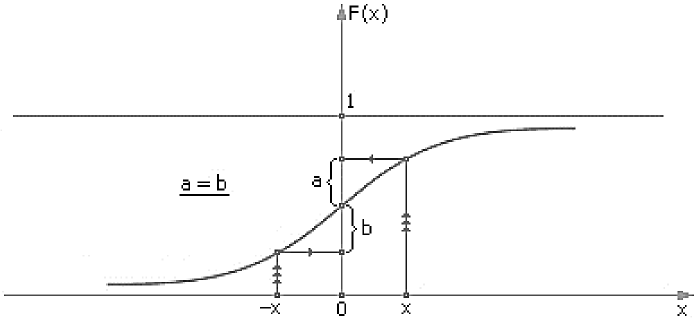


Рис. 5.23. Вигляд інтегральної функції Лапласа F(x)

Ця функція задана інтегралом від щільності ймовірності нормального розподілу

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

На жаль, цей інтеграл не береться в загальному вигляді, тому функція Лапласа задана у вигляді таблиці для $m_x = 0$ і $\sigma_x = 1$. Оскільки функція Лапласа симетрична відносно точки $(x = 0, y = 0{,}5)$ (як і функція самого нормального розподілу), $F(-x) = 1 - F(x)$, то в таблиці міститься тільки одна з її симетричних частин.

Якщо задається інтервал інтегрування функції Лапласа $[a; b]$, то

$$P(a < x < b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} dx = \int_{\frac{a-m}{\sigma}}^{\frac{b-m}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = F(b) - F(a).$$

Імовірність потрапляння $X$ в інтервал симетрична відносно $m_x$:

$$P(|x- m_x| < L) = \left((m_x - L) < x < (m_x + L)\right) = F\left(\frac{b - m_x}{\sigma}\right) - F\left(\frac{a - m_x}{\sigma}\right) =$$

$$= F\left(\frac{L}{\sigma}\right) - F\left(\frac{-L}{\sigma}\right) = 2F\left(\frac{L}{\sigma}\right) - 1.$$

Наприклад, для правила «трьох сигм»:

$$P(|x- m_x| < 3\sigma) = 2 \cdot F(3) - 1 = 2 \cdot 0{,}9987 - 1 = 0{,}9973$$

(як раніше ми і вказували). Число $F(3) = 0{,}9987$ узято з таблиці Лапласа.

**Приклад 3**. Знайти ймовірність виготовлення деталі з похибкою в її розмірах не більше 15 мм, якщо відомо, що виготовлення деталі з похибкою розподілене за нормальним законом $m = 0$ і $\sigma = 10$ мм.

$$P(|x| < 15) = F\left(\frac{(15-0)}{10}\right) - F\left(\frac{(-15-0)}{10}\right) = F(1{,}5) - F(-1{,}5) = F(1{,}5) -$$
$$-\left(1 - F(1{,}5)\right) = 2F(1{,}5) - 1 = 0{,}8664.$$

Тобто 8 664 деталей із 10 000 будуть мати похибку в розмірах не більше 15 мм.

Важливою властивістю закону є те, що нормальний розподіл є межею для різного виду розподілів, що випливає із центральної граничної теореми: «Для великої кількості $N$ випадкових величин $X$ із будь-яким законом розподілу їх сума є випадковим числом з нормальним законом розподілу»

$$\lim_{N\to\infty} P\left\{\alpha < \left(\sum_{i=1}^{n} x_i - N\alpha\right)\frac{1}{\sqrt{N\cdot\alpha}} < \beta\right\} = \frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-\frac{t^2}{2}} dt = F(\beta) - F(\alpha),$$

де $a$ – математичне сподівання в законі розподілу випадкової величини $X$; $\sigma$ – середньоквадратичне відхилення в законі розподілу випадкової величини $X$; $N$ – кількість випадкових чисел.

Розглянемо питання імітації випадкових величин, заданих нормальним законом розподілу.

**Табличний метод генерації нормально розподілених чисел**. Для цього нормальне число можна взяти з довідника в таблиці функції Лапласа й одержати випадкове число за методом взяття оберненої функції $x = F^{-1}(r)$, де $F$ – інтегральна функція Лапласа.

Технічно це означає, що потрібно розіграти випадкове рівномірно розподілене число $r$ з інтервалу [0; 1] стандартним ГВЧ, знайти число, яке дорівнює йому в таблиці значень функції Лапласа у стовпці $F$, за рядком визначити випадкову величину $x$, відповідну до цього числа.

Недоліком методу є необхідність зберігання в пам’яті комп’ютера всієї таблиці чисел функції Лапласа.

**Метод генерації нормально розподілених чисел, що використовує центральну граничну теорему (ЦГТ)**.Загальна суть методу полягає в такому: потрібно скласти випадкові числа з будь-яким законом розподілу, нормалізувати їх і перевести в потрібний діапазон нормального розподілу.

Припустимо, що нам потрібно з метою імітації одержати ряд випадкових чисел $x$, розподілених за нормальним законом із заданими математичним очікуванням $m_x$ і середньоквадратичним відхиленням $\sigma_x$.

Складемо $n$ випадкових чисел, використовуючи стандартний ГВЧ:

$$V = \sum_{i=1}^{n} r_i.$$

Згідно з ЦГТ числа $V$ утворюють ряд значень, розподілений за нормальним законом. Ці числа тим краще описують нормальний закон, чим більший параметр $n$. На практиці $n$ беруть такі, які дорівнюють 6 або 12. Зазначимо, що закон розподілу чисел $V$ має математичне очікування $m_v = n/2$, $\sigma_v = sqrt(n/12)$. Тому він є зміщеним відносно заданого довільного.

За допомогою формули $z = (V - m_v)/\sigma_v$ нормалізуємо цей ряд. Одержимо нормалізований закон нормального розподілу чисел $Z$. Тобто $m_z = 0, \sigma_z = 1$.

Формулою (зсув на $m_x$ і масштабування на $\sigma_x$) перетворимо ряд $Z$ у ряд $x$: $x = z \cdot \sigma_x + m_x$.

**Приклад 4.** Змоделювати потік заготовок для обробки їх на верстаті. Відомо, що довжина заготовки коливається випадково. Середня довжина заготовки становить 35 см, а середньоквадратичне відхилення реальної довжини від середньої – 10 см. Тобто за умовами задачі $m_x = 35, \sigma_x = 10$. Тоді значення випадкової величини буде розраховуватися за формулою:

$$V = r_1 + r_2 + r_3 + r_4 + r_5 + r_6,$$

де $r$ – випадкові числа із ГВЧ [0; 1], $n = 6$.

$$X = \sigma_x \left( sqrt\left(\frac{12}{n}\right)\left(V - \frac{n}{2}\right)\right) + m_x = 10 \cdot sqrt(2)(V - 3) + 35$$

або

$$X = 10 \cdot sqrt(2)\big((r_1 + r_2 + r_3 + r_4 + r_5 + r_6) - 3\big) + 35.$$

***Метод Мюллера.*** Зовсім простим методом одержання нормальних чисел є метод Мюллера, що використовує формули $Z = \sqrt{\big(-2 \cdot Ln(r_1)\big) \cdot \cos(2\pi \cdot r_2)}$ де $r_1$ і $r_2$ – випадкові числа із ГВЧ$_{рр}$ $[0;\ 1]$.

Можна також скористатися аналогічною формулою $Z = \sqrt{\big(-2 \cdot \mathrm{Ln}(r_1)\big) \cdot \sin(2\pi \cdot r_2)}$, де $r_1$ і $r_2$ – випадкові числа із ГВЧ$_{рр}$ $[0;\ 1]$.

## 5.8. Моделювання системи випадкових величин

Часто на практиці зустрічаються системи випадкових величин, тобто такі дві (і більше) різні випадкові величини $X,\ Y$ (та інші), які залежать одна від одної. Наприклад, якщо відбулася подія $X$ і набула якоюсь випадкового значення, то подія $Y$ відбувається хоча й випадково, але з урахуванням того, що $X$ уже набуло якогось значення.

Наприклад, якщо як $X$ випало велике число, то $Y$ повинно випасти теж досить великим числом (якщо кореляція додатна). Достатньо імовірно, що, якщо людина має велику вагу, то вона, швидше за все, буде й велика на зріст. Хоча це НЕ ОБОВ’ЯЗКОВО, тому що це НЕ ЗАКОНОМІРНІСТЬ, а кореляція випадкових величин. Оскільки бувають, хоча й рідко, люди з великою вагою, але невеликі на зріст або з маленькою вагою й високі. І все-таки основна маса гладких людей – високі, а низькі люди мають малу вагу.

За визначенням, якщо випадкові величини незалежні, то

$$f(\vec{x}) = f(x_1) \cdot f(x_2) \cdot \ \dots \ \cdot f(x_n),$$

де $x_i$ – випадкова незалежна величина; $f(x_i)$ – щільність імовірності випадання випадкової незалежної величини $x_i$; $f(\vec{x})$ – щільність імовірності випадання вектора $\vec{x}$ випадкових незалежних величин $x_1, x_2, \dots, x_n$.

Якщо випадкові величини залежні, то

$$f(\vec{x}) = f(x_1) \cdot f(x_2|x_1) \dots f(x_n) \cdot f(x_{n-1}|x_{n-2}, \dots, x_2, x_1),$$

де $\left(x_j \middle| x_{j-1}, \dots, , x_1\right)$ – випадкові залежні величини: випадання $x_j$ за умови, що випали $x_{j-1}, \dots, , x_1$; $f\left(x_j \middle| x_{j-1}, \dots, , x_1\right)$ – щільність умовної ймовірності появи $x_j$, якщо випали $x_{j-1}, \dots, , x_1$; $f(\vec{x})$ – імовірність випадання вектора $\vec{x}$ випадкових залежних величин.

Нехай, наприклад, є дві залежні події – $X$ і $Y$, розподілені за нормальним законом. Подія $X$ має математичне очікування $m_x$ і середньоквадратичне відхилення $\sigma_x$, а $Y$ має математичне очікування $m_y$ і середньоквадратичне відхилення $\sigma_y$. Коефіцієнт кореляції $q$ показує, наскільки тісно пов'язані події $X$ і $Y$. Якщо коефіцієнт кореляції дорівнює одиниці, то залежність подій $X$ і $Y$ взаємно є однозначною: одному значенню $X$ відповідає одне значення $Y$ (рис. 5.24).

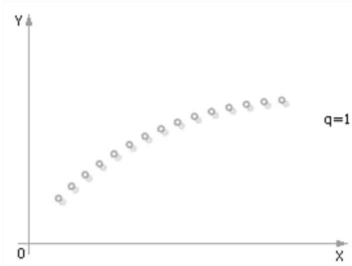


Рис. 5.24. Вигляд залежності двох випадкових величин при додатному коефіцієнті кореляції (q = 1)

При $q$, близьких до одиниці, виникає картина, показана на рис. 5.25, тобто одному значенню $X$ можуть відповідати вже декілька значень $Y$ (точніше, одне з декількох значень $Y$ зумовлене випадково); у цьому випадку події $X$ і $Y$ менш корельовані, менш залежні одна від одної.

І, нарешті, коли коефіцієнт кореляції наближається до нуля, виникає ситуація, при якій будь-якому значенню $X$ може відповідати будь-яке значення $Y$, тобто події $X$ і $Y$ не залежать або майже не залежать одна від одного, не корелюють одна з одною.

На всіх графіках кореляція була прийнята як додатна величина.

Насправді випадкові події ($X$ і $Y$) не можуть набувати з рівною ймовірністю будь-яких значень, як це має місце на рис. 5.24. Наприклад, у групі студентів не може бути людей надмалого або надвеликого зросту; переважно, люди мають якийсь середній зріст і розкид навколо цього середнього зросту. Тому на одних ділянках осі $X$ кількість подій розташована густіше, на інших – рідше. (Щільність випадкових подій, кількість точок на графіках більші біля величин $m_x$). Те саме правильно і для $Y$.

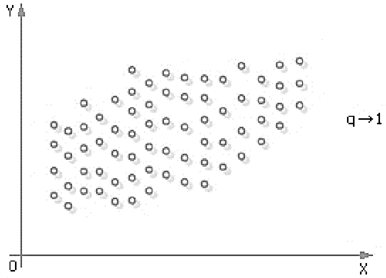


Рис. 5.25. Вигляд залежності двох випадкових величин при додатному коефіцієнті кореляції ($0 < q < 1$)

Для прикладу як найпоширеніший виберемо нормальний розподіл. Математичне очікування вказує на самі ймовірні події, де кількість подій більша й графік подій густіший. Додатна кореляція вказує на те, що великі випадкові величини $X$ стимулюють до генерації великі $Y$. Негативна кореляція вказує на те, що більші випадкові величини $X$ стимулюють до генерації менші випадкові величини $Y$. Нульова й близька до нуля кореляція показує, що величина випадкової величини $X$ ніяк не пов'язана з певним значенням випадкової величини $Y$. Легко зрозуміти сказане, якщо уявити собі спочатку розподіли $f(X)$ і $f(Y)$ окремо, а потім зв'язати їх у систему.

Припустимо, що $X$ і $Y$ розподілені за нормальним законом з відповідними значеннями $m_x$, $\sigma_x$ і $m_y$, $\sigma_y$. Заданий коефіцієнт кореляції двох випадкових подій $q$, тобто випадкові величини $X$ і $Y$ залежні одна від одної, $Y$ не зовсім випадковий.

Тоді можливий алгоритм реалізації моделі буде таким.

1. Розігрується шість випадкових рівномірно розподілених на інтервалі [0; 1] чисел $b_1, b_2, b_3, b_4, b_5, b_6$; знаходиться їх сума $S$: $S = b_1 + b_2 + b_3 + b_4 + b_5 + b_6$. Знаходиться нормально розподілене випадкове число $x$ за такою формулою: $x = sqrt(2)\sigma_x(S - 3) + m_x$.

2. За формулою $m_{y/x} = m_y + q \cdot \frac{\sigma_y}{\sigma_x} \cdot (x - m_x)$ знаходиться математичне очікування $m_{y/x}$ (знак $y/x$ означає, що набуватиме випадкових значень з урахуванням умови, що $x$ уже набув якихось певних значень).

3. За формулою $\sigma_{y/x} = \sigma_y \cdot sqrt(1- q^2)$ знаходиться середньоквадратичне відхилення $\sigma_{y/x}$ (знак $y/x$ означає, що $y$ набуватиме випадкових значеннь з урахуванням умови, що $x$ уже набув якихось певних значень).

4. Розігрується шість випадкових рівномірно розподілених на інтервалі [0; 1] чисел $r_1, r_2, r_3, r_4, r_5, r_6$; знаходиться їх сума $k$: $k = r_1 + r_2 + r_3 + r_4 + r_5 + r_6$. Знаходиться нормально розподілене випадкове число $y$ за такою формулою: $y = sqrt(2)\sigma_{y/x}(k- 3) + m_{y/x}$.

## 5.9. Розподіл Пуассона

Найбільш загальним випадком різного роду імовірнісних розподілів є біноміальний розподіл. Скористаємося його універсальністю для визначення частинних видів, які найчастіше зустрічаються на практиці розподілів.

**Біноміальний розподіл.** Нехай є якась подія $A$. Імовірність появи події $A$ дорівнює $P$, імовірність непояви події $A$ дорівнює $(1- P)$, іноді її непояву позначають як $Q = 1 - P$. Нехай $n$ – кількість випробувань, $m$ – частота появи події $A$ у цих $n$ випробуваннях.

Відомо, що сумарна ймовірність усіх можливих комбінацій результатів дорівнює одиниці, тобто

$$1 = P^n + nP^{n-1}(1 - P)^{n-1} + C_n^{n-2}P^{n-2}(1 - P)^{n-2} + \cdots +$$
$$+C_n^m P^m(1 - P)^{n-m} + \cdots + +(1 - P)^n(1–P),$$

де $P^n$ – імовірність того, що в $n$ випробуваннях подія $A$ відбудеться $n$ разів; $n \cdot P^{n-1}(1 - P)$ – імовірність того, що в $n$ випробуваннях подія $A$ відбудеться $(n- 1)$ разів і не відбудеться 1 раз; $C_n^{n-2} \cdot P^{n-2}(1- P)^2$ – імовірність того, що в $n$ випробуваннях подія $A$ відбудеться $(n- 2)$ разів й не відбудеться 2 рази; $P_m = C_n^m \cdot P^m(1- P)^{n-m}$ – імовірність того, що в $n$ випробуваннях подія $A$ відбудеться $m$ разів і не відбудеться $(n- m)$ разів; $(1- p)^n$ – імовірність того, що в $n$ випробуваннях подія $A$ не відбудеться жодного разу;
$C_n^m = \frac{n!}{m! \cdot (n-m)!}$ – кількість комбінацій з $n$ до $m$.

Математичне очікування $M$ біноміального розподілу дорівнює:

$$M = n \cdot P,$$

де $n$ – кількість випробувань; $P$ – імовірність появи події $A$.

Середньоквадратичне відхилення $\sigma$:

$$\sigma = sqrt(n \cdot P(1 - P)).$$

**Приклад 5**. Обчислимо ймовірність того, що подія, яка має ймовірність $P = 0{,}5$, у $n = 10$ випробуваннях відбудеться $m = 1$ раз. Маємо: $C_{10}^{1} = 10$ , і далі $P_1 = 10 \cdot 0{,}5^1 \cdot (1 - 0{,}5)^{10-1} = 0{,}0098$. Як бачимо, імовірність настання цієї події досить мала. Пояснюється це, по-перше, тим, що абсолютно не зрозуміло, чи відбудеться подія чи ні, оскільки ймовірність дорівнює 0,5 і шанси «50 на 50»; а по-друге, потрібно обчислити те, що подія відбудеться саме один раз (не більше й не менше) з десяти.

**Приклад 6.** Обчислимо ймовірність того, що подія, яка має ймовірність $P = 0{,}5$, у $n = 10$ випробуваннях відбудеться $m = 2$ рази. Маємо: $C_{10}^{2} = 45$, і далі: $P_2 = 45 \cdot 0{,}5^2 \cdot (1 - 0{,}5)^{10-2} = 0{,}044$. Імовірність настання цієї події стала більше!

**Приклад 7**. Збільшимо ймовірність настання самої події. Зробимо її більш імовірною. Обчислимо ймовірність того, що подія, яка має ймовірність $P = 0{,}8$, у $n = 10$ випробуваннях відбудеться $m = 1$ раз. Маємо: $C_{10}^{1} = 10$, і далі $P_1 = 0{,}000004$. Імовірність стала меншою, ніж у першому прикладі! Відповідь, на перший погляд, видається дивною, але оскільки подія має досить велику ймовірність, навряд чи вона відбудеться тільки один раз. Більш імовірно, що вона відбудеться більше, ніж один раз. Дійсно, підраховуючи $P_0, P_1, P_2, \dots, P_{10}$ (імовірність того, що подія в $n = 10$ випробуваннях відбудеться 0, 1, 2, 3, …, 10 разів), ми побачимо $P_0 = 0{,}0000\dots$; $P_1 = 0{,}0000\dots$; $P_2 = 0{,}0000\dots$; $P_3 = 0{,}0008\dots$; $P_4 = 0{,}0055\dots$; $P_5 = 0{,}0264\dots$; $P_6 = 0{,}0881\dots$; $P_7 = 0{,}2013\dots$; $P_8 = 0{,}3020\dots$ (найбільша ймовірність!); $P_9 = 0{,}2684\dots$; $P_{10} = 0{,}1074\dots$

Зрозуміло, що $P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 + P_8 + P_9 + P_{10} = 1$.

***Нормальний розподіл.*** Якщо зобразити величини $P_0, P_1, P_2, \dots, P_{10}$, які ми підрахували в прикладі 3, на графіку, то виявиться, що їх розподіл має вигляд, близький до нормального закону розподілу (рис. 5.26).

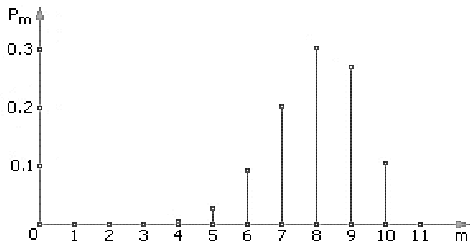

Рис. 5.26. Вигляд біноміального розподілу ймовірностей для різних $m$ при $P = 0{,}8; n = 10$

Біноміальний закон переходить у нормальний, якщо ймовірності появи й непояви події $A$ приблизно однакові, тобто умовно можна записати $P \approx 1 - P$. Для прикладу виберемо $n = 10$ і $P = 0{,}5$ (тобто $P = 1 - P = 0{,}5$).

Змістовно до такої задачі ми дійдемо, якщо, наприклад, захочемося теоретично порахувати, скільки буде хлопчиків і скільки дівчаток із 10 дітей народжених у пологовому будинку в один день. Точніше, рахувати будемо не хлопчиків і дівчаток, а ймовірність, що народяться тільки хлопчики, що народиться 1 хлопчик і 9 дівчаток, що народяться 2 хлопчика й 8 дівчаток і т. д. Візьмемо для простоти, що ймовірність народження хлопчика й дівчинки однакова та дорівнює 0,5 (але насправді, якщо чесно, це не так, див. курс «Моделювання систем штучного інтелекту»).

Зрозуміло, що розподіл буде симетричний, оскільки ймовірність народження 3 хлопчиків і 7 дівчаток дорівнює ймовірності народження 7 хлопчиків і 3 дівчаток. Найбільша ймовірність народження становитиме 5 хлопчиків і 5 дівчаток. Ця ймовірність дорівнює 0,25, до речі, не така вже вона велика за абсолютною величиною. Далі, імовірність того, що народяться відразу 10 або 9 хлопчиків набагато менша за імовірність того, що народиться $5 \pm 1$ хлопчик із 10 дітей. Саме біноміальний розподіл нам допоможе зробити цей розрахунок. Отже, $P_0 = 0{,}000977 \ldots$; $P_1 = 0{,}009766 \ldots$; $P_2 = 0{,}043945 \ldots$; $P_3 = 0{,}117188 \ldots$; $P_4 = 0.205078 \ldots$; $P_5 = 0{,}246094 \ldots$; $P_6 = 0{,}205078 \ldots$; $P_7 = 0{,}117188 \ldots$; $P_8 = 0{,}043945 \ldots$; $P_9 = 0{,}009766 \ldots$; $P_{10} = 0{,}000977 \ldots$

Зрозуміло, що $P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 + P_8 + P_9 + P_{10} = 1$.

Відобразимо на графіку величини $P_0, P_1, P_2, \ldots, P_{10}$ (рис. 5.27).

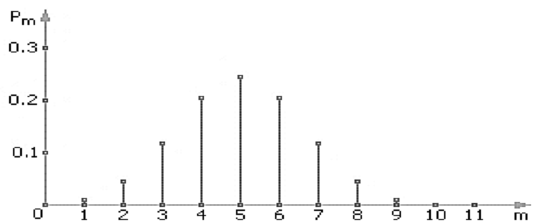


Рис. 5.27. Графік біноміального розподілу при параметрах $P = 0{,}5$ і $n = 10$, що наближають його до нормального закону

Отже, за умов $m \approx n/2$ і $P \approx 1-P$ або $P \approx 0{,}5$ замість біноміального розподілу можна використовувати нормальний. При більших значеннях $n$ графік зсувається праворуч, стає дедалі більш пологим, оскільки математичне очікування й дисперсія зростають зі збільшенням $n$: $M = n \cdot P$, $D = n \cdot P(1-P)$.

До речі, біноміальний закон наближається до нормального й при збільшенні $n$, що цілком природно відповідає до центральній граничній теоремі.

Тепер розглянемо, як зміниться біноміальний закон у випадку, коли $P \neq Q$, тобто $P \to 0$ . У цьому випадку застосувати гіпотезу про нормальність розподілу не можна, і біноміальний розподіл переходить у розподіл Пуассона.

***Розподіл Пуассона***. Розподіл Пуассона — це окремий випадок біноміального розподілу (при $n > 0$ і при $P \to 0$ (рідкі події)).

З математики відома формула, що дозволяє приблизно підрахувати значення будь-якого члена біноміального розподілу

$$P_m = C_n^m \cdot P^m (1-P)^{n-m} \cong \frac{\alpha^m \cdot e^{-\alpha}}{m!},$$

де $a = n \cdot P$ – параметр Пуассона (математичне очікування), а дисперсія дорівнює математичному очікуванню. Наведемо математичні викладки, що пояснюють цей перехід. Біноміальний закон розподілу

$$P_m = C_n^m \cdot P^m (1-P)^{n-m},$$

може бути написаний, якщо покласти $P = a/n$ у вигляді

$$P_m = C_n^m \left(\frac{a}{n}\right)^m \left(1-\frac{a}{n}\right)^{n-m};$$

або

$$P_m = \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{m-1}{n}\right)\left(1-\frac{\alpha}{n}\right)\frac{\alpha^m}{m!}.$$

Оскільки $P$ дуже мале, то слід брати до уваги тільки числа $m$, малі порівняно з $n$. Добуток

$$\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{m-1}{n}\right)$$

досить близький до одиниці. Це саме стосується величини $\left(1-\frac{\alpha}{n}\right)^{-m}$.

Величина

$$\left(1-\frac{\alpha}{n}\right)^{-n}$$

дуже близька до $e^{-a}$. Звідси одержуємо формулу:

$$P_m \cong \frac{\alpha^m \cdot e^{-\alpha}}{m!}.$$

**Приклад 8**. У ящику перебуває $n = 100$ деталей, як якісних, так і бракованих. Імовірність дістати бракований виріб становить $P = 0{,}01$. Допустимо, що ми виймаємо виріб, визначаємо, бракований він, чи ні, і кладемо його назад. Зробивши так, у нас вийшло, що зі 100 виробів, які ми перебрали, два виявилися бракованими. Яка ймовірність цього?

За біноміальним розподілом одержуємо

$$P_2 = C_n^m \cdot P^m (1-P)^{n-m} = C_{100}^2 \cdot 0{,}01^2 (1-0{,}01)^{100-2} \cong 0{,}185.$$

За розподілом Пуассона одержуємо

$$P_2 \cong \frac{\alpha^m \cdot e^{-\alpha}}{m!} = \frac{(n \cdot P)^m e^{-nP}}{m!} = \frac{(100 \cdot 0{,}1)^2 e^{-100 \cdot 0{,}01}}{2!} \cong 0{,}184.$$

Як бачимо, величини вийшли близькими, тому у випадку рідких подій цілком припустимо застосовувати закон Пуассона, тим більше, що він вимагає менших обчислювальних витрат.

Покажемо графічно вигляд закону Пуассона. Виберемо для прикладу параметри $P = 0{,}5$, $n = 10$. Тоді $P_0 = 0{,}5987 \ldots; P_1 = 0{,}3151 \ldots; P_2 = 0{,}0746 \ldots;$

$$P_3 = 0{,}0105 \ldots; P_4 = 0{,}00096 \ldots; P_5 = 0{,}00006 \ldots; P_6 = 0{,}0000 \ldots; P_7 = 0{,}0000 \ldots;$$

$$P_8 = 0{,}0000 \ldots; P_9 = 0{,}0000 \ldots; P_{10} = 0{,}0000 \ldots$$

Зрозуміло, що $P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 + P_7 + P_8 + P_9 + P_{10} = 1.$

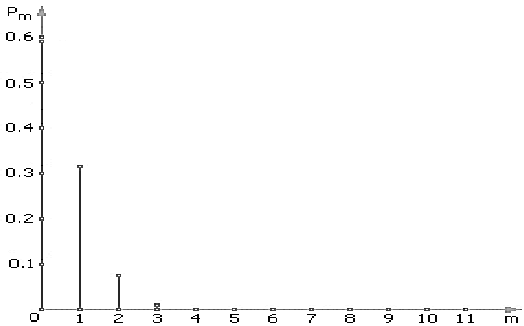


Рис. 5.28. Графік розподілу Пуассона при $P = 0{,}05$ і $n = 10$

При $n \to 0$ розподіл Пуассона (рис. 5.28) переходить у нормальний закон відповідно до центральної граничної теореми.

## 5.10. Потоки випадкових подій

Навчившись імітувати настання випадкових подій, ми можемо розіграти, яка з можливих подій настане й у якій кількості. Щоб це визначити, потрібно знати статистичні характеристики появи подій. Такою величиною, наприклад, може бути ймовірність появи події або розподіл імовірностей різних подій, якщо типів цих подій нескінченно багато.

Але часто ще важливо знати, коли конкретно настане та або інша подія в часі.

Коли подій багато, і вони відбуваються одна за одною, то утворюється ***потік***. Зазначимо, що події при цьому повинні бути однорідними, тобто схожими чимось одна на одну. Наприклад, поява водіїв на АЗС, які бажають заправити свій автомобіль. Тобто однорідні події утворюють певний ряд. При цьому вважається, що статистична характеристика цього явища (інтенсивність потоку подій) є заданою. Інтенсивність потоку подій указує, скільки *в середньому* відбувається таких подій за одиницю часу. Але коли саме відбудеться кожна конкретна подія, потрібно визначити методами моделювання. Важливо, що, коли ми згенеруємо, наприклад, за 200 годин 1000 подій, їх кількість буде дорівнювати приблизно величині середньої інтенсивності появи подій – 1000/200 = 5 подій за годину, що є статистичною величиною, яка характеризує цей потік загалом.

Інтенсивність потоку в деякому сенсі є математичним очікуванням кількості подій за одиницю часу. Але реально може так виявитися, що за одну годину відбудеться 4 події, за іншу – 6, хоча в середньому виходить 5 подій за годину, тому однієї величини для характеристики потоку недостатньо. Другою величиною, що характеризує, наскільки великий розкид подій відносно математичного очікування, є, як і раніше, дисперсія. Власне, саме ця величина визначає випадковість появи події, слабку передбачуваність моменту її появи.

Потік подій – це послідовність однорідних подій, які відбуваються одна за одною у випадкові проміжки часу. На осі часу ці події виглядають так, як показано на рис. 5.29.

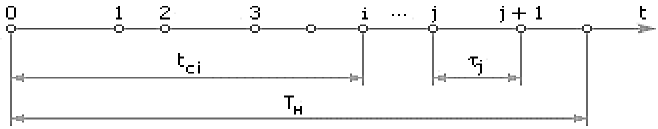


Рис. 5.29. Потік випадкових подій

На рис. 5.29: $\tau_j$ – інтервал між подіями (випадкова величина); $t_{zi}$ – момент здійснення $i$-тої події (відраховується від $t = 0$); $T_s$ – час спостереження.

Прикладом потоку подій можуть служити послідовність моментів торкання злітної смуги літаками, що прилітають в аеропорт.

**Інтенсивність потоку.** $\lambda$ – це середня кількість подій за одиницю часу. Інтенсивність потоку можна розрахувати експериментально за формулою: $\lambda = N/T_s$, де $N$ – кількість подій, що відбулися за час спостереження $T_s$.

Якщо інтервал між подіями $\tau_j$ дорівнює константі або визначений якою-небудь формулою у вигляді: $t_j = f(t_j - 1)$, то потік називається детермінованим. Інакше потік називається випадковим.

Випадкові потоки бувають:

– *ординарні:* імовірність одночасної появи двох і більше подій дорівнює нулю;

– *стаціонарні:* частота появи подій $\lambda(t) = const(t)$;

– *без післядії:* імовірність появи випадкової події не залежить від моменту здійснення попередніх подій.

**Пуассонівський потік.**Як еталон потоку в моделюванні прийнято брати пуассонівський потік.

***Пуассонівський потік*** – це ординарний потік без післядії. Як раніше було зазначено, імовірність того, що за інтервал часу $(t_0, t_0 + \tau)$ відбудеться $m$ подій, визначається із закону Пуассона:

$$P_m = \frac{\alpha^m e^{-\alpha}}{m!}, \ \alpha = \int_{t_0}^{t_0+\tau} \lambda(t)dt,$$

де α – параметр Пуассона.

Якщо $\lambda(t) = \mathrm{const}(t)$, то це **стаціонарний потік Пуассона** (найпростіший). У цьому випадку $a = \lambda t$. Якщо $\lambda = var(t)$, то це **нестаціонарний потік Пуассона**.

Для найпростішого потоку ймовірність появи $m$ подій за час $\tau$ дорівнює:

$$P_m = \frac{(\lambda \cdot \tau)^m e^{-\lambda\tau}}{m!}.$$

Імовірність непояви (тобто жодного, $m = 0$)події за час $\tau$ дорівнює:

$$P_0 = \frac{(\lambda \cdot \tau)^0 e^{-\lambda\tau}}{0!} = e^{-\lambda\tau}.$$

Рис. 5.30 ілюструє залежність $P_0$ від часу. Очевидно, що чим більший час спостереження, тим імовірність непояви жодної події менша. Крім того, чим більше значення λ, тим крутіше йде графік, тобто швидше убуває ймовірність. Це відповідає тому, що, якщо інтенсивність появи подій велика, то ймовірність непояви події швидко зменшується з часом спостереження.

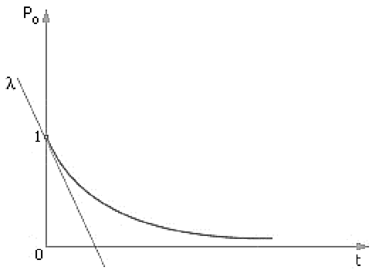


Рис. 5.30. Графік імовірності непояви жодної події в часі

Імовірність появи хоча б однієї події ($P_{\text{ХБ1П}}$) обчислюється так:

$$P_{\text{ХБ1П}} = 1 - P_0 = 1 - e^{-\lambda\tau}, \qquad (5.2)$$

оскільки $P_{\text{ХБ1П}} + P_0 = 1$ (або з'явиться хоча б одна подія, або не з'явиться жодної, – іншого не дано).

Із графіка на рис. 5.31 видно, що ймовірність появи хоча б однієї події наближається з часом до одиниці, тобто при відповідному тривалому спостереженні події таке обов'язково рано або пізно відбудеться. Чим довше ми спостерігаємо за подією (чим більше $t$), тим більша ймовірність того, що подія відбудеться, – графік функції монотонно зростає.

Чим більша інтенсивність появи події (чим більше λ), тим швидше настає ця подія, і тим швидше функція наближається до одиниці. На графіку параметр λ показаний крутістю лінії (нахил дотичної).

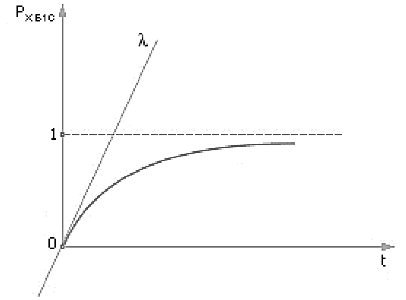


Рис. 5.31. Графік імовірності появи хоча б однієї події з часом

Якщо збільшувати λ, то при спостереженні за подією протягом того самого часу $\tau$ імовірність настання події зростає (рис. 5.32). Очевидно, що графік виходить із 0, оскільки, якщо час спостереження нескінченно малий, то ймовірність того, що подія відбудеться за цей час, не значна. І навпаки, якщо час спостереження нескінченно великий, то подія обов’язково відбудеться хоча б один раз, значить, графік прямує до значення ймовірності, яка дорівнює 1.

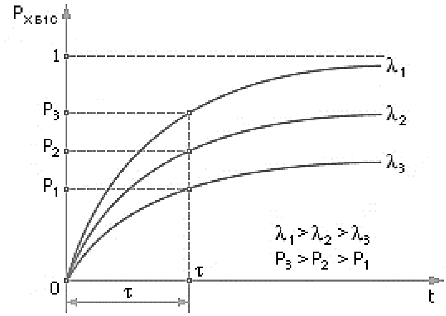


Рис. 5.32. Вплив величини інтенсивності потоку на ймовірність появи події протягом заданого інтервалу часу τ

Вивчаючи закон, можна визначити, що: $m_x = 1/\lambda,\ \sigma = 1/\lambda$, тобто для найпростішого потоку $m_x = \sigma$. Рівність математичного очікування середньоквадратичному відхиленню означає, що цей потік – потік без післядії. Дисперсія (точніше, середньоквадратичне відхилення) такого потоку є великою. Фізично це означає, що час появи події (відстань між подіями) погано передбачуваний, випадковий, перебуває в інтервалі $m_x - \sigma < \tau_j < m_x + \sigma$. Хоча зрозуміло, що в середньому він приблизно дорівнює: $\tau_j = m_x = T_s/N$. Подія може настати в будь-який момент часу, але в межах розкиду цього моменту $\tau_j$ відносно $m_x$ на $[-\sigma; +\sigma]$ (величину післядії). На рис. 5.33 показані можливі положення події 2 відносно осі часу при заданому $\sigma$. У цьому випадку говорять, що перша подія не впливає на другу, на третю і т. д., тобто післядія відсутня.

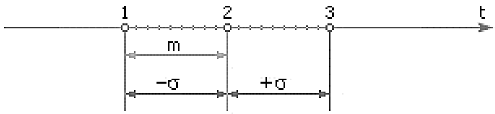


Рис. 5.33. Ілюстрація впливу величини $\sigma$ на положення події на часовій шкалі

За змістом $P$ дорівнює $r$, тому, виражаючи $\tau$ з формули (5.2), остаточно для визначення інтервалів між двома випадковими подіями маємо

$$\tau = \frac{1}{\lambda}\text{Ln}(r),$$

де $r$ – рівномірно розподілене від 0 до 1 випадкове число, яке беруть із ГВЧ, $\tau$ – інтервал між випадковими подіями (випадкова величина $\tau_j$).

**Приклад 9.** Розглянемо потік виробів, що приходять на технологічну операцію. Вироби приходять випадково – у середньому вісім штук за добу (інтенсивність потоку $\lambda = 8/24$ [од/год]). Необхідно промоделювати цей процес протягом $T_s = 100$ год. $m = 1/\lambda = 24/8 = 3$, тобто в середньому це одна деталь за три години. Зазначимо, що $\sigma = 3$. На рис. 5.34 наведений алгоритм, що генерує потік випадкових подій.

На рис. 5.355 показаний результат роботи алгоритму – моменти часу, коли деталі надходили на операцію. Як видно, усього за період $T_s = 100$ виробничий вузол обробив $N = 33$ вироби. Якщо запустити алгоритм знову, то $N$ може виявитися такою, що дорівнює наприклад, 34, 35 або 32. Але в середньому, за $K$ прогонів алгоритму $N$ буде дорівнювати 33,33… Якщо порахувати відстані між подіями $t_{si}$ і моментами часу, зумовленими як 3 $i$, то в середньому величина буде дорівнювати $\sigma = 3$.

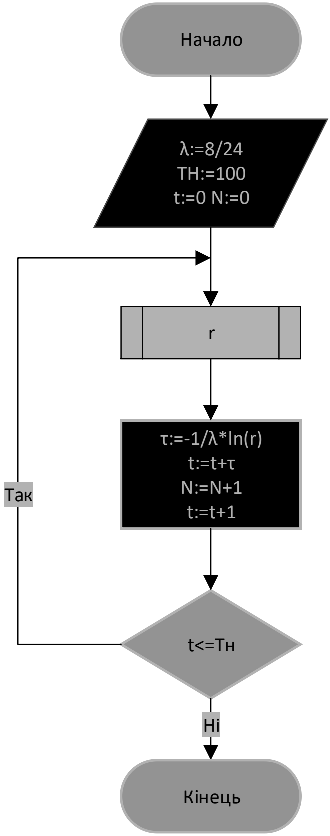


Рис. 5.34. Алгоритм, що генерує потік випадкових подій у заданому λ

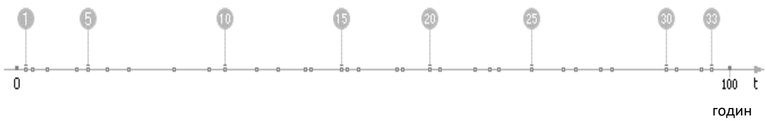


Рис. 5.35. Ілюстрація роботи алгоритму, що генерує потік випадкових подій

## 5.11. Моделювання неординарних і нестаціонарних потоків подій

Якщо відомо, що потік не є ординарним, то необхідно моделювати, крім моменту виникнення події, ще й кількість подій, яка могла з’явитися в цей момент. Наприклад, вагони на залізничну станцію прибувають у складі поїзда у випадкові моменти часу (ординарний потік поїздів). Але при цьому в складі поїзда може бути різна (випадкова) кількість вагонів. У цьому випадку про потік вагонів говорять як про потік неординарних подій.

Допустимо, що $M_k = 10, \sigma = 4$ (тобто у середньому в 68 випадках зі 100 приходить від 6 до 14 вагонів у складі поїзда) і їх кількість розподілена за нормальним законом.

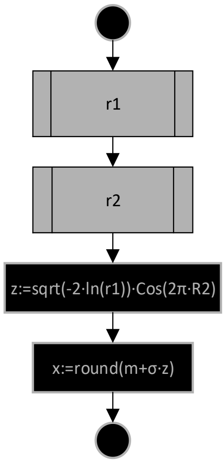


Рис. 5.36. Фрагмент алгоритму, що реалізує неординарний потік випадкових подій

**Приклад 10.** Дуже корисним у виробництві є розв'язання такої задачі. Який середній час добового простою устаткування технологічного вузла, якщо вузол обробляє кожний виріб у випадковий час, заданий інтенсивністю потоку випадкових подій $\lambda_2$? При цьому експериментально встановлено, що привозять вироби на обробку теж у випадкові моменти часу, задані потоком $\lambda_1$ партіями по 8 штук, причому розмір партії коливається випадково за нормальним законом з $m =$ $= 8, \sigma = 2$. До початку моделювання $T = 06$ на складі виробів не було. Необхідно промоделювати цей процес протягом $T_s = 100$ годин.

На рис. 5.37 наведений алгоритм, що генерує випадково потік надходження партій виробів на оброблення й потік випадкових подій – виходу партій виробів з оброблення.

На рис. 5.38 показаний результат роботи алгоритму – моменти часу, коли деталі надходили на операцію, і моменти часу, коли деталі залишали операцію. На третій лінії видно, скільки деталей стояло в черзі на обробку (лежало на складі вузла) у різні моменти часу.

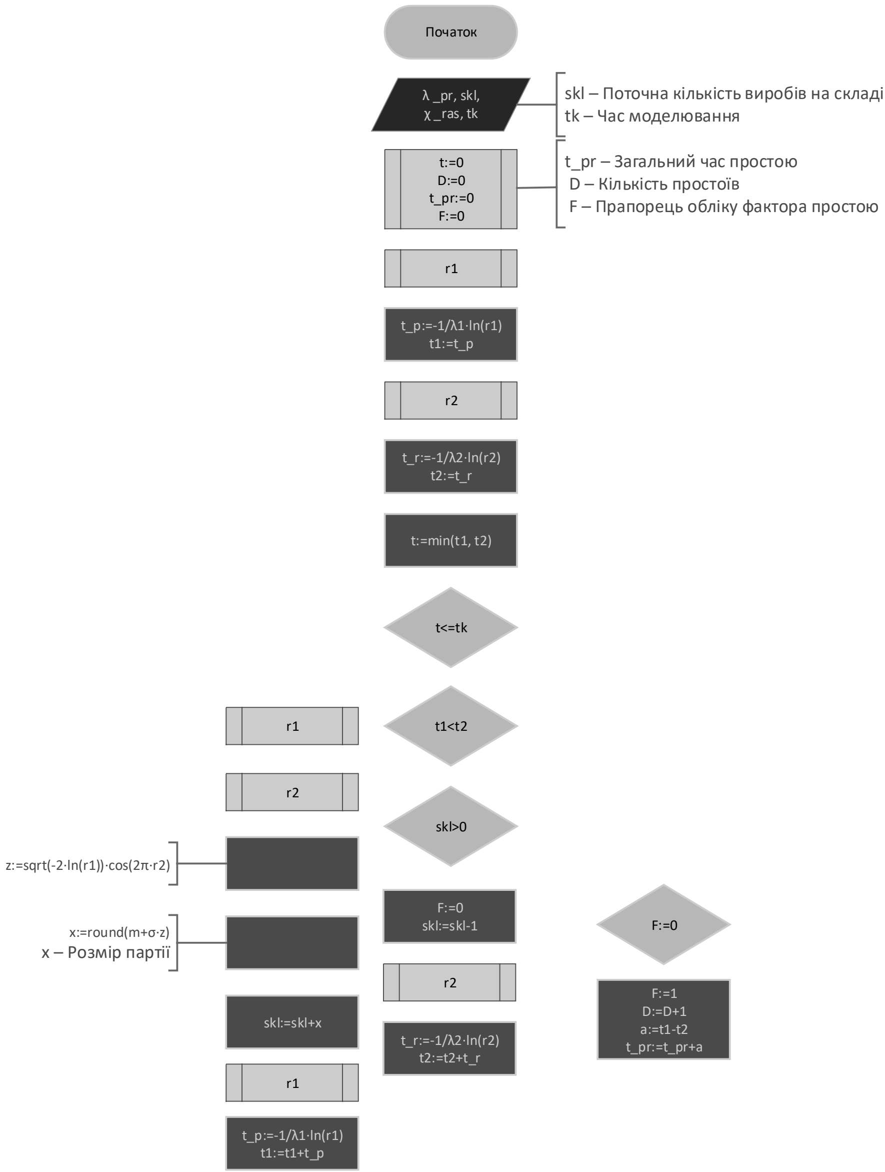


Рис. 5.37. Алгоритм імітації обробки партій виробів з урахуванням випадковості подій, що відбуваються

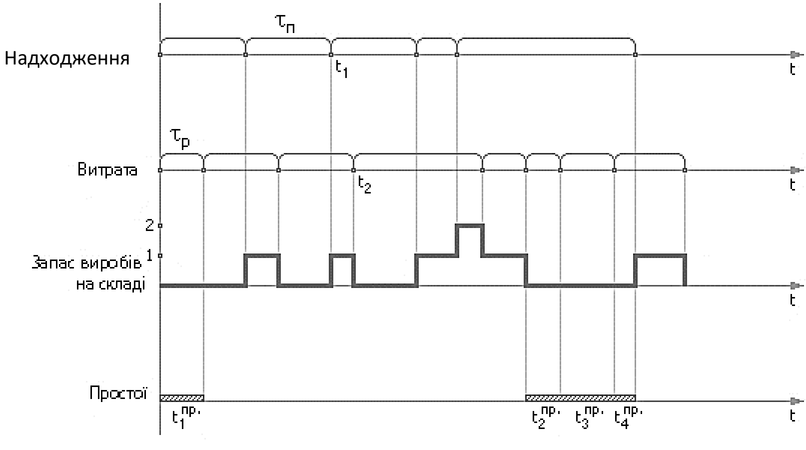


Рис. 5.38. Діаграма, що ілюструє рух виробів через вузол обробки

Відзначаючи для обробного вузла години, коли він простоював, чекаючи чергової деталі (на рис. 5.39 ділянки часу, виділені червоним штрихуванням), ми можемо порахувати сумарний час простоїв вузла за весь час спостереження, а потім розрахувати середній час простою протягом доби. Для цієї реалізації час обчислюється так:

$$T_{pr}^{cer} = 24\left(t_1^{in} + t_2^{in} + t_3^{in} + t_4^{in} + \cdots + t_N^{in}\right)/T_s.$$

**Завдання 1**. Змінюючи величину $\sigma$, встановити залежність $T_{pr}^{cer}(\sigma)$. Задаючи вартість за простій вузла 100 євро/година, установити річні втрати підприємства від нерегулярності в роботі постачальників. Запропонувати формулювання пункту договору підприємства з постачальниками «Величина штрафу за затримку поставки виробів».

**Завдання 2.** Змінюючи величину початкового заповнення складу, встановити, як зміняться річні втрати підприємства від нерегулярності в роботі постачальників залежно від прийнятої на підприємстві величини запасів.

У ряді випадків інтенсивність потоку може змінюватися з часом $\lambda(t)$. Такий потік називається нестаціонарним. Наприклад, середня кількість машин швидкої допомоги за годину, що залишають станцію за викликами населення великого міста, протягом доби може бути різною. Відомо, наприклад, що найбільша кількість викликів припадає на інтервали з 23 до 01 години ночі та з 05 до 07 ранку, тоді як в інші години вона вдвічі менша (рис. 5.39).

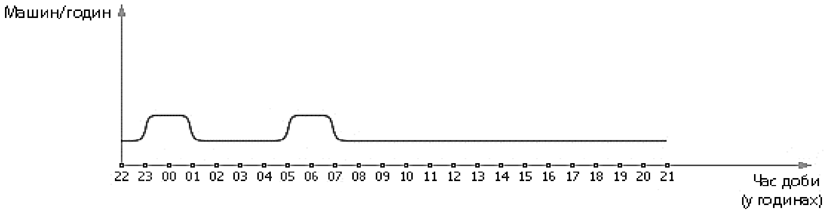


Рис. 5.39. Діаграма зміни інтенсивності потоку випадкових подій з часом

У цьому випадку розподіл $\lambda(t)$ може бути заданий або графіком, або формулою, або таблицею, або алгоритмі.

***Потоки з післядією (потоки Ерланга).*** Як ми раніше зазначили, інтенсивність потоку в деякому сенсі є математичним очікуванням кількості подій за одиницю часу (обернена до неї величина вказує на середній час між подіями). Другою величиною, яка характеризує наскільки великим є розкид у часі надходження подій відносно математичного очікування, є дисперсія.

Допустимо, події в потоці відбуваються точно одна за одною кожні 12 хвилин без відхилень. Інтенсивність такого потоку буде дорівнювати 5 подій за годину. Але, якщо події будуть відбуватися випадково, наприклад, 12 ± 2 хвилини, то в середньому вони дадуть також 5 подій за годину. Наприклад, за 200 годин відбудеться 1000 подій, величина інтенсивності – 1000/200 = 5 подій за годину. Тому за цією характеристикою потоки не можна відрізнити один від одного. Але очевидно, що другий потік все-таки буде більш випадковим, ніж перший, тим більше, якщо в потоці події будуть відбуватися одна за одною 12 ± 10 хвилин.

Перший потік ми назвемо детермінованим, регулярним, другий і третій — випадковими. Причому захід випадковості зі збільшенням дисперсії (розкиду величини інтервалу між подіями) буде зростати. У першому потоці дисперсія дорівнює нулю. Цей факт проілюстровано на рис. 5.40.

Потік Ерланга 1 порядку

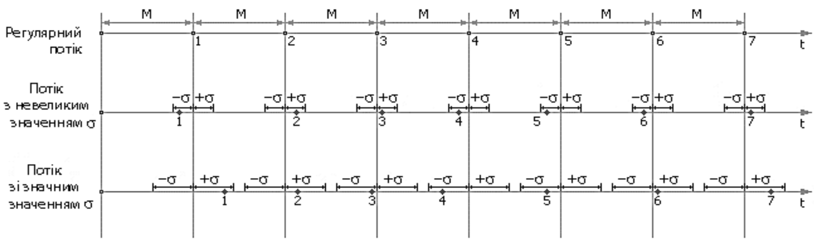


Рис. 5.40. Ілюстрація регулярного й випадкового потоків

Власне, саме дисперсія визначає випадковість появи події, слабку передбачуваність моменту її появи. Важливо вміти управляти цією величиною при моделюванні випадкових потоків. Якщо передбачити кожну наступну подію важко, то потік – без післядії (або з малою післядією, зв'язок між подіями відсутній, події випадкові), якщо потік детермінований, то післядія є великою – кожній події практично передує момент появи наступної.

Потік Ерланга $k$-го порядку – це потік випадкових подій, що виходить, якщо в найпростішому випадковому потоці зберегти кожну $k$-ту подію, а інші відкинути (рис. 5.41). Порядок потоку – захід післядії потоку. Тобто оберненою величиною до міри випадковості потоку є його порядок.

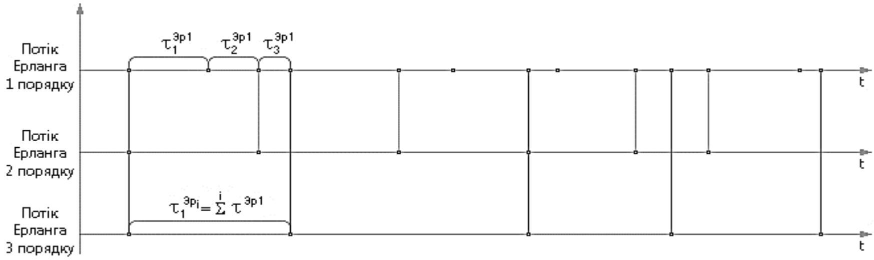


Рис. 5.41. Ілюстрація методу одержання потоків Ерланга

Просівання подій починає приводити до того, що між точками з'являється післядія, детермінація, яка тим вища, чим більше $k$. Зі збільшенням $k$ точки лягають на вісь часу дедалі більш рівномірно, розкид їх зменшується, регулярність збільшується.

Засновано це на тому простому й раніше вивченому нами факті, що сума випадкових величин є величиною невипадковою (центральною граничною теоремою). Чим більше ми складемо випадкових величин, тим передбачуваний буде результат (їх сума).

Очевидно, що

$$\tau^{Ep_k} = \sum_{i=1}^{k} \tau_i^{Ep_1},$$

інтервал між подіями в потоці Ерланга $k$-го порядку.

Щільність імовірності розподілу інтервалів між випадковими подіями в потоці Ерланга $k$-го порядку:

$$f(T_{ек}) = \frac{\lambda \cdot (\lambda \cdot t)^{k-1}}{(k-1)!} \cdot e^{-\lambda t},$$

$\lambda k = \lambda/k$ – інтенсивність потоку Ерланга $k$-го порядку, де λ – інтенсивність найпростішого потоку Пуассона, а $\lambda k$ інтенсивність просіяного $k$ разів потоку, тобто в $k$ разів менше.

Параметри закону Ерланга обчислюються за формулами

$$M_k = {}^{1}\!/_{\lambda_k}, \sigma_k = {}^{1}\!\Big/\frac{sqrt(k)}{\lambda_k}.$$

Слід звернути увагу, що в потоці Ерланга $M \neq \sigma$, тобто в потоках із післядією, рівність $M$ і $\sigma$ неможлива.

Більш того, при $k \to \infty$ подія відбувається строго у виміряний час, тому оскільки $\sigma \to 0$.

Порівняємо: Потік Ерланга 1-го порядку: $m = \sigma_1$ – потік без післядії; Потік Ерланга $i$-го порядку: $m \neq \sigma_2$, при цьому $(\sigma_2 > 0)$ і $(\sigma_1 > \sigma_2)$ розкид зменшується, післядія збільшується; Потік Ерланга ∞-го порядку: $m \neq \sigma = 0$ – регулярний потік.

Із цього випливає, що порядок потоку Ерланга – це захід післядії потоку.

**Приклад 11.** Розглянемо приклад виходу з ладу лампочок на опорі вуличного освітлення. Візьмемо час спостереження – 100 років. З паспортних даних на ці вироби відомо, що середній час роботи виробу на відмову становить 1,5 року; середньоквадратичне відхилення – 0,5 року.

Тобто задано: $M_k = 1{,}5$, $\sigma_k = 0{,}5$.

Оскільки $M_k \neq \sigma_k$, то $k \neq 1$, тобто ми маємо справу з потоком із післядією. Інтенсивність цього потоку $\lambda_k = 1/M_k = 1/1{,}5 = 0{,}67$. Обчислена інтенсивність потоку свідчить нам про те, що протягом року в середньому перегоряє 0,67 лампочки або 67 лампочок за 100 років.

Оскільки $\sigma_k = 1/\mathrm{sqrt}(k)/\lambda_k$ і дорівнює 0,5, то обчислимо порядок потоку Ерланга: $k = 1/\sigma_2/\lambda_k^2 = 1/0{,}52/0{,}672 \approx 9$.

Обчислимо інтенсивність порджувального потоку *Пуассона* $\lambda = \lambda_k \cdot k = 0{,}67 \cdot 9 = 6$.

## 5.12. Фіксація і обробка статистичних результатів

Вище ми докладно ознайомилися зі схемою статистичного комп'ютерного експерименту, а також розглянули практичну реалізацію всіх основних блоків (див. рис. 5.3) цієї схеми. Зараз важливо навчитися організовувати роботу останніх двох блоків – блока обчислення статистичних характеристик (БОСХ) і блоку оцінки достовірності статистичних результатів (БОД).

Отже, розглянемо, як слід фіксувати статистичні величини в результаті експерименту, щоб отримати надійну інформацію про властивості об'єкта, що моделюється. Нагадаємо, що узагальненими характеристиками випадкового процесу або явища є середні величини.

***Обчислення середніх.*** Обчислення середніх величин під час експерименту, який багаторазово повторюється, а результат його усереднюється, може бути організовано кількома способами:

- вся статистика обчислюється наприкінці обчислення;
- вся статистика обчислюється в процесі обчислення (за рекурсивними співвідношеннями);
- вся статистика обчислюється в класових інтервалах (цей метод поєднує універсальність першого методу і економічність другого).

*Спосіб 1. Обчислення всієї статистики в кінці.* Для цього в процесі експерименту значення $x_i$ вихідний (досліджуваної) випадкової величини $x$ накопичується в масиві даних. Після закінчення експерименту підраховується математичне очікування (середнє) $m_x$ і дисперсія $D_x$ (характерний розкид величин щодо цього математичного очікування)

$$m_x = \frac{1}{n}\sum_{i=1}^{n} x_i \, , \tag{5.3}$$

$$D_x = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)^2 \, . \tag{5.4}$$

Часто використовують середньоквадратичне відхилення $\sigma = sqrt(D)$.

Зауважимо, що недоліком методу є неефективне використання пам'яті, тоскільки доводиться накопичувати і зберігати велику кількість значень вихідної величини протягом усього експерименту, який може бути досить тривалим.

Другий мінус полягає в тому, що доводиться двічі зчитувати масив $x_i$, оскільки як скористатися формулою (5.4) в тому вигляді, як вона тут записана, ми можемо, тільки прорахувавши формулу (5.3) (від 1 до $n$), а потім ще раз прогнавши для формули (5.4) масив $x_i$.

Позитивним моментом є збереження всього масиву даних, що дає можливість більш детального його вивчення надалі при необхідності розслідування тих чи інших ефектів і результатів.

*Спосіб 2. Обчислення всієї статистики в процесі обчислення (за рекурсивним співвідношенням).* Цей спосіб передбачає можливість зберігати тільки поточне значення математичного очікування $x_i$ і дисперсії $D_i$, підправляти на кожній ітерації. Це позбавляє нас від необхідності постійного зберігання всього масиву експериментальних даних. Кожне нове дане $x_i$ враховується в сумі з ваговим коефіцієнтом – чим більше доданків $i$ накопичено в сумі $x_i$, тим більше її значення важливо по відношенню до чергової поправки $x_i$, тому співвідношення вагових коефіцієнтів $i/(i\ +\ 1)\ :\ 1/(i\ +\ 1)$.

$$x_{i+1} = \frac{x_i \cdot i + x_{i+1}}{i+1} = x_i \frac{i}{i+1} + x_{i+1} \frac{i}{i+1};$$

$$D_{i+1} = \frac{(x_{i+1} - x_i) + i \cdot D_i}{i+1} = \frac{(x_{i+1} - x_i)}{i+1} + \frac{i}{i+1} D_i,$$

де $x_i$ – чергове значення експериментальної вихідної величини.

*Спосіб 3. Обчислення всієї статистики в класових інтервалах*. Цей спосіб передбачає, що в масив будуть накопичувати в повному обсязі значення $x_i$, тільки за значущим інтервалами, в яких розподілена випадкова вихідна величина $x_i$. Загальний інтервал зміни $x_i$ розбивається на $m$ підінтервалів, в кожному з яких фіксується кількість $n_i$, що показує, скільки разів $x_i$ набуло значення з $i$-го інтервалу. При невеликій кількості інтервалів ($m \approx 1$) ми отримуємо спосіб 1, при кількості інтервалів $m = n$ ми отримуємо спосіб 2. У випадку $1 < m < n$ отримуємо середній розв'язок – компроміс між займаною пам'яттю й інформативністю масиву вихідних даних

$$\bar{x} = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} n_i x_i;\ \ D = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} (n_i x_i - \bar{x})^2.$$

**Обчислення геометрії розподілу.** Ще більш інформативним є обчислення геометрії розподілу випадкової величини. Воно необхідне для того, щоб уявити собі більш точно характер розподілу. Відомо, що за значенням статистичного моменту можна приблизно судити про геометричний вигляд розподілу.

*Перший момент (або середнє арифметичне)* обчислюється так:

$$m_1 = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} (x_i - A)^1;\ \text{де } A = \begin{cases} 0 \text{ (початковий момент)}, \\ \bar{x}, \text{(центральний момент)}. \end{cases}$$

Якщо $A$ набуває значення 0, то перший момент називається *початковим моментом*, якщо $A$ набуває значення $\bar{x}$, то перший момент називається *центральним*. (У принципі $A$ може бути будь-яким числом, що задається дослідником.)

На практиці прийнято використовувати не сам перший момент, а нормовану його величину $R_1 = m_1 / \sigma_1$.

Перший момент вказує на центр ваги в геометрії розподілу, рис. 5.42.

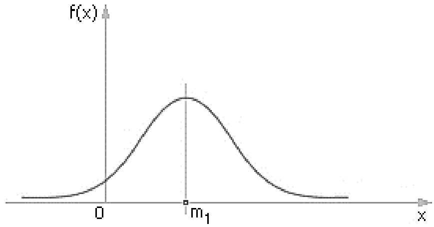


Рис. 5.42. Характерне положення першого моменту на графіку розподілу статистичної величини

**Другий момент (або дисперсія)** обчислюється так:

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

На практиці прийнято використовувати не сам другий момент, а нормовану його величину $R_2 = m_2 / \sigma_2$.

Дисперсія характеризує величину розкиду експериментальних даних щодо центру ваги $m_1$. Таким чином, за величиною $m_2$ можна судити про другий параметр геометрії розподілу (рис. 5.43).

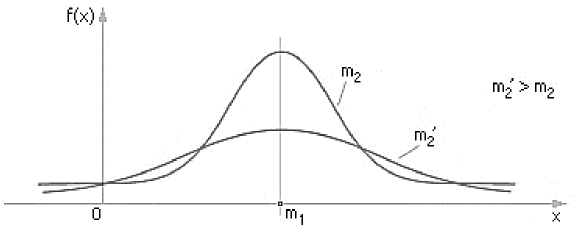


Рис. 5.43. Характерна зміна виду розподілу статистичної величини залежно від величини другого моменту

*Третій момент* характеризує асиметрію (або скісність) (рис. 5.44) обчислюється так:

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3.$$

На практиці прийнято використовувати не сам другий момент, а нормовану його величину $R_3 = m_3 / \sigma_3$.

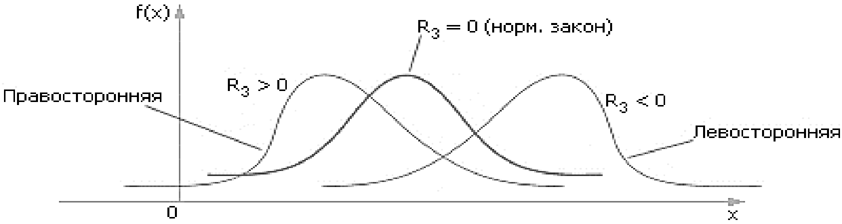


Рис. 5.44. Характерна зміна виду розподілу статистичної величини в залежності від величини третього моменту

Визначаючи знак $R_3$, можна з'ясувати, чи є асиметрія у розподілу (див. рис. 5.3), а якщо є ($R_3 \neq 0$), то в який бік.

*Четвертий момент* (рис. 5.45) характеризує ексцес (або гостровершинності) і обчислюється так:

$$m_4 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4.$$

Нормований момент дорівнює: $R_4 = m_4 - \sigma^4$.

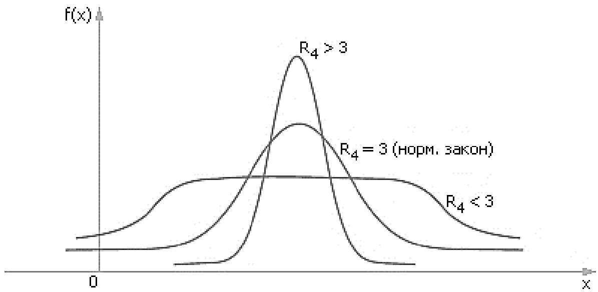


Рис. 5.45. Характерна зміна виду розподілу статистичної величини в залежності від величини четвертого моменту

Дуже важливим є з'ясування того, на який розподіл найбільше походить отриманий експериментально розподіл випадкової величини. Оцінка ступеня збігу емпіричного закону розподілу з теоретичним проводиться в два етапи:

– визначають параметри експериментального розподілу;

– роблять оцінку за Колмогоровим з умови відповідності експериментального розподілу вибраному теоретичному.

Оцінка (за Колмогоровим) збігу емпіричного закону розподілу з теоретичним полягає в такому:

1. Обчислюємо моменти $m_1,\ m_2, m_3, ...$ Кількість моментів дорівнює кількості невідомих у теоретичному законі розподілу.

2. Перш за все, оскільки оцінка стосується безперервного розподілу, а ми маємо справу з дискретним розподілом, знятим експериментально, то необхідно вирішити, на скільки інтервалів потрібно розбити при дискретизації і той, і інший розподіл. Для цього рекомендується користуватися правилом Стерджеса, добре зарекомендували себе на практиці: $K = 1 + \log_2 n = 1 + 3{,}322 \cdot \log_{10} n$, де $n$ – кількість випадкових значень (дослідів); $k$– кількість інтервалів розподілу.

3. Будується інтегральний закон (рис. 5.46) для емпіричного розподілу $F(x) = P(x \leq x_i)$.

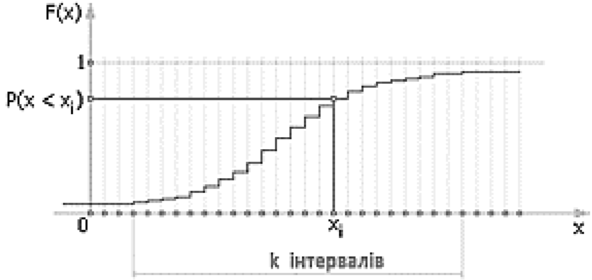


Рис. 5.46. Інтегральний закон емпіричного розподілу, дискретний варіант (приклад)

4. Залежності від кількості експериментів $n$ і кількості інтервалів $1 \leq i \leq k$ можна порахувати кількість випадків у кожному з інтервалів: $N_i = P_i \cdot n$.

5. Далі слід розрахувати теоретичний розподіл частоти: $N_i^{teor} = P_i \cdot n$. Якщо як теоретичний взяти нормальний закон розподілу, то можна зробити так:

$$P_i(x < x_i) = F\left(\frac{x-a}{\sigma}\right) + \frac{1}{2},$$

де $F$ – функція Лапласа, а параметри $a$ і $\sigma$ закону обчислені в п. 1.

6. Порівняємо отримані частоти: $N_i^{teor}$ і $N_i$ у всіх $k$ інтервалах (рис. 5.47) і виберемо найбільше відхилення експериментального розподілу від теоретичного:

$$D = \max_i |N_i - N_i^{teor}|.$$

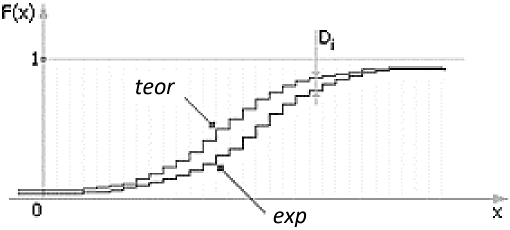


Рис. 5.47. Порівняння теоретичного та емпіричного інтегральних розподілів випадкової величини

7. Параметр Колмогорова λ характеризує відхилення теоретичного розподілу від експериментального:

$$\lambda = \left(\sqrt{k} + 0{,}12 + \frac{0{,}11}{\sqrt{k}}\right) D.$$

Далі, використовуючи табл. 5.9 Колмогорова, слід прийняти або відкинути гіпотезу про те, чи є емпіричним розподіл із заданою нами ймовірністю $Q$ теоретичним чи ні. Для прийняття гіпотези маєї= бути: $\lambda < \lambda_{tabl}$.

Таблиця 5.8 – Таблиця критерію Колмогорова

| $Q$ | 0,85 | 0,90 | 0,95 | 0,99 |
|---|---|---|---|---|
| $\lambda$ | 1,14 | 1,22 | 1,36 | 1,63 |

Слід зауважити, що критерій Колмогорова не єдиний можливий до застосування при оцінюванні; можна використовувати критерій Хі-квадрат, критерій Андерсона-Дарлінга та інші.

## 5.13. Оцінка точності статистичних характеристик

Вкрай важливим є питання, скільки експериментів слід зробити, щоб можна було довіряти знятим характеристикам. Якщо експериментів мало, то характеристика не достовірна. Зазвичай дослідник задає надійну ймовірність, тобто ймовірність, з якою він готовий довіряти знятим характеристикам. Чим більше буде задана надійна ймовірність, тим більше експериментів потрібно зробити. Раніше ми користувалися й іншими способами оцінки необхідної кількості експериментів.

Наша оцінка буде ґрунтуватися на центральній граничній теоремі, згідно з якою сума (або середнє) випадкових величин є величина не випадкова. Відповідно до ЦГТ значення обчисленої нами статистичної характеристики будуть розподілені за нормальним законом, $n_i$ – кількість $i$-тих результатів значення статистичної характеристики в $n$ експериментах, $P_i = n_i/n$ – частота $i$-го результату.

Якщо $n \to \infty$, то $p \to P$ (частота $p$ прямує до теоретичної ймовірності $P$) і емпіричні характеристики будуть прямувати до теоретичних (рис. 5.48). Отже, згідно з ЦПТ $p$ буде розподілена за нормальним законом з математичним очікуванням $m$ і середньоквадратичним відхиленням $\sigma$.

При цьому $m = P, \sigma = sqrt\,(p\ (1 - p)\ /\ n)$.

Позначимо як $Q$ надійну ймовірність, тобто ймовірність того, що частота $p$ відрізняється від ймовірності $P$ не більше, ніж на $\varepsilon$. Тоді за теоремою Бернуллі:

$$Q(|p - P| \le \varepsilon) = F\left(\frac{\varepsilon}{\sigma}\right) = F\left(\frac{\varepsilon\sqrt{n}}{\sqrt{p(1-p)}}\right).$$

Величина $\varepsilon$ називається **довірчим інтервалом**. Сенс $\varepsilon$ полягає в тому, що в серії (кожна вибірка $n$) в середньому $\varepsilon \cdot 100\ \%$ довірчих інтервалів містять справжнє значення статистичної характеристики $p$. Як і раніше $F$ – інтеграл від функції нормального закону розподілу, інтегральна функція Лапласа.

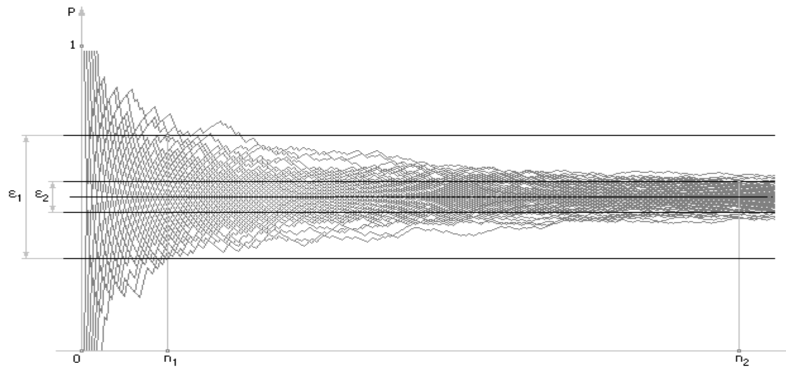


Рис. 5.48. Ілюстрація до обчислення кількості експериментів за величиною довірчого інтервалу згідно з центральною граничною теоремою

Звідси можна висловити необхідну для надійної ймовірності кількість експериментів ($F^{-1}$ – зворотна функція Лапласа):

$$n = \frac{p(1-p)}{\varepsilon^2}\left(F^{-1}(q)\right)^2.$$

**Приклад 12.** При моделюванні продукції, що випускається підприємством в результаті імітації його роботи протягом 50-ти днів були отримані такі вихідні дані (див. табл. 5.9).

Таблиця 5.9 – Експериментальні статистичні дані моделювання

| Якість виробів у балах (випадкова подія $i$) | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Кількість результатів ($n_i$) | 15 | 10 | 5 | 20 |
| Частість результату ($p_i = n_i/n$) | 0,3 | 0,2 | 0,1 | 0,4 |

Тобто всього було проведено: 15 + 10 + 5 + 20 = 50 експериментів ($n = 50$). З таблиці експериментів випливає відповідь завдання, що частість (ймовірність) випуску виробів 1 сорту дорівнює 15/50, частість (ймовірність) випуску виробів 2 сорту дорівнює 10/50, частість (ймовірність) випуску виробів 3 сорти дорівнює 5/50, частість (ймовірність ) випуску виробів 4 сорти дорівнює 20/50.

Задамося надійною ймовірністю до відповідей моделі $Q = 0{,}9$ і довірчим інтервалом $\varepsilon = 0{,}05$.

Тепер потрібно відповісти на запитання: чи можна довіряти з ймовірністю $Q$ обчисленій відповіді?

Будемо оцінювати результат статистичних експериментів за найгіршою імовірністю, такою в нашій задачі є $p = 0{,}4$, оскільки ймовірність, наприклад, 0,1 визначена набагато краще.

Взагалі ймовірності (частості), близькі до 0 або 1, вельми привабливі як відповідь, через те що цілком визначають розв'язок. Ймовірності близькі до 0,5 свідчать про те, що відповідь дуже не визначена, подія трапиться «50 на 50». Таку відповідь задовільною назвати складно, вона мало інформативна.

Формула

$$n = \frac{p(1-p)}{\varepsilon^2}[F^{-1}(Q)]^2$$

після підстановки значень $F^{-1}(0{,}9) = 1{,}65$ (див. таблицю Лапласа), далі $\left(F^{-1}(0{,}9)\right)^2 = 2{,}7$, $p = 0{,}4$, $\varepsilon = 0{,}05$ дає $N = 0{,}4 \cdot 0{,}6\ \frac{2{,}7}{0{,}052}$ або остаточно $N = 250$.

Тобто наш експеримент і його відповідь не достовірні щодо заданих $Q$ і $\varepsilon$: 50, експериментів не достатньо для відповіді, потрібно 250. Отже, необхідно продовжувати експерименти і провести ще 200 експериментів, щоб досягти необхідної точності.

Формула використовує себе рекурентно. Відразу обчислити з її допомогою кількість експериментів $n$ не вдається. Щоб обчислити $n$, необхідно провести пробну серію експериментів, оцінити значення шуканої статистичної характеристики $p$, підставити це значення в формулу і визначити необхідну кількість експериментів.

Для впевненості цю процедуру слід провести кілька разів при різних одержуваних послідовно значеннях $n$.

Отже, в блоці оцінки достовірності (БОД) аналізують ступінь достовірності статистичних експериментальних даних, знятих з моделі (беручи до уваги точність результату $Q$ і $\varepsilon$, задані користувачем) і визначають необхідну для цього кількість статистичних випробувань $n$.

При великій кількості дослідів $n$ частота появи події $p$, отримана експериментальним шляхом, наближається до значення теоретичної ймовірності появи події $P$. Якщо коливання значень частоти появи подій щодо теоретичної ймовірності менші від заданої точності, то експериментальну частоту будуть відповідь, інакше генерацію випадкових вхідних впливів продовжують , і процес моделювання повторюється. При малій кількості випробувань результат може виявитися недостовірним. Але чим більше випробувань, тим точніша відповідь, згідно з центральною граничною теоремою. Кількість необхідних експериментів $n$ дані для порівняння в табл. 5.10 і табл. 5.11 при різних комбінаціях $p$ і ε.

На рис. 5.49 відображено графік залежності $n\ (\varepsilon)$ при $Q = 0{,}95$ і $p = 0{,}5$.

Слід зауважити, що оцінювання ведуть за гіршою з частот. Це забезпечує достовірний результат відразу за всіма характеристиками, що знімаються з моделі.

Слід мати на увазі, що ця оцінка кількості експериментів за ЦПТ не єдина з існуючих. Відомі аналогічні близькі за змістом оцінки Бернуллі, Муавра-Лапласа, Чебишева.

Таблиця 5.10 – Кількість експериментів, необхідних для обчислення достовірної відповіді з надійною ймовірністю $Q = 0,95, (F^{-1}(0,95))^2 = = 3,84, p = 0,1$

| $\varepsilon$ | 0,001 | 0.005 | 0,010 | 0,050 | 0,100 |
|---|---|---|---|---|---|
| Критична кількість експериментів $n$ | 345 600 | 13 824 | 3456 | 138 | 35 |

Таблиця 5.11 – Кількість експериментів $n$, необхідних для обчислення достовірної відповіді з надійною ймовірністю $Q = 0,95, (F^{-1}(0,95))^2 = 3,84, p = 0,5$

| $\varepsilon$ | 0,001 | 0,005 | 0,010 | 0,050 | 0,100 |
|---|---|---|---|---|---|
| Критична кількість експериментів $n$ | 960 000 | 38 400 | 9600 | 384 | 96 |

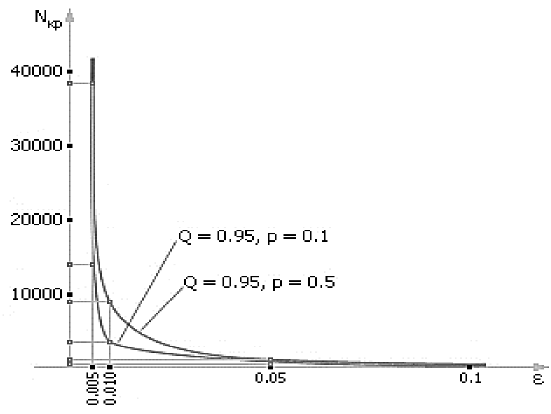


Рис. 5.49. Залежність кількості необхідних експериментів від величини надійрої ймовірності ε і довірчого інтервалу Q для випадку частості випадання випадкової події $p = 0,5$

Як пояснити, чому так дивно поводиться крива знятої експериментально статистичної характеристики (рис. 5.48 і рис. 5.49)? При великому $n$ крива вкрай повільно підходить до істинного значення, хоча спочатку (при малих $n$) процес триває з великою швидкістю – ми швидко входимо в область наближеної відповіді (великі $\varepsilon$), але повільно наближаємося до точної відповіді (малі $\varepsilon$).

Наприклад, припустимо, що ми провели $N$ випробувань. Випадінь події в цих випробуваннях становило кількість $N1$. Нехай ймовірність випадання події близька до $N1\ /\ N\ =\ 0{,}5$ або $N\ =\ N1\ \cdot\ 2$.

Припустимо, що ми хочемо провести ще одне випробування $(N+1)$-е. Взявши відповідь (частість $N1\ /\ N$) при $N$ за 100 %, оцінимо, наскільки відсотків зміниться відповідь після наступного досліду? Складемо пропорцію:

$$\frac{N1}{N} — 100\ \%$$

$$\frac{N1\ +\ 1}{(N\ +\ 1)} — X\ \%\,.$$

Звідси маємо: $X\ =\ (N1\ +\ 1)\ 100\ \cdot\ N\ /\ (N1\ (N\ +\ 1))$, при $N1\ =\ N\ /\ 2$ (ймовірність 0,5) отримуємо, що $X\ =\ 100\ \ (N\ +\ 2)\ /\ (N\ +\ 1)$.

І величина $X$ утворює ряд: $150\ \%,\ 133\ \%,\ 125\ \%,\ 120\ \%, \ldots, 100.1\ \%, \ldots, \ldots \to\ 100\ \%$. Значить, спочатку поліпшення відповіді на один додатковий експеримент становило 50 %, на 2–33 %, на 3–25 %, на 4–20 %, ..., на 100-му – лише на 0,1 %.

Видно, що поліпшення точності на кожен новий експеримент (значення $X$) спочатку дуже гарне, а потім – незначне, після 100 експериментів ця величина змінюється всього на частки відсотка в розрахунку на один додатковий експеримент! Підсумок: зміна оцінки, заснованої на сумі, після серії дослідів перестає сильно змінюватися !!!

*З викладеного вище матеріалу можливо зробити такі висновки.*

1. Як відповідь статистичного експерименту приймається частість $p$ появи деякої вихідної події, яка є оцінкою ймовірності. Чим більше експериментів $n$, тим ближче частість $p$ до ймовірності $P$, а експериментальна відповідь до теоретичної.

2. Частості $p$, близькі за значенням до 0 або 1, більш кращі щодо інформативності, ніж частості, близькі до 0,5, які мало інформативні і дають максимально невизначену відповідь.

3. У моделюванні важливою метою є зниження дисперсії відповіді, розкиду вихідної величини моделі щодо частості. Дійсно, якщо розкид випадкової вели-

чини $m_2$ малий, то обчислена відповідь досить достовірна. Якщо в ряду випадкової величини зустрічаються значення досить віддалені один від одного (див. рис. 5.42), то $m_2'$ велика, і відповідь недостатньо визначена.

4. Статистична відповідь оцінюється не тільки значеннями частості і розкиду, а й точністю, функцію якої виконують надійна ймовірність $Q$ і заданий довірчий інтервал $\varepsilon$. Ці величини пов'язані з розкидом $m_2$.

5. Необхідна кількість статистичних експериментів $n$ залежить від заданої точності $(Q, \varepsilon)$ і характеристик процесу (частості $p$ і розкиду $m_2$). Підвищення вимог за точністю, погані характеристики істотно підвищують витрати на дослідження моделі, збільшуючи кількість експериментів.

## Контрольні запитання

1. Що є метою статистичного моделювання?
2. Що є статистичною моделлю випадкового процесу?
3. Назвіть етапи статистичного моделювання методом Монте-Карло.
4. Які складові має структура методу статистичного моделювання?
5. Яка послідовність випадкових чисел використовується як базова при статистичному моделюванні на ЕОМ?
6. Суть моделювання випадкової події й повної групи несумісних подій.
7. Чому послідовності чисел, що генеруються на ЕОМ, називаються псевдовипадковими?
8. Які методи можуть використовуватися при моделюванні випадкової величини із заданим законом розподілу?
9. Властивості та методи генерації нормально розподілених випадкових величин.
10. Суть моделювання системи випадкових величин.
11. Потоки випадкових подій та їх характеристики.
12. Фіксація і обробка статистичних результатів.
13. Оцінка точності статистичних характеристик.
14. Що є надійною ймовірністю та довірчим інтервалом?

# 6. ІМІТАЦІЙНЕ МОДЕЛЮВАННЯ СИСТЕМ

## 6.1. Загальні відомості про імітаційне моделювання

Імітаційне моделювання є методом дослідження поведінки систем. Цей метод полягає в тому, що за допомогою ЕОМ відтворюється поведінка досліджуваної системи, а дослідник, управляючи ходом процесу імітації й переглядаючи одержувані результати, робить висновок про її властивості та якість поведінки. Тому імітацію слід розуміти як числовий метод проведення на ЕОМ експериментів із математичними моделями, які описують поведінку системи для визначення функціональних характеристик, що цікавлять нас. Поява імітаційного моделювання й перетворення його в ефективний засіб аналізу складних систем були, з одного боку, зумовлені потребами практики, а з іншого – забезпечені розвитком методу статистичних випробувань, що відкрив можливість моделювання випадкових факторів, яких багато в реальних системах, а також розвитком електронної обчислювальної техніки, що є базою для проведення статистичних експериментів.

Досить широке застосування методу імітації при дослідженні поведінки системи зумовлено такими причинами:

- складністю моделі поведінки системи, наявністю великої кількості випадкових факторів, які обмежують ефективність застосування традиційних аналітичних методів дослідження, а в ряді випадків взагалі виключають можливість їх застосування, у результаті чого імітаційне моделювання виявляється єдиним способом дослідження;
- новими можливостями, що дозволяють здійснювати:
  – *спостереження* за поведінкою системи в таких умовах, у яких натурний експеримент просто неможливий (або з суто фізичних причини, або через обмеженість часових і вартісних ресурсів);
  – *проведення імітаційних експериментів* у широкому діапазоні зміни параметрів системи й зовнішнього середовища, що дозволяє одержати корисну інформацію в умовах інформаційної невизначеності, яка завжди супроводжує початкові етапи розв’язання системотехнічних задач;
  – *прогнозування* поведінки системи дозволяє одержувати до того ж відповідь у стислому масштабі часу;
- детальне спостереження за поведінкою імітованої системи дозволяє краще

зрозуміти зміст самої системи й розробити такі пропозиції щодо її поліпшення, які були б неможливі без імітації;

- імітаційне моделювання дозволяє створити уявлення про те, які з параметрів системи є найбільш істотними;
- імітаційне моделювання може бути використане як педагогічний прийом для навчання студентів та інженерів основним навичкам теоретичного аналізу, статистичного аналізу й теорії прийняття рішень.

Але, як і в будь-якого інструмента дослідження, у методу імітації є переваги й недоліки. До недоліків можна віднести таке:

- у ряді випадків імітаційні моделі виявляються досить складними, що вимагає більших часових і вартісних витрат на програмування, налагодження моделей та проведення експериментів;
- «імітаційний світ», як і реальна дійсність, виявляється важко збагненним, оскільки складна імітаційна модель призводить до такої кількості різноманітних наслідків, що в результаті одержувану інформацію не так вже й легко інтерпретувати;
- аналіз результатів імітації оснований тільки на використанні математичної статистики, а як відомо, для одержання статистичної достовірності результатів, достатньої для обґрунтування вибору варіанта керування, варіанта побудови системи й інших, потрібне багаторазове повторення імітаційних експериментів, що в ряді випадків потребує більших часових витрат. Однак незважаючи на це флуктуації результатів, що все одно залишаються, змушують проявляти обережність при підбитті підсумків машинних імітаційних експериментів;
- імітаційне моделювання поки не володіє добре методично обґрунтованими принципами побудови моделей для широкого класу систем керування, а тому кожний конкретний випадок потребує значного спеціального опрацювання.

При імітаційному моделюванні можна виділити такі основні етапи дослідження:

- формулювання задачі;
- побудова математичної моделі функціонування системи;
- складання й налагодження програми для ЕОМ, включаючи й розробку процедур моделювання різних випадкових факторів;
- планування імітаційних експериментів;
- проведення експериментів і обробка результатів дослідження.

Розглянемо більш докладно зміст кожного з етапів.

***Формулювання задачі*** передбачає визначення або питань, на які потрібно відповісти, або гіпотез, які треба перевірити, або впливів, які потрібно оцінити, що загалом визначає мету імітації, відповідно до якої повинні бути визначені й критерії, за якими оцінюють результати імітації.

***Побудова математичної моделі*** містить у собі визначення вхідних, вихідних, керуючих змінних та їх взаємозв'язки в загальному алгоритмі функціонування системи з метою оцінки значень вибраних критеріїв. У випадку машинної імітації математична модель часто подається у вигляді алгоритмічного опису модельованого процесу. Основою для цього є змістовний опис процесу функціонування системи. При побудові моделі необхідно враховувати два суперечливі фактори. Ускладнення моделі, тобто включення в модель великої кількості змінних, призводить до більших часових витрат на складання, налагодження моделі; збільшується й сам час проведення імітаційних експериментів, а в деяких випадках виникають труднощі з інтерпретацією результатів. У результаті може бути втрачена цінність отриманих результатів через їх великий час запізнювання. Спрощення моделі може призвести до втрати змістовності, тобто модель стає неадекватною системі.

На цьому етапі визначають, які зі змінних є випадковими, а які – детермінованими. Після визначення структури моделі проводиться оцінка значень її параметрів, якій передує етап збору необхідної вихідної інформації. Цей етап повинен обов'язково закінчитися перевіркою адекватності моделі об'єкта. Загальної методики перевірки адекватності не існує. Модель вважається адекватною об'єкту дослідження за наявності позитивних відповідей на ряд таких запитань:

- чи немає в моделі несуттєвих змінних, які не поліпшують здатність передбачати поведінку системи?
- чи всі істотні вхідні й керуючі змінні включені в модель?
- чи правильно сформульовані функціональні зв'язки між вхідними й вихідними змінними системи?
- чи правильно визначені параметри системи?
- чи є оцінки випадкових параметрів побудованої моделі статистично значущими?

Правильне й обґрунтоване розв'язання задач 1 і 2-го етапів багато в чому визначає успіх імітаційного експерименту.

***Складання машинної програми.*** Під час складання машинної програми вирішуються такі задачі:

- складання самої програми з використанням як універсальних алгоритмічних мов, так і проблемно-орієнтованих на розв'язання задач імітації;
- розробка програмних процедур імітації різних випадкових факторів, наявних у системі;

➢ налагодження програми.

***Планування експериментів.*** Під час планування експерименту вирішуються такі основні завдання:

➢ вибір способів прискорення збіжності статистичних оцінок критеріїв, які нас цікавлять, до дійсних значень;

➢ визначення обсягу імітаційних експериментів;

➢ складання плану проведення машинних експериментів, що особливо важливо при розв'язанні задач оптимізації на основі імітації. Розв'язання задач вказаних вище і становить зміст етапу планування експериментів.

***Проведення експериментів та обробка результатів.*** При проведенні експериментів та обробки результатів мається на меті: використовуючи все різноманіття статистичних критеріїв і максимум інформації, отриманої в процесі експерименту, зробити висновки за результатами імітаційного експерименту й визначити їх точність.

Перейдемо до розкриття змісту основних питань, пов'язаних із розбором принципів побудови математичних моделей функціонування систем, орієнтованих на використання ЕОМ, припускаючи, що задачу дослідження сформульовано й критерії оцінки визначено. Функціонування багатьох систем можна розглядати як процес переходу системи з одного стану в інший, причому зміні стану завжди передує поява певної дискретної події. Тому такі системи набули назву «системи з дискретними подіями».

Імітаційні моделі, методика побудови яких розглядається нижче, є дискретними, динамічними й стохастичними. Такі моделі називаються дискретно-подієвими імітаційними моделями.

## 6.2. Сутність методу імітаційного моделювання та сфери його застосування

Основний принцип методу імітаційного моделювання полягає в побудові штучного імовірнісного процесу, параметри якого давали б розв'язок поставленої задачі, причому сама задача може й не бути імовірнісною. При цьому можна сформулювати дві задачі: пряму й зворотну.

***Пряма задача.*** Необхідно реалізувати операцію, результат якої визначається випадковими факторами. Тому ця операція являє собою випадковий процес. Шляхом проведення серії експериментів із подальшою статистичною обробкою результатів окремих спостережень можна знайти наближене значення імовірнісних характеристик цього процесу. Для цього здійснюється багаторазове проведення цієї операції за фіксованих умов. Кожне спостереження являє собою реалізацію випадкового процесу, а кожний результат – випадкову величину, підпорядковану тому закону розподілу, який і відшукується при проведенні експериментів.

***Зворотна задача.*** Відомі імовірнісні характеристики процесу й необхідно відтворити одну з його реалізацій. Процес цього відтворення називають статистичним випробуванням. При багаторазовому випробуванні й подальшій обробці можна знайти статистичні оцінки вихідних імовірнісних характеристик.

Нехай відомо, що ймовірність здійснення якоїсь події дорівнює 0,6 (наприклад, це ймовірність безвідмовної роботи знову встановленого обладнання без відповідного налаштування). Елементарною статистичною моделлю процесу роботи обладнаних може бути звичайна традиційна урна з 10 кулями (6 білих і 4 чорних). Реалізація одного випробування полягає в добуванні кулі, подальшому її поверненні в урну й перемішуванні куль. При великій кількості випробувань статистична оцінка ймовірності дорівнює 0,6, тобто збігається з імовірнісною характеристикою процесу. Але, очевидно, не доцільно будувати модель із метою одержання заздалегідь відомого результату. Оскільки складний процес може складатися з комплексу простих процесів і при цьому зв’язок між ними може бути невідомим, крім їх імовірнісних характеристик, то виявляється можливим знаходження за допомогою імітаційного моделювання імовірнісних характеристик усього процесу.

Нехай, наприклад, будуються моделі функціонування систем масового обслуговування. Відомо, що складність побудови імовірнісних моделей, у яких необхідно враховувати вплив ряду випадкових факторів, різко зростає зі збільшенням їх кількості. Тому побудова аналітичних моделей систем або процесів, основаних на математичному апарату теорії ймовірностей, вдається тільки при значних спрощеннях або навіть при відмові від тих або інших факторів.

Суть методу імітаційного моделювання проста й полягає в такому. Замість опису випадкового явища аналітичними залежностями проводиться моделювання (розіграш) цього випадкового явища за допомогою певної процедури, що дає випадковий результат. Таким чином виходить одна реалізація випадкового явища. Після проведення великої кількості таких реалізацій цього явища проводиться обробка результатів моделювання методами математичної статистики. Цими розіграшами можна вирішувати будь-які імовірнісні задачі, але виправданими (вигідними) вони стають тоді, коли процедура моделювання простіша за застосування аналітичних, обчислювальних методів.

Нехай досліджується багатоканальна СМО, але процес, що протікає в ній, не є марківським. Закон розподілу інтервалів між моментами появи заявок не показовий, а довільний. Тривалість часу обслуговування заявок теж не є показовою, а має якийсь інший розподіл. Крім того, канали можуть іноді виходити з ладу, а час безвідмовної роботи має довільний розподіл. Потрібно знайти ймовірності станів СМО і показники її ефективності. Створення аналітичної моделі такої СМО може бути ускладненим, тому потрібно (легше) побудувати імітаційну модель або використовувати вже готову.

Імітаційне моделювання базується на використанні генераторів випадкових величин у комбінації з інтегральною функцією розподілу ймовірностей для досліджуваного процесу. Розподіл імовірностей, що розігрується, оснований на емпіричних даних або є відомим теоретичним законом. В основі загальної схеми методу імітаційного моделювання є центральна гранична теорема теорії ймовірностей.

Отже, коротка послідовність етапів проведення імітаційного моделювання:

- розробляється й реалізується на ЕОМ детермінована математична модель системи, яка відображає зв’язок значень вихідних параметрів системи зі значеннями вхідних впливів, початкових умов і структури системи;
- забезпечується одержання на ЕОМ окремих реалізацій випадкових подій, величин, функцій, тобто моделюється випадкове явище із заданими характеристиками, відповідними до характеристик випадкових явищ, що супроводжують функціонування реальної досліджуваної системи;
- проводиться багаторазова реалізація детермінованого процесу, де в кожній із реалізацій враховується вплив випадкових явищ;
- проводиться статистична обробка отриманих результатів відповідно до характеру імітованого процесу.

## 6.3. Моделі систем масового обслуговування

Формальний опис процесів у вигляді системи масового обслуговування широко застосовується в різних галузях науки й практики. Теоретичною базою побудови й дослідження СМО є теорія масового обслуговування (ТМО).

Теорія масового обслуговування – це область прикладної математики, що займається аналізом процесів у СМО, у яких події повторюються багаторазово. За допомогою цієї теорії розробляються методи розв’язання типових задач масового обслуговування, будуються моделі СМО й визначаються їх кількісні характеристики.

Методами ТМО аналізують функціонування об'єкта, а потім вирішують питання щодо синтезу обслуговуючих обладнань і вибору оптимальних параметрів системи.

### *6.3.1. Структурна схема СМО та її компоненти*

Можна розглянути роботу деякого складу, в якому є декілька *К* розвантажувальних терміналів. Використовуючи теорію масового обслуговування, можна сказати, що є СМО з *К* каналами обслуговування. Тривалість обслуговування транспортних засобів (тривалість розвантаження) залежно від типу засобу і його вантажу коливається від 1 до 2 годин. Коли всі термінали зайняті, то прибулий транспорт стає в чергу на розвантаження. Кожний транспортний засіб має графік прибуття, але через складну обстановку на дорогах (непередбачені обставини) цей графік зазвичай порушується. Тому вважається, що прибуття транспорту є випадковою подією з якимись значеннями математичного очікування й дисперсії.

У випадку, коли весь транспорт однаковий (з погляду пріоритетності обслуговування), вважають, що в СМО надходить однорідний потік заявок, тобто всі заявки на розвантаження рівноцінні. Але часто виникають і деякі неоднорідності. Наприклад, у разі надходження швидкопсувного вантажу заявці на обслуговування привласнюється деякий пріоритет (той, що прийшов останнім, обслуговується першим). Найпростіша схема СМО зображена на рис. 6.1.

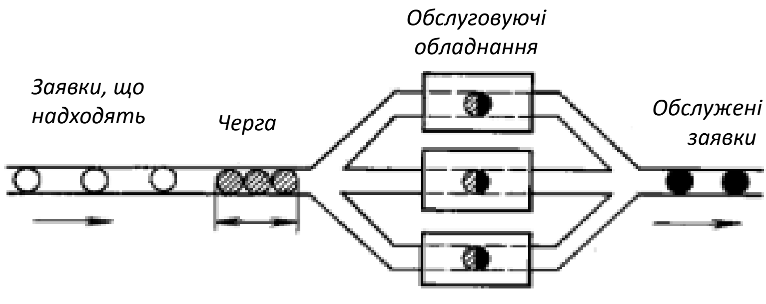


Рис. 6.1. Структура найпростішої СМО

У загальному випадку інтервали між заявками, які надходять, не однакові: це випадкові величини, зумовлені ймовірнісним законом розподілу вхідного потоку. Заявки, що стали в чергу, чекають початку обслуговування відповідно до дисципліни черги (дисципліни обслуговування). Вони можуть бути різними: перший прийшов – перший іде на обслуговування або навпаки – останній прийшов – перший обслуговується.

Обслуговуючі обладнання (прилади, канали) здійснюють обслуговування своєї заявки (що надійшла на вхід) відповідно до заданого детермінованого або випадкового закону розподілу. Обслужена заявка надходить у вихідний потік, який відрізняється від вхідного й залежить від дисциплін черги та обслуговування.

Характерна риса СМО – усі явища описуються за допомогою подій, що з'являються в ті або інші моменти часу. На часовій осі вхідний потік заявок позначається як послідовність подій; вибірка заявок із черги або закінчення обслуговування – це теж події у відповідні моменти часу. Важливою обставиною у цьому випадку тут є абстрагування від усіх інших несуттєвих властивостей реальної системи, що не вписуються у схему послідовності подій.

### *6.3.2. Класифікація систем масового обслуговування*

Можна використовувати велику різноманітність моделей СМО й підходів до їхньої класифікації. Спочатку моделі розділяють на марківські й немарківські. Цей напрямок пов'язаний із класом певно марківських випадкових процесів. Моделі СМО, у яких протікають саме марківські випадкові процеси, описуються системою диференціальних рівнянь, або, у граничному випадку, системою лінійних алгебраїчних рівнянь. Розв'язання цих рівнянь дає можливість легко одержати набір «гарних» виразів, що легко запам'ятовуються, які визначають показники ефективності функціонування СМО. За наявності немарківських процесів у СМО аналітичного дослідження зазнають лише деякі окремі випадки.

Як правило, використовуються такі класифікаційні ознаки (рис. 6.2):

- організація потоку заявок;
- характер утворення черги;
- обмеження черги;
- кількість обслуговуючих каналів;
- дисципліна черги.

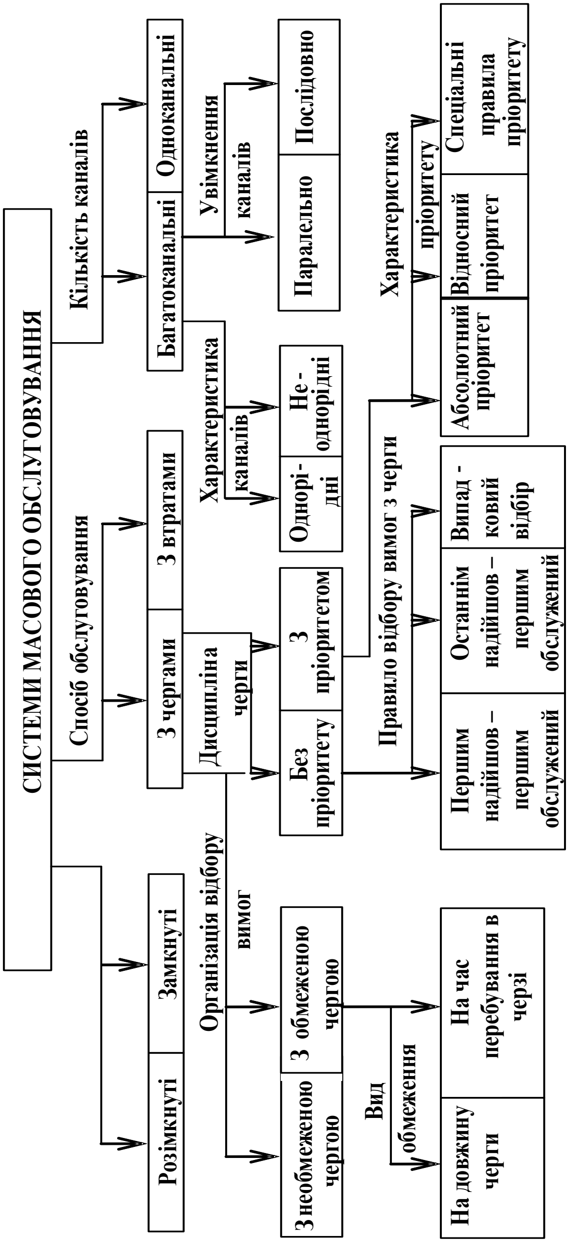


Рис. 6.2. Класифікація систем масового обслуговування

За кількістю каналів СМО поділяються на одноканальні й багатоканальні системи. Багатоканальні СМО, у свою чергу, поділяються на системи з однаковими параметрами каналів обслуговування (рівноцінними каналами) і системи з різними параметрами каналів обслуговування (нерівноцінними каналами).

Залежно від взаємного розташування каналів СМО підрозділяються на системи з послідовними й паралельними каналами. У СМО з паралельними каналами вхідний потік є загальним, і заявка може обслуговуватися будь-яким вільним каналом. У СМО з послідовним розташуванням каналів кожний канал можна розглядати як одноканальну СМО, і тоді вихідний потік одного з них одночасно є вхідним для подальшого каналу. Такі системи називаються відповідно багатофазними й однофазними. У загальному випадку з деякими допущеннями багатофазна СМО є декількома послідовно з'єднаними однофазними системами. Надалі будуть розглядатися тільки однофазні СМО й, якщо вони багатоканальні, то – з рівноцінними каналами.

Залежно від характеру утворення черги СМО можна розділити на два типи:

- системи з відмовами в обслуговуванні;
- системи з очікуванням обслуговування, тобто із чергою.

У першому випадку заявка, що надійшла в СМО, і застала канал (для одноканальної системи) зайнятим або – канали (для багатоканальної системи) зайнятими, одержує відмову та йде зі СМО необслуженою.

Особливістю СМО з очікуванням є характеристика самої черги (обмеження на чергу). Такі системи підрозділяються на СМО з необмеженим очікуванням (необмеженою кількістю місць у черзі) і СМО з обмеженим очікуванням (обмеженою кількістю місць у черзі). Перший варіант обмеження полягає в тому, що заявка завжди ставиться в чергу й коли-небудь буде обслужена. Другий варіант для постановки заявки в чергу потребує наявності вільних місць у ній. І тільки в цьому випадку заявка ставиться в чергу. Але можливий і інший вид обмеження – за часом очікування заявки початку обслуговування. У цьому випадку заявка залишає систему необслуженою, якщо тривалість її можливого перебування в черзі перевищує заданий ресурс часу очікування (це так звані «нетерплячі» заявки). За дисципліною черги (тобто за критерієм вибору із черги чергової заявки) СМО підрозділяються на системи із пріоритетом і без пріоритету. У першому випадку заявці, яка надійшла у СМО, привласнюється деякий пріоритет, відповідно до якого на обслуговування відправляється заявка з найвищим пріоритетом, незважаючи на те, що, можливо, вона надійшла в систему після надходження інших заявок. У цьому випадку насправді утворюється декілька черг із різним рівнем пріоритету. При реалізації дисципліни черги без пріоритету можливі різні правила відбору заявок на обслуговування: перший прийшов – перший обслужений, останній прийшов – перший обслужений або у випадковому порядку (наприклад, із заданими ймовірностями вибірки із черги).

Системи із пріоритетом, у свою чергу, можуть підрозділятися на СМО з перериванням і без переривання. У системах з перериванням заявка з вищим пріоритетом перериває обслуговування заявки з більш низьким пріоритетом. Перервана заявка може бути поставлена в чергу й, можливо, коли-небудь буде обслужена відповідно до свого пріоритету.

Зручно користуватися формальним класифікаційним описом СМО як абстрактною системою. При цьому абстрактну СМО розуміють як математичний об'єкт $\alpha|\beta|\gamma|\delta$, що характеризується такими чотирма елементами: $\alpha$ – вхідним потоком вимог; $\beta$ – потоком обслуговування; $\gamma$ – структурою системи (кількістю обслуговуючих каналів), $\delta$ – дисципліною обслуговування (характером утворення черги й наявністю обмеження на чергу).

Абстрактна система, наприклад, виду $M|M|3|5$ – це СМО з марківськими вхідним потоком і потоком обслуговування, кількістю обслуговуючих приладів, яке дорівнює 3, місць у черзі, що дорівнює 5.

Слід зазначити, що теорія СМО оперує не самими системами розподілу інформації (СРІ), а їх математичними моделями. Для повного опису СРІ необхідно вказати імовірнісні процеси, що описують вхідний потік вимог, структуру системи та дисципліну обслуговування. Таким чином, математична модель СРІ містить такі основні елементи.

1. Вхідний потік вимог на обслуговування (трафік) класифікується за ознаками стаціонарності, ординарності та післядії. Основними характеристиками потоку вимог є його параметр й інтенсивність.

2. Структура системи розподілу інформації – це інформація про кількість обслуговуючих пристроїв або серверів (пристроїв, що надають послуги), їх взаємне з'єднання (схема) та доступність для вхідних вимог.

3. Дисципліна обслуговування потоку вимог характеризує взаємодію потоку вимог із системою розподілу інформації. У теорії СМО дисципліна обслуговування описується:

- способом обслуговування вимог;
- порядком обслуговування вимог;
- режимами пошуку виходів схеми (наприклад, довільний або груповий);
- законами розподілу тривалості обслуговування;

➢ наявністю переваг (пріоритетів) в обслуговуванні деяких вимог;

➢ наявністю обмежень при обслуговуванні (наприклад, за тривалістю очікування або обслуговування, кількістю вимог, що очікують);

➢ законами розподілу імовірностей виходу з ладу елементів схеми.

У загальному випадку вхідний потік вимог на обслуговування описується функцією розподілу ймовірностей інтервалів часу між сусідніми вимогами $P(\leq z)$:

$$A(z) = P(\leq z),$$

де $P(\leq z)$ – імовірність того, що час між послідовними вимогами $(\leq z)$.

Якщо інтервали часу між послідовними вимогами є не залежними та однаково розподіленими випадковими величинами, то вхідний потік вимог утворює стаціонарний процес відновлення. Це є ознакою незмінності в часі імовірнісних характеристик випадкових процесів, яка в більшості випадків добре відображає реальні процеси в СМО за короткі проміжки часу. Таким чином, функція розподілу інтервалів $A(z)$ є достатньою для опису потоку вимог.

Час, протягом якого вимога перебуває на обслуговуванні, описується функцією розподілу ймовірностей тривалості обслуговування $B(x)$:

$$B(x) = P(\leq x),$$

де $P(\leq x)$ – імовірність того, що час обслуговування $\leq x$.

Для опису інтервалу часу між послідовними вимогами або тривалості обслуговування застосовуються різні закони. Найбільше використовуються розподіли, що подані далі та позначені певними літерами:

– $M$ – експонентний ($M$ – марківська модель);

– $H$ – гіперекспонентний ($H$yper–exponential);

– $D$ – детермінований (Determined);

– $U$ – рівномірний(Uniform);

– $E$ – розподіл Ерланга;

– $G$ – довільний або узагальнений (General).

Дисципліна обслуговування потоку вимог визначає правила обслуговування та частку вимог при їх надходженні до системи на обслуговування. Розрізняють такі типи СМО (див. рис. 6.2), які визначаються способом обслуговування вимог:

1. Системи з втратами – вимоги, які при надходженні до системи, не знаходячи у ній жодного вільного обслуговуючого апарата, отримують відмову в обслуговуванні та втрачаються.

2. Системи із чергами – вимоги, які не можуть бути обслужені відразу через зайнятість всіх обслуговуючих апаратів системи, стають у чергу, і за допомогою певної дисципліни обслуговування черги визначається, в якому порядку вимоги, що очікують, вибираються із черги для обслуговування. Найбільш поширеними дисциплінами обслуговування черги є:

- FF (FIFO – first in first out) – вимоги із черги обслуговуються в порядку їх надходження (упорядкована черга);
- LF (LIFO – last in first out) – щоразу переваги для обслуговування має вимога, що надійшла до черги останньою;
- SR (SIRO – service in random order) – наступна вимога для обслуговування із черги вибирається випадково (випадкова черга).

3. Комбіновані системи із чергами та втратами (системи із чергою при обмеженнях). Наприклад, очікувати може тільки кінцева кількість вимог, зумовлена кількістю місць очікування, меншою за нескінченність. Можливо й так – вимога втрачається тоді, коли час очікування в черзі або перебування в системі перевищує задані межі.

4. Пріоритетні системи. Для вимог передбачено різні пріоритети в обслуговуванні. Якщо вимога, що надійшла, має високий пріоритет, а всі сервери зайняті, то вона або займає одне з перших місць у черзі, або тимчасово припиняє обслуговування вимоги низького пріоритету й займає її місце в сервері. При цьому можуть бути застосовані такі пріоритетні правила:

- абсолютний пріоритет із перериванням (pre-emptive discipline) – вимога високого пріоритету перериває обслуговування вимоги низького пріоритету. Може бути: абсолютний пріоритет із втратами (pre-emptive loss discipline), абсолютний пріоритет із дообслуговуванням (pre-emptive resume discipline) та абсолютний пріоритет із обслуговуванням заново (pre-emptive repeat different discipline);
- відносний пріоритет (head of the line priority discipline) – вимога високого пріоритету займає перше місце в черзі й переривань немає.

Змішані пріоритети зумовлюють вибір абсолютного або відносного пріоритетного правила залежно від уже реалізованої частини тривалості обслуговування, а динамічні – залежно від типу поточних вимог і співвідношення кількості вимог різних пріоритетів, що є в серверах та в черзі.

Основні характеристики, що становлять структуру СРІ, такі:

- кількість обслуговуючих пристроїв (серверів, ліній, каналів, портів);
- кількість місць очікування або максимальна довжина черги (ємність пам'яті, в якій накопичуються вимоги, що очікують);
- доступність – спосіб увімкнення серверів, за якого кожній вимозі доступні

всі або не всі (хоча всім вимогам у сукупності доступні всі) ВК. Схема може бути повнодоступною або неповнодоступною;

- взаємне з'єднання (схема) – спосіб увімкнення серверів, за якого кожна вимога обслуговується одним сервером або декількома, але поетапно.

Структурні характеристики системи частково зумовлюють дисципліну обслуговування потоку вимог. Наприклад, за кількості місць очікування $r = 0$ буде система з втратами, при $0 < r < \infty$ – комбінована система із чергою та зі втратами, а при $r = \infty$ – чиста система з чергою.

Отже, базова математична модель СМО позначається послідовністю символів: перший вказує функцію розподілу інтервалів часу між вимогами, другий – функцію розподілу тривалості обслуговування, третій і наступний (необов'язковий) символи – схему й дисципліну обслуговування.

Побудова математичної моделі, що адекватно відображає реальну систему розподілу інформації, у багатьох випадках не є простою задачею. Від правильного вибору моделі залежить точність розв'язання задач аналізу, синтезу та оптимізації.

### *6.3.3. Показники ефективності роботи СМО*

Якість функціонування будь-якої системи визначається кількісними показниками ефективності її роботи. Залежно від цілей дослідження ці показники можуть бути різними. Наприклад, для оцінки роботи певної СМО можна використовувати середню величину втрат $C$ за одиницю часу системи, що відбуваються в процесі експлуатації

$$C = k_1 c_1 + k_2 c_2,$$

де $k_1$ – середня кількість заявок у черзі на обслуговування; $c_1$ – вартість очікування в черзі за одиницю часу; $k_2$ – середня кількість вільних каналів; $c_2$ – вартість простою одного каналу за одиницю часу.

Прийнято оцінювати ефективність СМО за такими показниками:

- імовірність обслуговування клієнта системою;
- пропускна здатність системи;
- імовірність відмови клієнтові в обслуговуванні;
- імовірність зайнятості кожного з каналу й усіх разом;
- середній час зайнятості кожного каналу;
- імовірність зайнятості всіх каналів;
- середня кількість зайнятих каналів;

- імовірність простою кожного каналу;
- імовірність простою всієї системи;
- середня кількість заявок, що стоять у черзі;
- середній час очікування заявки в черзі;
- середній час обслуговування заявки;
- середній час знаходження заявки в системі.

Наведений невеликий перелік характеристик уже свідчить про їх тісний взаємозв'язок і суперечливості. Поліпшення показників одних із них зазвичай спричиняє погіршення інших. Звідси випливає, що при проектуванні СМО бажано встановити розумну рівновагу між різними показниками ефективності системи.

Для розв'язання цієї задачі варто вибрати узагальнений показник $(C)$ ефективності СМО, наприклад, критерій економічної ефективності, – який включав би й витрати обігу $(C_{vo})$, і витрати споживання $(C_{vs})$:

$$C = (C_{vo} + (C_{vs}) \rightarrow \min.$$

Перший доданок узагальненого показника включає в себе витрати, наприклад, на експлуатацію системи й простій обслуговуючих каналів, другий доданок – втрати, пов'язані з відходом необслужених заявок і з перебуванням у черзі.

Методика розрахунку окремих характеристик СМО буде висвітлена нижче при побудові моделей конкретних СМО.

## 6.4. Методологія побудови моделюючих алгоритмів систем масового обслуговування

### *6.4.1. Моделювання потоків заявок*

Основними характеристиками СМО, що визначають особливості її функціонування, є характер вхідного потоку заявок (тобто послідовність подій, спеціально упорядкованих у часі), а також дисципліни очікування й обслуговування.

*Вхідний потік заявок на обслуговування*. Якщо стосовно обслуговування всі заявки потоку рівноправні, то в цей момент часу важливий лише сам факт настання події (надходження заявки). Такі потоки називаються потоками однорідних подій. Кожна подія потоку характеризується моментом часу $t_j$, в який вона настає. У випадку цілком детермінованого потоку однорідних подій послідовність $t_j$ можна одержати, перерахувати (наприклад, сформувавши одновимірний масив у пам'яті комп'ютера), задавши функцію $t_j = t$ або, нарешті, використавши рекурентні співвідношення, що дозволяють одержати поточне значення $t_j$ за попереднім.

Для опису випадкових потоків однорідних подій задається закон розподілу, що характеризує послідовність випадкових величин $t_1,\ t_2, \ldots, t_m, \ldots$. Зазвичай $t_j$ доцільно заміняти величинами $\xi_j$, що визначають довжину інтервалів часу між моментами уведення заявок $t_j$:

$$t_1 = \xi_1,\ t_2 = \ \xi_1 + \xi_2\ , \ldots t_m\ = \xi_1\ + \xi_w + \ldots + \xi_m.$$

Сукупність $\xi_j$ задається спільним законом розподілу. Випадковий потік однорідних подій називають *потоком з обмеженою післядією*, якщо $\xi_j$ є незалежними випадковими величинами. У цьому випадку інтервал $\xi_j$ може бути заданий своєю функцією щільності $f_j\ (z)$.

Численні застосування мають стаціонарні ординарні потоки з обмеженою післядією (потоки Пальма). Потік однорідних подій називається стаціонарним, якщо ймовірність $P_k(t_0, t)$ появи $k$ подій на інтервалі часу $(t_0,\ t_0\ +\ t)$ не залежить від $t_0$, а залежить тільки від $t$ і $k$ (тобто імовірнісний режим потоку не залежить від часу). Потік називається *ординарним*, якщо ймовірність $\varphi(t_0, t)$ появи двох та більше заявок за проміжок часу $(t_0,\ t_0\ +\ t)$ для будь-якого $t_0$ мала порівняно з $t$:

$$\lim_{t \to 0} \varphi(t_0, t)/t = 0.$$

Для стаціонарних потоків з обмеженою післядією є правдивим таке співвідношення: $f_2(z) = f_3(z) = \ldots = \ f_k(z) =\ f(z),$ яке означає, що при $j > 0$ інтервали $\xi_j$ розподілені однаково.

Математичне сподівання $m_\xi$ випадкової величини $\xi_j$ при $j > 0$ дорівнює

$$m_\xi = \int_0^\infty z f(z) dz.$$

Величиною $m_\xi$, по суті, є середня довжина інтервалу між послідовними заявками. Тоді для стаціонарних потоків з обмеженою післядією величина

$$\lambda = 1/m_\xi,$$

яка визначає середню кількість подій за одиницю часу, називається *інтенсивністю потоку*. Для потоків, які розглядаються, правдивою є формула Пальма, що зв'язує $f_1(z_1)$ і $f(z)$ :

$$f_1(z_1) = \lambda\left[1 - \int_0^\infty f(u)du\right] \qquad (6.1)$$

Це співвідношення дозволяє одержати функцію щільності $f_1(z_1)$ за відомою щільністю $f(z)$.

Випадковий потік однорідних подій з обмеженою післядією називається потоком без післядії, якщо ймовірність $P_k(t_0, t)$ надходження $k$ подій за проміжок часу $(t_0,\ t_0\ +\ t)$ не залежить від чергування подій до моменту $t_0$.

При моделюванні СМО важливу роль відіграє найпростіший (стаціонарний, ординарний, без післядії) потік, для якого ймовірність $P_k(t)$ настання $k$ подій за інтервал часу $t$ виражається законом розподілу Пуассона

$$P(t) = \frac{(\lambda t)^k}{k!}\, e^{-\lambda t}.$$

Звідси найпростіший потік іноді називають пуассонівським.

Крім розглянутих потоків, практично цікаві потоки Ерланга. В окремих випадках розглядають нестаціонарні вхідні потоки.

*Дисципліна очікування.* У загальному випадку СМО має $n$ каналів (ліній), що паралельно обслуговують в один потік заявок. Кожний канал може бути або вільний, або зайнятий. Якщо в СМО є вільні канали, то заявка обслуговується, а інакше – очікує в черзі час $\tau$, після чого одержує відмову й залишає систему.

Залежно від значення $\tau$, розрізняють: системи з очікуванням $(\tau = \infty)$, системи з відмовами $(\tau = 0)$ і змішані системи $(0\ \leq \tau \leq\ h)$.

У деяких випадках тривалість очікування або взагалі можливість встати в чергу залежать від обмежень на довжину черги. Звідси розрізняють СМО без обмежень за довжиною черги та з обмеженням за довжиною черги.

За характером обслуговування черги розрізняють СМО з упорядкованою й абсолютно неупорядкованою чергою. У перших – дисципліна черги підпорядковується правилу «раніше прийшов – раніше надійшов на обслуговування», в інших – будь-яка із заявок, що стоять у черзі, при вивільненні каналу з рівною ймовірністю може надійти на обслуговування.

Нарешті, якщо заявки, що надходять у СМО, не рівнозначні, тобто мають різні пріоритети, то говорять про системи без пріоритету на обслуговування й із пріоритетами. Необхідно зазначити, що система пріоритетів і відповідні їй дисципліни очікування можуть бути досить складними.

*Дисципліна обслуговування.* За кількістю каналів СМО підрозділяються на одно- і багатоканальні. Останні, у свою чергу, поділяються на системи з рівноцінними й нерівноцінними каналами.

Тривалість обслуговування є випадковою величиною з показовим законом розподілу й рідше – ерлангівським.

За часом обслуговування розрізняють СМО з необмеженою тривалістю обслуговування й з обмеженням за тривалістю обслуговування.

Крім того, часто враховується надійність каналів обслуговування. У зв'язку з цим розрізняють системи з виходом з ладу каналів. Нарешті, останні підрозділяються на СМО без відновлення каналів обслуговування, які вийшли з ладу, і СМО з відновленням каналів.

*Моделювання потоків заявок*. Нехай у СМО надходить стаціонарний потік однорідних подій з обмеженою післядією (потік типу Пальма).

Потік типу Пальма задається функцією щільності $f(z)$ випадкових інтервалів $z_j, j = 2,3, \ldots$ між послідовними моментами $t_j = t_{j-1} + z_j$ надходження заявок. Для його імітації на комп'ютері потрібно одержати необхідну кількість реалізацій, тобто таких невипадкових послідовностей $t_1^*, t_2^*, \ldots$ моментів надходження заявок у СМО, інтервали між якими $t_j^* - t_{j-1}^*$ були б значеннями випадкової величини $z_j$, описуваними функцією щільності $f(z)$.

Розглянемо процедуру формування послідовності $t_1^*, t_2^*, \ldots$;

1) формуємо $t_1 = z_1$. Функція щільності $f_1(z_1)$ визначається за (6.1);

2) одержимо випадкове число $\xi_j$ з рівномірним розподілом на інтервалі [0,1];

3) формуємо $z_1$, для чого одним з описаних в § 5.5 методів перетворимо $\xi_j$ у випадкову величину $z_1^*$ що має функцію щільності $f_1(z_1)$;

4) вважаємо, що $t_1^* = z_1^*$;

5) формуємо $t_2 = t_1 + z_2$. Для чого, одержавши $\xi_{j+1}$, перетворимо його в $z_2^*$, що має функцію щільності $f(z)$;

6) вважаємо, що $t_2^* = t_1^* + z_2^*$;

7) чергове $t_j^*$ обчислюється за формулою $t_j^* = t_{j-1}^* + z_j^*$.

### *6.4.2. Обробка результатів моделювання СМО*

При вивченні початкового нестаціонарного періоду функціонування СМО моделювання являє собою, по суті, розігрування великої кількості множини реалізацій процесу, а потрібні характеристики (імовірність відмови в обслуговуванні, середня кількість заявок у черзі) виходять обробкою емпіричних даних як статистичне середнє за множиною реалізацій.

При моделюванні стаціонарних процесів можна скористатися тільки однією, але достатньо довгою реалізацією, а характеристики, що цікавлять, розрахувати як середні за часом.

Припустимо є двоканальна ($n = 2$) СМО із чергою, максимальна довжина якої – три заявки. Якщо всі три місця в черзі зайняті, заявка, що знову надійшла, одержує відмову й залишає систему. Потік заявок – пальмівський. Час обслуговування однієї заявки – випадкова величина, розподілена за показовим законом.

Визначимо за однією, але досить довгою реалізацією такі характеристики: імовірність зайнятості $0, 1, 2$ каналів – $P_0, P_1, P_2$; імовірності того, що в черзі буде $0, 1, 2, 3$ заявки – $\overline{P_0}$ , $\overline{P_1}$ , $\overline{P_2}$ , $\overline{P_3}$; середня кількість зайнятих каналів – $z$ середній час очікування заявки в черзі – $t_{och}$; середній час обслуговування заявок – $t_{ob}$ ; імовірність відмови в обслуговуванні – $P_{vidm}$.

Часову діаграму функціонування СМО зображено на рис. 6.3. Угорі дано вісь часу (0) з відзначеними моментами надходження заявок ($t_1, \dots, t_{14}$). На осях 1 і 2 показані стани 1-го й 2-го каналів (відповідно тонка лінія – «вільний», жирна – «зайнятий»). На осях 3-5 фіксуються стани 1-3 місць у черзі відповідно.

Перед появою першої заявки $t_1$ усі канали й місця в черзі вільні. У момент $t_1$ надходить заявка і займає перший канал. Час обслуговування $\tau_1$ формується в результаті роботи генератора. У момент $t_2$ надходження наступної заявки перший канал зайнятий, і вона надходить на другий.

Генерується ще одне значення часу обслуговування $\tau_2$. Заявка $t_3$ надходить у систему, коли обидва канали зайняті, надходить у чергу й чекає звільнення одного з них. Раніше закінчує обслуговування канал 2 і третя заявка починає ним оброблятися (час обслуговування $\tau_3$ виходить аналогічно $\tau_1$ і $\tau_2$). Далі процес повторюється. Для наочності на рис. 6.3 під кожною ділянкою зайнятості каналу проставлений номер заявки, що дозволяє простежити її «шлях» у СМО.

Обробимо результат моделювання. У першу чергу знайдемо $P_0, P_1, P_2$. Для цього розділимо часову вісь на ділянки й позначимо їх цифрою 0, якщо не зайнятий жоден канал, 1 – якщо зайнято один, і 2 – якщо зайнято два канали.

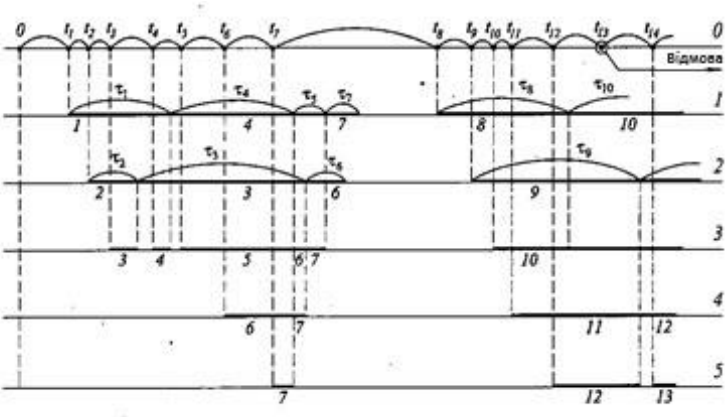


Рис. 6.3. Часова діаграма функціонування двоканальної СМО з відмовами

Виберемо велику ділянку часу моделювання $T$ (рекомендується брати не із самого початку процесу, де позначається вплив початкових умов), на якому просумовуємо довжини ділянок, позначених нулем, одиницею та двійкою. Очевидно, що

$$T_0 + T_1 + T_2 = T.$$

Тоді

$$P_0 \approx \frac{T_0}{T}, P_1 \approx \frac{T_1}{T}, P_2 \approx \frac{T_2}{T}.$$

Аналогічно зробимо для визначення характеристик $\overline{P_0}$ , $\overline{P_1}$ , $\overline{P_2}$ , $\overline{P_3}$:

$$\overline{P_0} \approx \frac{\overline{T_0}}{T}, \overline{P_1} \approx \frac{\overline{T_1}}{T}, \overline{P_2} \approx \frac{\overline{T_2}}{T}, \overline{P_3} \approx \frac{\overline{T_3}}{T},$$

де $\overline{T_i}$ − сума довжин ділянок, позначених $i$-тою цифрою.

Середня кількість зайнятих каналів $z$ визначиться як

$$\bar{z} = 1P_1 \, + \, 2P_2.$$

Середній час очікування заявки в черзі $\bar{t}_{och}$ визначиться так: виберемо велику ділянку $T$ і розглянемо заявки, що надійшли в моменти $t_k, \, t_{k+1}, \dots, t_{k+i}, \dots, t_{k+N}$, для кожної з яких за емпіричними даними підрахуємо час очікування в черзі $t_{och}^{k+i}$, який дорівнює нулю, якщо $k+i$ -та заявка була відразу обслужена (або одержала відмову), або за сумою часів очікування цієї заявки на осях 3-5. Тоді

$$\bar{t}_{och} = \frac{1}{N}\sum\nolimits_{i=0}^{N} t_{och}^{k+i}.$$

Середній час обслуговування заявки за аналогією з $\bar{t}_{och}$ можна знайти за формулою

$$\bar{t}_{obs} = \frac{1}{N}\sum\nolimits_{i=0}^{N} t_{obs}^{k+i},$$

де $t_{obs}^{k+i}$ – час обслуговування $(k+i)$-тої заявки.

Імовірність відмови можна розрахувати за формулою

$$P_{vidm} = \frac{N^*}{N},$$

де $N^*$ – кількість заявок, що одержали відмову; $N$ – загальна кількість заявок за час $T$.

### *6.4.3. Найпростіші моделі одноканальної й багатоканальної СМО*

Як уже було сказано, з усього різноманіття імітаційних моделей СМО будуть розглядатися тільки дискретно-подієві моделі. Це такі моделі, які розбудовуються в часі, а стани системи змінюються миттєво в конкретні моменти часу.

При моделюванні дослідник має справу з різними поданнями часу. По-перше, потрібно враховувати реальний час, у якому функціонує імітована система. По-друге, робота моделі здійснюється в певному масштабі – це модельний (системний) час. І, нарешті, по-третє, – витрати машинного часу (часу роботи комп'ютера) на реалізацію самої імітації. Перші два виміри часу мають найважливіше значення при побудові моделі. З урахуванням цих часових факторів реалізуються переходи системи зі стану в стан; здійснюється синхронізація функціонування елементів системи, задається масштаб часу функціонування («життя») імітованої системи; здійснюється керування технологічним процесом модельного експерименту.

При реалізації моделюючих алгоритмів на цифрових машинах є певні загальні закономірності.

Основна задача при складанні алгоритмів на машині з послідовною обробкою процесів полягає в тому, що при моделюванні необхідно відслідковувати велику кількість процесів, які в реальному часі відбуваються паралельно. У зв'язку з цим алгоритми моделювання мають свої особливості:

- просування системи в часі, відстеження часової координати;
- забезпечення синхронної роботи об'єктів, з яких складається модельована система.

На сьогодні відомі чотири основні принципи регламентації подій:

1) принцип $\Delta t$ (постійний крок зміни часу);
2) принцип особливих станів;
3) принцип послідовного проведення заявок;
4) принцип паралельної роботи об'єктів (об'єктний принцип моделювання).

Розглянемо на прикладах, як реалізується в моделюючих алгоритмах кожний принцип окремо.

***Принцип*** $\Delta$t. Цей принцип полягає в тому, що алгоритмом моделювання імітується рух, тобто зміна стану системи, у фіксовані моменти часу: $t,\ t\ +\ \Delta t,\ t\ +\ 2\Delta t,\ t\ +\ 3\Delta t, \ldots$

Для цього заводиться лічильник часу (годинник), який, починаючи з нуля (початок моделювання), на кожному циклі збільшує своє значення $t$ на величину кроку в часі $\Delta t$. Таким чином, зміни системи відслідковуються такт за тактом у задані моменти: $t, t + \Delta t, t + 2\Delta t, t + 3\Delta t, \ldots$

Це найбільш універсальний із розглянутих принципів, оскільки застосовується для дуже широкого класу систем. Також він є найбільш простим у реалізації, оскільки принцип $\Delta t$ збігається з уявленням людини про час як про послідовне явище, що плине з постійним темпом.

Однак це самий неекономічний принцип, оскільки вся система аналізується моделюючим алгоритмом на кожному такті, навіть якщо в ній не відбувається жодних змін.

Інший недолік цього принципу полягає в тому, що часи подій округляються до величини $\Delta t$, що призводить до похибок у визначенні змінних, які характеризують систему.

*Принцип особливих станів*. Назвемо стан, у якому зазвичай перебуває система, звичайним. Такі стани інтересу не становлять, хоча займають більшу частину часу.

**Особливі стани** – це такі стани в ізольовані моменти часу, в яких характеристики системи змінюються стрибкоподібно. Для зміни стану системи потрібна певна причина, наприклад, прихід чергового вхідного сигналу. Зрозуміло, що з погляду моделювання інтерес становить саме зміна характеристик системи, тобто принцип вимагає від нас відслідковувати моменти переходу системи з одного особливого стану в інший, що приводить до істотної економії часу моделювання. Принцип особливих станів відрізняється від принципу $\Delta t$ тим, що включає в себе процедуру визначення моменту часу, що відповідає наступному особливому стану за відомих характеристик цього або попереднього станів.

*Принцип послідовного проведення заявок.* Цей принцип полягає в тому, що кожна заявка відслідковується від моменту надходження її в систему до моменту її виходу із системи. І тільки потім розглядається наступна заявка.

І, звичайно, треба пам'ятати, що чим більшим буде час моделювання, тим точнішим буде обчислений результат.

Слід нагадати про те, що необхідно спостерігати за поведінкою статистичної характеристики, якою, наприклад, є $P_{obs}$. Раніше ми зазначали, що статистична величина змінюється залежно від часу спостереження. Як тільки статистична величина перестане змінюватися в межах оголошення точності, тобто крива входить у коридор, відведений їй точністю, то це сигналізує про достатню кількість експериментів.

Необхідно уважно стежити, щоб усі шукані змінні ввійшли в інтервал оголошеної точності; тільки тоді можна припинити моделювання й бути впевненим у результаті.

Для підвищення ефективності алгоритму (зменшення часу його роботи) можна відкинути нехарактерну частину реалізації – зазвичай це початкова ділянка роботи системи, «вихід її на режим».

Зазначимо також, що не важливо, чи маємо ми справу з однією довгою реалізацією або з більшою кількістю коротких реалізацій (у яких зазвичай вирізана ділянка «вихід на режим»), у сумі, що дають реалізацію такої ж довжини, – статистичний результат буде тим самим. Це міркування встановлює рівність усереднень за ансамблем реалізацій усереднення за часом.

Слід зазначити, що на практиці зазвичай застосовують комбінації всіх трьох методів.

*Об'єктний принцип моделювання.* Як правило, алгоритми, спроектовані за першими трьома принципами, погано модернізуються. А виробнича ситуація часто змінюється й вимагає відповідних моделей для знаходження раціональних рішень у процесі керування.

Таким чином, виникла необхідність у прийомах моделювання, що забезпечують незалежність складання моделей елементів складної системи від зміни задачі або структури виробництва. Такий підхід до моделювання окремих об'єктів незалежно один від одного дозволяє складати як завгодно складні системи без зміни їх складових. Принцип об'єктного моделювання забезпечує модернізацію складних систем, подовжуючи життєвий цикл АСУ.

Ці моделі можуть складатися в будь-які конфігурації без зміни їх змісту.

Розглянемо імітаційну модель одноканальної СМО з обмеженим часом очікування заявок початку обслуговування.

Процес функціонування системи розглядається за період часу $[0, T]$. На вхід у випадкові моменти часу надходять заявки. Якщо в момент надходження чергової заявки канал вільний, то заявка приймається до негайного обслуговування, а якщо ні, то заявка стає в чергу на обслуговування. Час перебування заявки в черзі обмежений. Тривалість обслуговування заявок також є випадковою величиною. Показниками ефективності СМО можна вважати величини ймовірності того, що заявка, яка надійшла в систему, буде обслужена, або продуктивності розглянутої СМО. Суть підходу до імітації роботи СМО проста – випадковий процес роботи системи в інтервалі$[0, T]$ розглядається як детермінований, але з випадково вибраними параметрами. Для забезпечення статистичної стійкості шуканих результатів процес моделювання повторюється задану кількість разів $N_1$. Показники ефективності формуються на основі статистичної обробки результатів окремих реалізацій процесу моделювання.

Отже, розглядається одноканальна СМО, в яку надходять заявки, що утворюють ординарний потік однорідних подій із заданим законом розподілу $f(t)$ інтервалів між моментами появи заявок. Заявки надходять у випадкові моменти часу $t_j$; вони обслуговуються протягом часу $\tau_j^*$, і обслуговування закінчується в момент $t_j^*$. Випадкова величина тривалості обслуговування $\tau^*$ визначається відповідно до заданого закону розподілу $f(\tau^*)$. Частка наступної заявки залежить від моменту її появи $t_j$ відносно до моменту закінчення обслуговування попередньої $t_{j-1}^*$. Так, якби розглядалася СМО з відмовами й чергова заявка надійшла б раніше настання моменту закінчення обслуговування попередньої, тобто $t_j < t_{j-1}^*$, то така заявка одержала б відмову, якщо ж $t_j \geq t_{j-1}^*$, то заявка була б обслужена.

У зв'язку з тим, що використовується СМО, яке розглядається, має враховуватися обмеження часу можливого очікування, необхідно формувати $\tau_j$ – випадкову величину тривалості очікування $j$-тої заявки початку обслуговування з відповідним законом розподілу $f(\tau)$. Коли $t_j < t_{j-1}^*$, то $j$-та заявка одержує відмову, якщо $t_j' < t_{j-1}^*$, де $t_j' = t_j + \tau_j$ – момент закінчення очікування $j$-тої заявки початку обслуговування. Якщо ж $t_j' \geq t_{j-1}^*$, то обслуговування $j$-тої заявки починається в момент часу $t_j^n \geq t_{j-1}^*$.

Вихідні дані для моделювання: закони розподілу $f(t), f(\tau)$ і $f(\tau^*)$, величини $T$ і $N$, а також стан системи при $t = 0$, наприклад, $t_{j-1}^* = t_j = 0, m = n = N = 0$. Вважається, що заявки, для яких момент появи $t_j \geq T$, у систему не потрапляють і не обслуговуються. Заявки, для яких час закінчення обслуговування $t_j^* > T$, вважаються такими, що одержали відмову, а для яких $t_j^* \leq T$, є обслуженими.

Фрагмент блок-схеми моделюючого алгоритму одноканальної СМО наведений на рис. 6.4, а нижче поданий опис більшості використовуваних блоків:

– формування чергового моменту $t_j$ заявки, що подається в систему;

– перевірка умови $t_j < T$ приналежності моменту надходження чергової заявки $t_j$ інтервалу дослідження системи $[0, T]$;

– перевірка умови $t_j < t_{j-1}^*$, де $t_{j-1}^*$ – момент закінчення обслуговування попередньої заявки;

– формування випадкового значення тривалості очікування $\tau_j$ відповідно до закону розподілу $f(\tau)$;

– обчислення моменту $t_j' = t_j + \tau_j$ (у момент $t_j'$ заявка залишає систему, якщо вона не буде прийнята до обслуговування);

– перевірка умови $t_j' < t_{j-1}^*$ (заявка залишає систему раніше, ніж вивільниться канал);

– вибір як момент початку обслуговування $j$-тої заявки моменту закінчення обслуговування $(j-1)$-тої заявки $t_j^n = t_{j-1}^*$;

– вибір як момент початку обслуговування $j$-тої заявки моменту її надходження в систему $t_j^n = t_j$;

– формування тривалості обслуговування $\tau_j^*$ заявки (часу зайнятості каналу) відповідно до закону $f(\tau^*)$;

– обчислення моменту закінчення обслуговування $j$-тої заявки (моменту звільнення каналу) $t_j^* = t_j^n + \tau_j^*$;

– перевірка умови $t_j^* \leq T$;

– підрахування кількості обслужених заявок $m$;

– обчислення тривалості перебування в черзі $j$-ї заявки, $t_{ochj} = t_j^n + t_j$;

– підрахування кількості заявок $n$, що одержали відмову;

– підрахування кількості реалізацій $N$;

– перевірка умови $N < N1$, де $N1$ – кількість реалізацій, необхідна для забезпечення заданої точності.

Як д перетворити описану модель СМО у модель із відмовами, в якій заявка, що застала канал зайнятим, одержує відмову? Для цього потрібно в моделі (рис. 6.4) ліквідувати блоки *4 … 7* та із блока *3* за ознакою «Так» перейти в блок *14*.

Як перетворити описану модель СМО у модель із необмеженим очікуванням заявки початку обслуговування? Для цього потрібно в моделі ліквідувати блоки *4 … 6* і із блока *3* за ознакою «Так» перейти в блок *7*.

Нехай багатоканальна СМО має $K$ каналів, цілком ідентичних (з погляду тривалості обслуговування) і рівноправних щодо можливості використання для обслуговування цієї заявки будь-якого вільного каналу. У систему, як і раніше, надходить ординарний потік заявок із заданим законом розподілу $f(t)$. Час можливого очікування до початку обслуговування $\tau$ – випадкова величина із законом розподілу $f(\tau)$. Тривалість обслуговування $\tau^*$ має закон розподілу $f(\tau^*)$. Процес функціонування системи розглядається в інтервалі часу $[0, T]$.

Заявки приймаються до обслуговування по черзі (у порядку надходження в систему). Якщо для обслуговування цієї заявки є декілька вільних каналів, то канали залучаються до обслуговування по черзі (у першу чергу залучається той канал, який раніше за інших звільнився від обслуговування).

Описана система близька за характером до одноканальної СМО. Тому багато блоків можуть бути використані для побудови моделюючого алгоритму в цьому випадку. У моделі також зберігаються всі використані вище позначення. Блок-схема алгоритму багатоканальної СМО наведена на рис. 6.5. У ній додатково використані такі блоки.

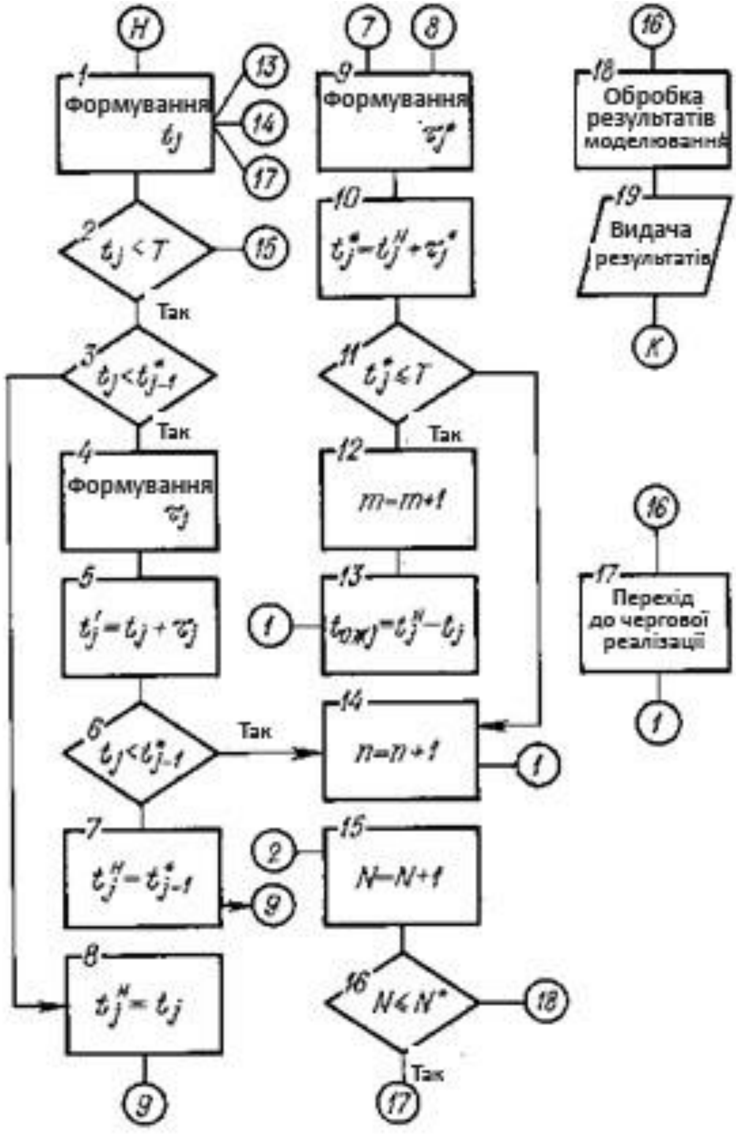


Рис. 6.4. Модель одноканальної СМО

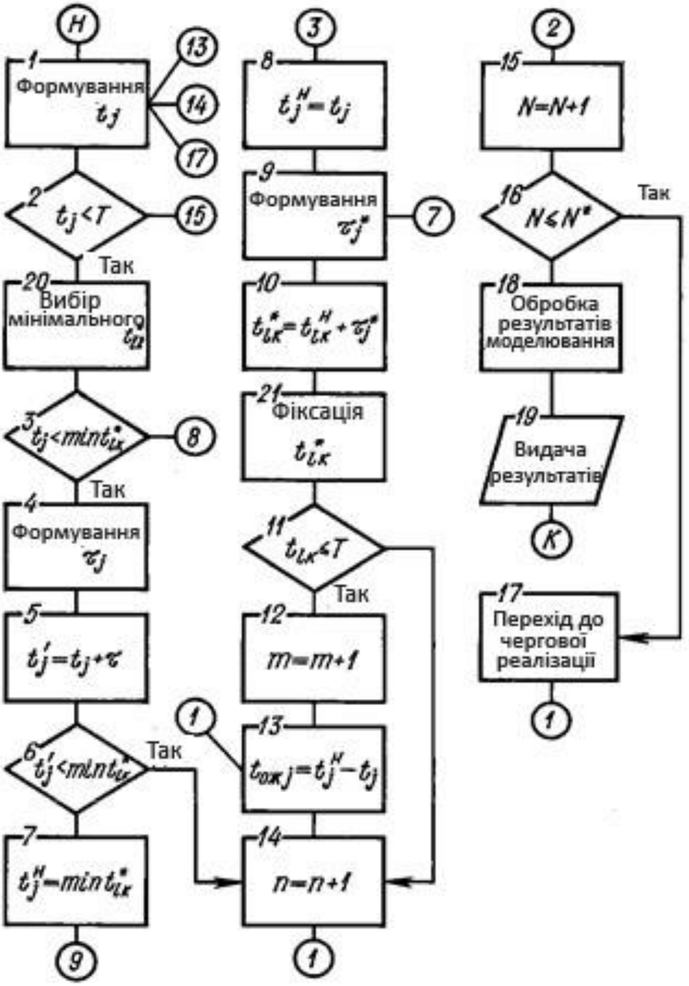


Рис. 6.5. Модель багатоканальної СМО

*3* – перевірка умови $t_j < \min t^*_{lk}$;

*6* – перевірка $t'_j < \min t^*_{lk}$;

*7* – вибір як моменту початку обслуговування $t^{н}_j$, величини $\min t^*_{lk}$;

*10* – обчислення $t^*_{lk}$ – моменту закінчення обслуговування $l$-тої заявки $k$-м каналом (моменту вивільнення $k$-го каналу, $k = \overline{1, K}$),

$$t_{lk}^{*} = t_{lk}^{\text{н}} + \tau_j^{*},$$

де $t_{lk}^{\text{н}}$ – момент початку обслуговування $l$-ї заявки $k$-м каналом.

## 6.5. Модифікації моделей систем масового обслуговування

Різновиди СМО, що зустрічаються на практиці, не вичерпуються розглянутими вище найпростішими моделями. Можна виділити особливості моделювання деяких інших класів СМО. Фрагменти алгоритмів, що наводяться нижче, слід розбирати разом із найпростішими моделями.

*Системи з обмеженням за довжиною черги*. У модельованій СМО може встановлюватися обмеження за довжиною черги. При цьому в момент надходження заявки необхідно перевіряти умову $l' < l$, де $l$ – кількість місць у черзі, $l'$ – кількість зайнятих місць у черзі в момент $t_j$. Якщо умова, що перевіряється, виконується, то заявка ставиться в чергу й чекає звільнення каналу, а якщо не виконується, то одержує відмову. Така перевірка здійснюється тільки в тому випадку, якщо всі канали зайняті. Якщо хоча б один канал вільний, це свідчить про те, що черги немає й перевірка не потрібна, а заявка йде відразу на обслуговування.

*Системи з обмеженнями за тривалістю обслуговування й тривалістю перебування.* Для забезпечення моделювання цих обмежень у модель уводяться й реалізуються відповідні перевірки. Так, при обмеженні за тривалістю обслуговування перевіряється умова $\tau_j^{*} < \tau_1$, де $\tau_j^{*}$ – розрахункова тривалість обслуговування $j$-тої заявки, а $\tau_1$ – гранична тривалість обслуговування заявок. При виконанні цієї умови обслуговування такої заявки триває, а при невиконанні обслуговування – переривається і заявка одержує відмову.

Для реалізації обмеження за тривалістю перебування заявки в СМО необхідно перевіряти умову $\tau_{jsist} \leq \tau_2$, де $\tau_{jsist} = t_j^{*} - t_j$ – тривалість перебування заявки в системі, а $\tau_2$ – гранична тривалість перебування заявок у СМО. При виконанні цієї умови обслуговування $j$-тої заявки триває, а при невиконанні заявка одержує відмову.

*Модель СМО із переважаючим обслуговуванням заявок, що раніше за інших залишають систему.* У розглянутих прикладах одно- і багатоканальної СМО заявки обслуговуються за надходженням. Дуже часто заявки, що мають право (по черзі) на негайне обслуговування, можуть перебувати в системі досить довго й обслуговуватися після заявок з малим $\tau$ і тими, що швидко йдуть із системи.

Розглянемо модифікацію алгоритму багатоканальної СМО (рис. 6.5). Блок-схема алгоритму, зображена на рис. 6.6, дозволяє ввести пріоритет для заявок, що раніше за раніше інших залишають систему.

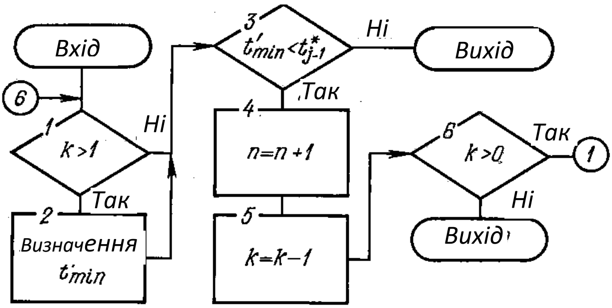


Рис. 6.6. Фрагмент моделі СМО із переважаючим обслуговуванням заявок

Якщо черга складається з $k$ заявок, то необхідно знайти таку, яка має найменший момент закінчення очікування $t'_{\min}$. Після визначення $t'_{\min}$ керування передається блоку *3*, що встановлює можливість обслуговування цієї заявки. При невиконанні умови керування передається операторові, який формує $t^n$ – момент початку обслуговування в основному алгоритмі. У разі відмови (тобто заявка обслужена не буде) обчислюється кількість відмов $n$, а черга зменшується на одиницю: $k - 1$. Якщо черга ще не вичерпана, тобто $k > 0$, то процедура повторюється, а інакше керування передається операторові основного алгоритму, якщо формує $t_{j+1}$ – момент надходження чергової заявки.

*Модель СМО з випадковим порядком обслуговування.* На рис. 6.7 зображено блок-схему моделюючого алгоритму.

Група операторів 1 … 7 вибирає заявки, які можуть бути обслужені, тобто для яких $t'_j > t^*_{j-1}$. Моменти часу $t'_j$ формують масив, з якого потім у блоці *8* здійснюється вибір заявки. Після перегляду всіх заявок, що стоять у черзі, блок *6* передає керування блоку *7*, і, якщо масив $t'_j$ не порожній ($a > 0$), то блок *8* вибирає заявки відповідно до заданого закону розподілу. Для цієї заявки блок *9* формує час початку обслуговування. Якщо регістр порожній (тобто жодна заявка із черги не може бути обслужена), то керування від блока *7* передається в основний алгоритм операторові, який формує час надходження чергової заявки.

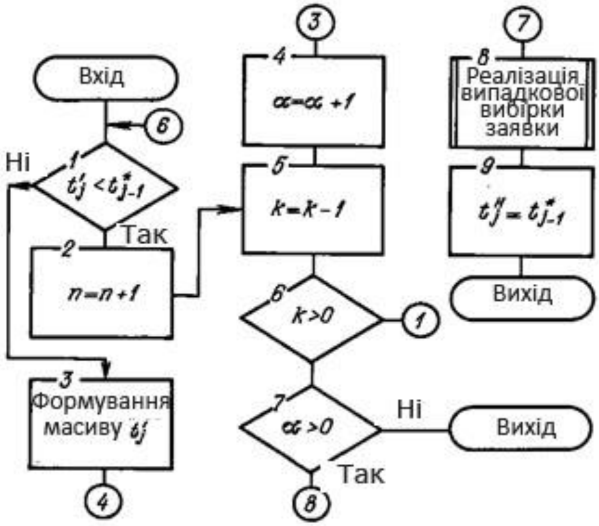


Рис. 6.7. Фрагмент моделі СМО з випадковим порядком обслуговування

Якщо відповідно змінити випадкову схему вибору заявки (блок *8*), наприклад, жорсткою детермінованою системою пріоритетів, то приходимо до чергової модифікації СМО.

*Модель СМО із заданим порядком вибору обслуговуючих каналів*. На рис. 6.8 зображена блок-схема моделюючого алгоритму.

Для кожної заявки, що надійшла в $t_j$ момент часу, проглядаються всі канали й визначаються вільні. Перевіряється умова $i < n$ ($i$ – поточне*,* $n$ – кількість каналів у СМО). Зайнятість каналу перевіряється блоком *3.* Незайняті канали сформують масив $t_i^{sv}$ (блок 4)*,* а їх кількість накопичується в лічильнику $n^* = n^* + 1$. Після того як вільні до моменту $t_j$ канали відібрані, керування передається блоку *7*, що реалізує правило вибору каналу. Якщо $n^* = 0$, тобто немає жодного вільного до моменту $t_j$ каналу, переходимо в основний алгоритм до блока, який формує нові значення $t_j$*,* і повторюємо відбір каналів для моменту $t_{j+1}$.

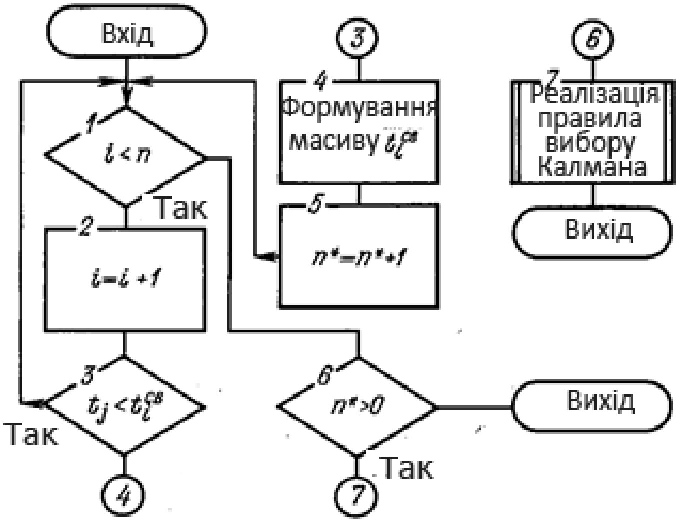


Рис. 6.8. Фрагмент моделі СМО із заданим порядком вибору обслуговуючих каналів

*Модель СМО з виходом із ладу каналів*. Особливістю цього класу СМО є можливість перебою в роботі каналу. Канал може бути відновлений за час $\tau^p$ і стати до ладу в момент $t = t^{sb} + \tau^p$, де $t^{sb}$ – момент перебою. Надійність каналу задається функцією розподілу інтервалу безвідмовної роботи $P(t)$, а $\tau^p$ визначається як випадкова величина з певним законом розподілу.

Перебій і ремонт каналу можна інтерпретувати як обслуговування фіктивної заявки з моментом початку обслуговування $t^{sb}$, тривалістю обслуговування $\tau^p$ і першим пріоритетом. Якщо заявка приходить у момент перебою, то вона одержує відмову або може бути обслужена після ремонту каналу.

Розглянемо цей клас СМО на прикладі одноканальної системи з такими властивостями: канал виходить із ладу в моменти $t_j^{sb}$, визначені законом $P(t)$; канал ремонтується випадковий час $\tau_j^p$, заданий законом розподілу $f(\tau^p)$; заявка, при обслуговуванні якої відбувся перебій, одержує відмову.

На рис. 6.9 зображено блок-схему моделюючого алгоритму. Блок *1* перевіряє умова $\gamma > 0$, тобто визначає, чи сформоване й занесене в масив значення $t^{sb}$. Якщо це так, то керування передається блоку *3*, а інакше в блоці *2* формується $t^{sb}$. Далі перевіряється умова $t^{sb} < t^{sv}$. Якщо вона не виконана, тобто перебій відбувся після закінчення обслуговування, то заявка може бути обслужена. У цьому випадку $t^{sb}$ записується в масив, а $\gamma = 1$.

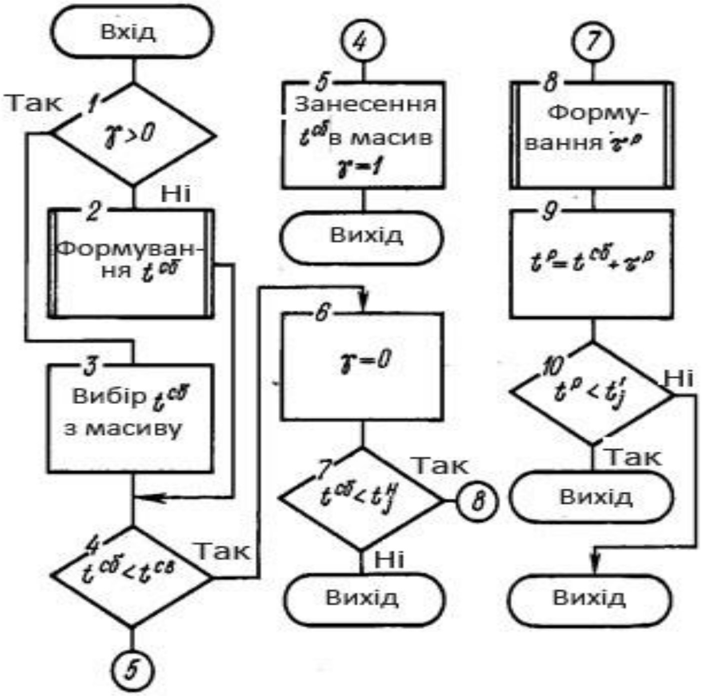


Рис. 6.9. Фрагмент моделі СМО з виходом із ладу каналів

Якщо ж $t^{sb} > t^{sv}$, керування передається блоку 6. Якщо $t^{sb} \leq t_j^n$ (перебій відбувся після початку обслуговування), то заявка одержує відмову й керування передається в основний алгоритм блока, що фіксує кількість відмов. Інакше, якщо перебій відбудеться до початку обслуговування, керування передається блоку *8*, де формується час ремонту, і далі блоку *9,* де обчислюється момент закінчення ремонту. Далі, якщо $t^p < t_j'$, тобто ремонт закінчиться раніше, ніж мине час очікування, керування передається в основний алгоритм блокам, які займаються обслуговуванням заявки, інакше – блоку, що підраховує відмови.

## 6.6. Приклад імітаційного моделювання системи масового обслуговування

Наведене вище свідчить, що при імітаційному моделюванні систему випробовують випадковими вхідними сигналами, підпорядкованими заданому статистичному закону, а результат приймають статистичні показники, усереднені за часом розгляду або за кількістю дослідів.

Фахівець із систем повинен добре розуміти ресурси продуктивності й ефективності проектованих ним систем, приховані в оптимізації параметрів, структур і дисциплінах обслуговування. Моделювання допомагає виявити ці приховані резерви.

При аналізі результатів моделювання важливо також указати інтереси й ступінь їх виконання. Розрізняють інтереси клієнта й інтереси власника системи.

Судити про результати роботи СМО можна за сукупністю значень показників. При аналізі результатів моделювання (показників) важливо також звертати увагу на інтереси клієнта й інтереси власника системи, тобто мінімізувати або максимізувати потрібно той або інший показник, а також ступінь їх виконання. Зазначимо, що найчастіше інтереси клієнта й власника між собою не збігаються або збігаються не завжди. Показники будемо позначати далі $H = \{h_1, h_2, \dots\}$.

Параметрами СМО можуть бути: інтенсивність потоку заявок; інтенсивність потоку обслуговування; середній час, протягом якого заявка готова очікувати обслуговування в черзі, кількість каналів обслуговування; дисципліна обслуговування і т. д. Параметри – це те, що впливає на показники системи. Параметри будемо позначати далі як $R = \{r_1, r_2, \dots\}$.

Розглянемо метод моделювання СМО на конкретному прикладі та план її дослідження. Припустимо, що для розглянутої СМО відома інтенсивність потоку заявок на обслуговування, яка в середньому становить 5 за одиницю часу.

Система масового обслуговування має два обслуговуючі пристрої, статистична продуктивність кожного з яких відома. Перший обслуговуючий пристрій в середньому обслуговує 1 заявку за одиницю часу, другий у середньому – 3 заявки за одиницю часу. При цьому є можливість запам'ятовувати чергу із двох заявок. Чергу будемо вважати загальною. Як тільки один з обслуговуючих пристроїв звільниться, то перша із черги заявка може почати обслуговуватися. У цьому випадку очевидно, що шляхи руху потоків заявок по СМО можна зобразити у вигляді еквівалентної схеми, а додавши значення й позначення характеристик кожного елемента СМО, одержати остаточно схему, зображену на рис. 6.10.

Застосуємо в нашому прикладі принцип послідовного проведення заявок. Його суть полягає в тому, що заявку проводять через усю систему від входу до виходу, і тільки після цього беруться за моделювання наступної заявки.

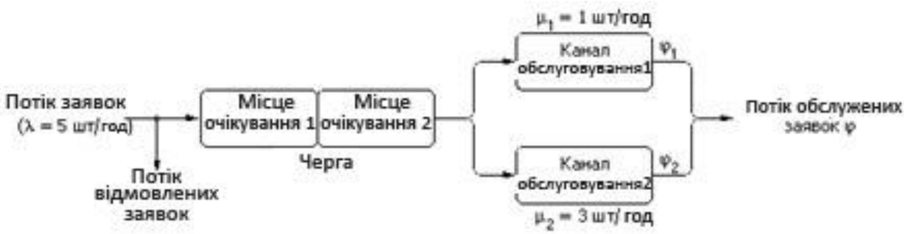


Рис. 6.10. Еквівалентна схема об'єкта моделювання

Для наочності побудуємо часову діаграму роботи СМО, відбиваючи на кожній лінійці (вісь часу) $t$ стан окремого елемента системи. Часових лінійок проводиться стільки, скільки є різних місць у СМО, потоків. У нашому прикладі їх 7 (потік заявок, потік очікування на першому місці в черзі, потік очікування на другому місці в черзі, потік обслуговування в каналі 1, потік обслуговування в каналі 2, потік обслужених системою заявок, потік заявок з відмовою).

Для генерації часу надходження заявок використаємо формулу обчислення інтервалу між моментами надходження двох випадкових подій:

$$t = \frac{1}{\lambda}\ln(r).$$

У цій формулі величина потоку λ повинна бути задана (до цього вона має бути визначена експериментально на об'єкті як статистичне середнє), $r$ – випадкове рівномірно розподілене число від 0 до 1 із ГВЧ, у якій випадкові числа потрібно брати підряд (не вибираючи спеціально).

**Задача.** Згенерувати потік із 10 випадкових подій з інтенсивністю появи подій 5 шт/год.

**Розв'язання задачі**. Візьмемо випадкові числа, рівномірно розподілені в інтервалі від 0 до 1, і обчислимо їхні натуральні логарифми (табл. 6.1).

Таблиця 6.1 – Фрагмент таблиці випадкових чисел та їх логарифмів

| $r_{\text{pp}}[0;\ 1]$ | $\ln(r_{\text{pp}}[0;\ 1])$ |
|---|---|
| 0,0333 | –3,4022 |
| . | –1,0337 |
| 0,2172 | –1,5269 |
| 0,5370 | –0,6218 |

Формула пуассонівського потоку визначає відстань між двома випадковими подіями в такий спосіб: $t =- \mathrm{Ln}(r_{\text{pp}})/\lambda$. Тоді, враховуючи, що $\lambda = 5$, маємо відстані між двома випадковими сусідніми подіями: 0,68; 0,21; 0,31; 0,12 години. Тобто події настають: перша – у момент часу $t = 0$, друга – у момент часу $t = 0{,}68$, третя – у момент часу $t = 0{,}89$, четверта – у момент часу $t = 1{,}20$, п'ята – у момент часу $t = 1{,}32$ і т. д. Події – надходження заявок відобразимо на першій лінійці (рис. 6.11).

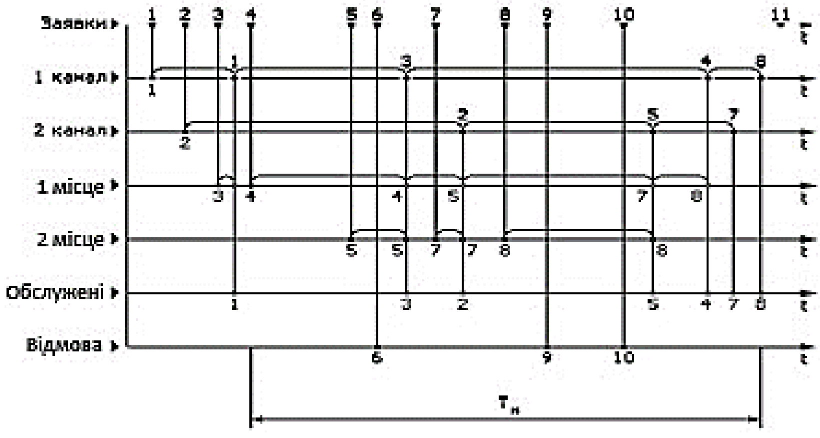


Рис. 6.11. Часова діаграма роботи СМО

Береться перша заявка й, оскільки в цей момент канали вільні, установлюється на обслуговування в перший канал. Заявка 1 переноситься на лінійку «1 канал».

Час обслуговування в каналі теж випадковий та обчислюється за аналогічною формулою:

$$t = -\frac{1}{\lambda}\ln(r),$$

де функцію інтенсивності виконує величина потоку обслуговування $\mu_1$ або $\mu_2$ залежно від того, який канал обслуговує заявку. Знаходимо на діаграмі момент закінчення обслуговування, відкладаючи згенерований час обслуговування моменту початку обслуговування, та опускаємо заявку на лінійку «Обслужені».

Заявка пройшла в СМО весь шлях. Тепер можна, відповідно до принципу послідовного проведення заявок, так само проімітувати шлях другої заявки.

Якщо в певний момент виявиться, що обидва канали зайняті, то слід установити заявку в чергу. На рис. 6.10 це заявка з номером 3. Зазначимо, що за умовами задачі в черзі, на відміну від каналів заявки, перебувають не випадковий час, а очікують, коли звільниться якийсь із каналів. Після звільнення каналу заявка піднімається на лінійку відповідного каналу, й там організовується її обслуговування.

Якщо всі місця в черзі в момент, коли надійде чергова заявка, будуть зайняті, то заявку слід відправити на лінійку «Відмовлені». На рис. 6.11 це заявка з номером 6.

Процедуру імітації обслуговування заявок продовжують певний час спостереження $T_s$. Чим більший цей час, тим точнішими надалі будуть результати моделювання. Реально для простих систем вибирають $T_s$, що дорівнює 50–100 і більше годин, хоча іноді краще виміряти цю величину кількістю розглянутих заявок.

Проведемо аналіз часової діаграми. Спочатку потрібно дочекатися режиму, що встановився. Відкидаємо перші чотири заявки як нехарактерні, що протікають під час процесу встановлення роботи системи. Вимірюємо час спостереження, припустимо, що в нашому прикладі він становить $T_s = 5$ одиниць часу. Підраховуємо з діаграми кількість обслужених заявок $N_{obs}$, години простою та інші величини. У результаті можемо обчислити показники, які характеризують якість роботи СМО.

1. Імовірність обслуговування: $P_{obs} = \frac{N_{obs}}{N} = \frac{5}{7} = 0{,}714$. Для розрахунку ймовірності обслуговування заявки в системі достатньо розділити кількість заявок, яким вдалося обслужитися за час $T_s$ (див. лінійку «Обслужені») $N_{obs}$, на кількість заявок $N$, що хотіли обслужитися за той самий час. Як і раніше ймовірність експериментально визначаємо відношенням подій, які відбулися, до загальної кількості подій, що могли відбутися!

2. Пропускна здатність системи: $A = \frac{N_{obs}}{T_s} = \frac{7}{5} = 1{,}4$ [шт/од. часу]. Для розрахунків пропускної здатності системи достатньо розділити кількість обслужених заявок $N_{obs}$ на час $T_s$, за який відбулося це обслуговування (див. лінійку «Обслужені»).

3. Імовірність відмови: $P_{vidm} = N_{vidm}/N = 3/7 = 0{,}43$. Для розрахунку ймовірності відмови заявці в обслуговуванні достатньо розділити кількість заявок $N_{vidm}$, яким відмовили за час $T_s$ (див. лінійку «Відмовлені»), на кількість заявок $N$, що хотіли обслужитися за той самий час, тобто надійшли в систему. Слід звернути увагу на те, що $P_{vidm} + P_{obs}$ в теорії повинне дорівнювати 1. Насправді експериментально вийшло, що $P_{vidm} + P_{obs}$ = 0,714 + 0,43 = 1,144. Ця неточність пояснюється тим, що час спостереження $T_s$ малий і накопичена кількість статистики не достатня для отримання точної відповіді. Похибка цього показника на даний момент становить 14 %!

4. Імовірність зайнятості одного каналу: $P_1 = T_{zan}/T_s$ = 0,05/5 = 0,01, де $T_{zan}$ – час зайнятості тільки одного каналу (першого або другого). Вимірюванням підлягають часові відрізки? в яких відбуваються певні події. Наприклад, на діаграмі шукають такі відрізки, під час яких зайняті або перший, або другий канал. У цьому прикладі є один такий відрізок в кінці діаграми завдовжки 0,05 одиниць часу. Частка цього відрізка в загальному часі розгляду ($T_s$=5 од. часу) визначається діленням і складає шукану імовірність зайнятості.

5. Імовірність зайнятості двох каналів: $P_2 = T_{zan}/T_s = 4{,}95/5 = 0{,}99$. На діаграмі шукаються такі відрізки, під час яких одночасно зайняті і перший, і другий канали. У цьому прикладі таких відрізків чотири, їх сума дорівнює 4,95 годин. Доля тривалості цих подій у загальному часі розгляду ($T_s$=5) визначається діленням і становить шукану імовірність зайнятості.

6. Середня кількість зайнятих каналів: $\overline{N}_k = 0 \cdot P_0 + 1 \cdot P_1 + 2 \cdot P_2 = 0{,}01 + 2 \cdot 0{,}99 = 1{,}99$. Для того щоб підрахувати, скільки каналів зайнято в системі в середньому, достатньо знати частку (імовірності зайнятості одного каналу) і помножити на вагу цієї частки (один канал), знати частку (імовірність зайнятості двох каналів) і помножити на вагу цієї частки (два канали) і т. д. Отримана цифра 1,99 свідчить про те, що з можливих двох каналів у середньому завантажено 1,99 каналу. Це високий показник завантаження, а саме 99,5 %, і система добре використовує цей ресурс.

7. Імовірність простою хоча б одного каналу:

$$P_1^* = T_{prost}\frac{1}{T_s} = \frac{0{,}05}{5} = 0{,}01.$$

8. Імовірність простою двох каналів одночасно: $P_2^* = T_{prost}.2/T_s = 0$.

9. Імовірність простою всієї системи: $P_s^* = T_{prost}.s/T_s = 0$.

10. Середня кількість заявок у черзі: $\overline{N}_z = 0 \cdot P_{0z} + 1 \cdot P_{1z} + 2 \cdot P_{2z} = 0{,}34 + 2 \cdot 0{,}64 = 1{,}62$ [шт]. Щоб визначити середню кількість заявок у черзі, потрібно визначити окремо ймовірність того, що в черзі буде одна заявка $P_{1z}$; імовірність того, що в черзі буде стояти дві заявки $P_{2z}$ і т. д., і знову з відповідними вагами їх скласти.

11. Імовірність того, що в черзі буде одна заявка: $P_{1z} = T_{1z}/T_s = 1{,}7/5 = 0{,}34$ (усього на діаграмі чотири таких відрізки, які у сумі дають 1,7 годин).

12. Імовірність того, що у черзі буде стояти одночасно дві заявки: $P_{2z} = T_{2z}/T_s = 3{,}2/5 = 0{,}64$ (усього на діаграмі три таких відрізки, які у сумі дають 3,25 годин).

13. Середній час очікування заявки в черзі

$$\overline{T}_{och} = \frac{\sum_{i=1}^{N} T_{oc.i}}{N}.$$

(Скласти всі часові інтервали, протягом яких будь-яка заявка перебувала в черзі, і розділити на кількість заявок). На часовій діаграмі таких заявок 4.

14. Середній час обслуговування заявки:

$$\overline{T}_{obs} = \frac{\sum_{i=1}^{N} T_{obs.i}}{N}.$$

(Скласти всі часові інтервали, протягом яких будь-яка заявка перебувала на обслуговуванні в будь-якому каналі, і розділити на кількість заявок).

15. Середній час знаходження заявки в системі $\overline{T}_{sist} = \overline{T}_{och} + \overline{T}_{obs}$.

16. Середня кількість заявок у системі

$$\overline{N} = \frac{\sum_{i=1}^{K}(N_{2i}+N_{3i}+N_{4i}+N_{5i})}{K}.$$

Розіб'ємо інтервал спостереження, наприклад, на десятихвилинки. Вийде на п'ятьох годинах $K$ підінтервалів (у нашому випадку $K = 30$). У кожному підінтервалі визначимо за часовою діаграмою, скільки заявок у цей момент перебуває в системі. Дивитися потрібно на 2, 3, 4 і 5-ти лінійки – які з них зайняті в цей момент. Потім суму $K$ доданків усереднити.

Далі слід оцінити точність кожного з отриманих результатів. Тобто відповісти на запитання: наскільки ми можемо довіряти цим значенням?

Якщо точність не є задовільною, то слід збільшити час експерименту й тим самим поліпшити статистику. Можна зробити й по-іншому. Знову декілька разів запустити експеримент на час $T_s$. А у висновку усереднити значення цих експериментів. І знову перевірити результати на критерій точності. Цю процедуру слід повторювати доти, доки не буде досягнута необхідна точність.

Далі слід скласти таблицю результатів і оцінити значення кожного з них із погляду клієнта й власника СМО та, враховуючи сказане в кожному пункті, слід зробити загальний висновок.

## 6.7. Синтез систем масового обслуговування

Ми пробили аналіз існуючої системи. Це дало можливість побачити її недоліки й визначити напрями поліпшення її якості. Але залишаються незрозумілими відповіді на конкретні запитання, що саме треба робити – збільшувати кількість каналів або збільшувати їх пропускну здатність, або збільшувати кількість місць у черзі, і, якщо збільшувати, то наскільки? Є й такі питання, що краще – створити три канали із продуктивністю 5 шт/год або один із продуктивністю 15 шт/год?

Щоб оцінити чутливість кожного показника до зміни значення певного параметра, роблять так. Фіксують усі параметри крім одного, обраного. Потім знімають значення всіх показників при декількох значеннях цього обраного параметра. Звичайно, доводиться повторювати знову й знову процедуру імітації й усереднювати показники при кожному значенні параметра, оцінювати точність. Але в результаті виходять надійні статистичні залежності характеристик (показників) від параметра.

Наприклад, для 12-ти показників нашого прикладу можна одержати 12 залежностей від одного параметра: залежність імовірності відмов $P_{vidm}$ від кількості місць у черзі (КМЧ), залежність пропускної здатності *A* від кількості місць у черзі і т. д. (рис. 6.12).

Потім так само можна зняти ще 12 залежностей показників $P$ від іншого параметра $R$, зафіксувавши інші параметри, і т. д. Утворюється своєрідна матриця залежностей показників $P$ від параметрів $R$, за якою можна провести додатковий аналіз про перспективи руху (поліпшення показників) у ту або іншу сторону. Нахил кривих добре показує чутливість, ефект від руху за певним показником. У математиці цю матрицю називають якобіаном $J$, у якій функцію нахилу кривих виконують значення похідних $P_i/\Delta R_j$.

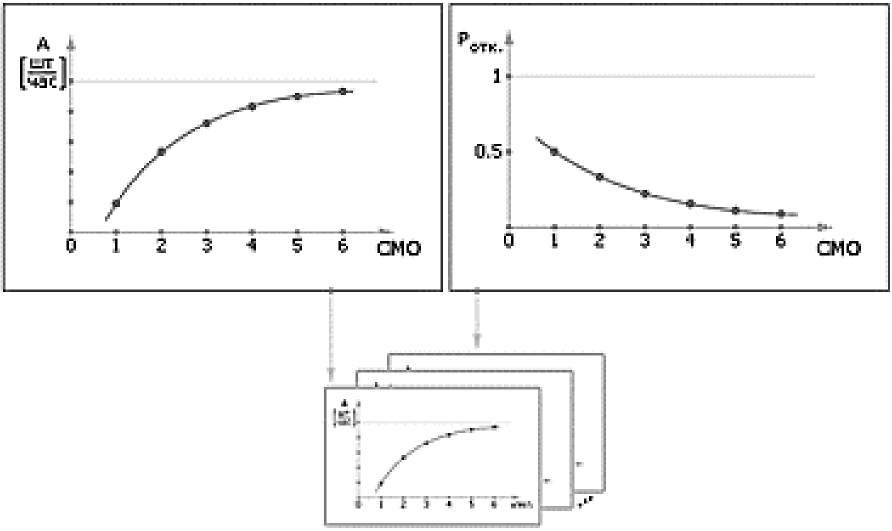


Рис. 6.12. Приблизний вигляд залежностей показників від параметрів СМО

Якщо показників 12, а параметрів, наприклад, 5, то матриця має розмірність 12 x 5. Кожний елемент матриці – крива, залежність $i$-го показника від $j$-го параметра. Кожна точка кривої – середнє значення показника на досить значному відрізку $T_s$ або усереднене за декількома експериментами.

Слід розуміти, що криві знімалися в припущенні того, що всі параметри крім одного в процесі їх зняття були не змінні. (Якби всі параметри змінювали значення, то криві були б іншими. Але так не роблять, оскільки вийде повна плутанина й залежностей не буде видно).

Тому, якщо на основі розгляду знятих кривих приймається рішення про те, що деякий параметр буде в СМО змінений, то всі криві для нової точки, в якій знову буде досліджуватися питання про те, який параметр необхідно змінити, щоб поліпшити показники, потрібно знімати заново.

Так, крок за кроком можна спробувати поліпшити якість системи. Але поки що ця методика не може відповісти на низку запитань. Справа в тому, що, по-перше, якщо криві монотонно зростають, то виникає запитання, де все ж таки слід зупинитися. По-друге, можуть виникати протиріччя; один показник може поліпшуватися при зміні вибраного параметра, у той час як інший буде одночасно погіршуватися. По-третє, ряд параметрів складно виразити числово, наприклад, зміна дисципліни обслуговування, зміна напрямків потоків, зміна топології СМО. Пошук рішення у двох останніх випадках проводиться із застосуванням методів експертизи й методами штучного інтелекту.

Тому зараз розглянемо тільки перше запитання. Як прийняти рішення? Яким повинне бути все-таки значення параметра, якщо з його зростанням показник увесь час монотонно поліпшується? Навряд чи значення нескінченності влаштує інженера.

Параметр $R$ – керування, це те, що перебуває в розпорядженні власника СМО. Змінюючи керування, можна впливати на значення показника $P$, ціль, критерій (імовірність відмов, пропускну здатність, середній час обслуговування і т. д.). З рис. 6.13 видно, що, якщо збільшувати керування $R$, то можна добитися завжди поліпшення показника $P$. Але очевидно, що будь-яке керування пов'язане з витратами $Z$. І чим більше докладають зусиль для керування, чим більше значення керуючого параметра, тим більші витрати. Зазвичай витрати на керування зростають лінійно $Z = C_1 \cdot R$. Хоча бувають випадки, коли, наприклад, в ієрархічних системах вони зростають експоненційно, іноді – зворотно експоненційно (знижки за опт) і т. д.

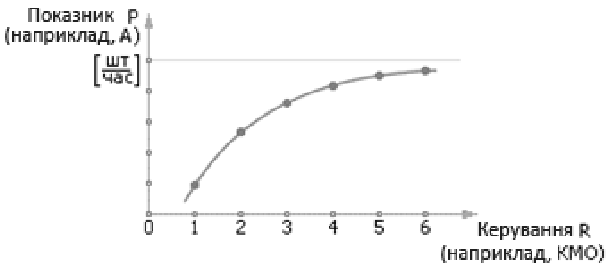


Рис. 6.13. Залежність показника *P* від керованого параметра *R* (приклад)

У будь-якому випадку зрозуміло, що колись вкладення дедалі нових витрат просто перестане себе окупати. Іншими словами, показник $P$ у складних системах не може зростати нескінченно. Рано або пізно його зростання уповільнюється, а витрати $Z$ зростуть (рис. 6.14).

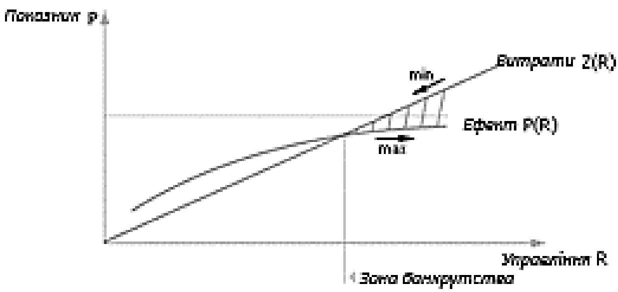


Рис. 6.14. Залежності ефекту від застосування показника $P$
і витрат $Z$ на його одержання як функції керованого параметра $R$

З рис. 6.14 видно, що при призначенні ціни $C_1$ за одиницю витрат $R$ і ціни $C_2$ за одиницю показника $P$, ці криві можна скласти. Криві складають, якщо їх потрібно одночасно мінімізувати або максимізувати. Якщо одна крива підлягає максимізації, а інша – , мінімізації, то слід знайти їх різницю, наприклад, за точками. Тоді результуюча крива (рис. 6.15), що враховує й ефект від керування, й витрати на це, буде мати екстремум. Значення параметра $R$, що доставляє екстремум функції, і є розв'язком задачі синтезу.

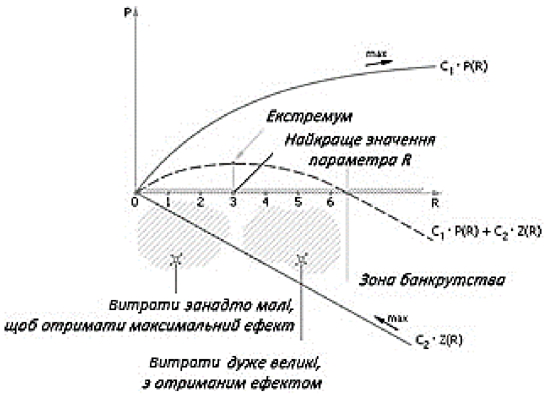


Рис. 6.15. Сумарна залежність ефекту від застосування показника $P$ і витрат $Z$ на його одержання як функції керованого параметра $R$

Крім керування $R$ і показника $P$ у системах діє збурення. Збурення позначимо як $D = \{d_1, d_2, \ldots\}$. Збурення – це вхідний вплив, який, на відміну від керуючого параметра, не залежить від волі власника системи. Наприклад, низькі температури на вулиці, конкуренція, на жаль, знижують потік клієнтів, несправність обладнання прикро знижують продуктивність системи. Керувати цими величинами безпосередньо власник системи не може. Зазвичай збурювання діє «назло» власникові, знижуючи ефект $P$ від керуючих зусиль $R$. Це відбувається тому, що у загальному випадку система створюється для досягнення цілей, недосяжних самих по собі в природі. Людина, організовуючи систему, завжди сподівається за допомогою неї досягти певної мети $P$. На це вона витрачає зусилля $R$, ідучи всупереч природі. Система — організація доступних людині, вивчених нею природних компонентів для досягнення певної нової мети, недосяжної раніше іншими способами.

Отже, якщо ми знімемо залежність показника $P$ від керування $R$ ще раз (як показано на рис. 6.13), але в умовах збурювання $D$, що з'явилося, можливо, характер кривої зміниться. Швидше за все, показник буде при однакових значеннях керувань перебувати нижче, оскільки збурювання має «протилежний» характер, знижуючи показники системи (рис. 6.16). Система, надана сама собі, без зусиль керуючого характеру перестає забезпечувати мету, для досягнення якої вона була створена. Якщо, як і раніше, побудувати залежність витрат, співвіднести її із залежністю показника від параметра керування, то знайдена точка екстремуму зміститься (рис. 6.17) порівняно з випадком «збурювання = 0» (див. рис. 6.15).

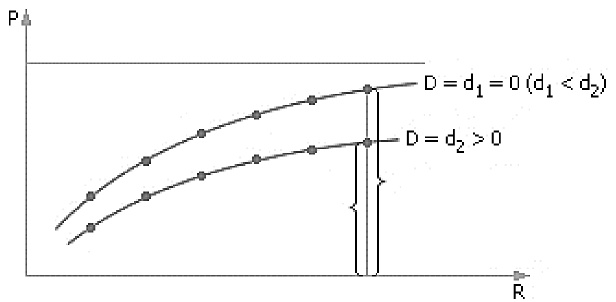


Рис. 6.16. Залежність показника $P$ від керуючого параметра $R$ при різних значеннях діючих на систему збурювань $D$

Якщо знову збільшити збурювання, то криві зміняться (див. рис. 6.16) і, як наслідок, знову зміниться положення точки екстремуму (рис. 6.17).

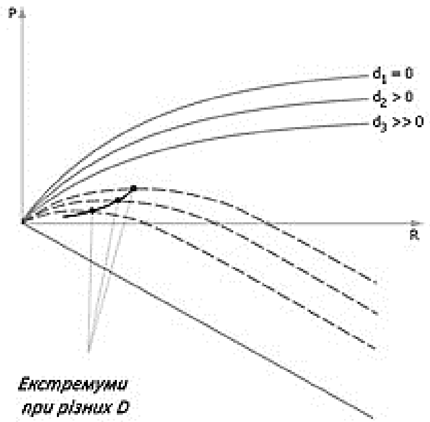


Рис. 6.17. Знаходження точки екстремуму на сумарній залежності при різних значеннях діючого збурного фактора D

В остаточному підсумку, усі знайдені положення точок екстремуму переносяться на новий графік, де утворюють залежність показника *P* від керуючого параметра *R* при зміні збурювань *D* (рис. 6.18).

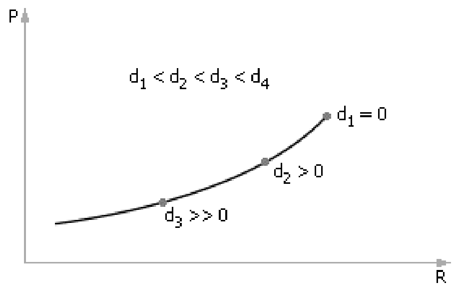


Рис. 6.18. Залежність показника *P* від керуючого параметра *R* при зміні значень збурювань D (крива складається тільки із точок екстремумів)

Слід звернути увагу на те, що насправді на цьому графіку можуть бути й інші робочі точки (графік пронизаний як би сім'ями кривих), але нанесені нами точки задають такі координати керуючого параметра, при яких при заданих збуруваннях (!) досягається найбільше з можливих значення показника $P$.

Цей графік (див. рис. 6.18) зв'язує показник $P$, керування (ресурс) $R$ і збурювання $D$ у складних системах, указуючи на те, як діяти щонайкраще ОПР (особі, що приймає рішення) в умовах виниклих збурювань. Тепер користувач, знаючи реальну обстановку на об'єкті (значення збурювання), може швидко за графіком визначити, який керуючий вплив на об'єкт необхідний, щоб забезпечити найкраще значення його показника, що цікавить.

Варто звернути увагу на те, що якщо керуючий вплив буде меншим за оптимальний, то сумарний ефект знизиться, виникне ситуація недоотриманого прибутку. Якщо керуючий вплив буде більшим за оптимальний, то ефект також знизиться, оскільки заплатити за чергове збільшення керуючих зусиль потрібно буде за величиною більшою, ніж та, яку ми одержимо в результаті її використання.

## Контрольні запитання

1. В чому полягає суть імітаційного моделювання?
2. Назвіть основні етапи дослідження систем при імітаційному моделюванні.
3. Сутність методу імітаційного моделювання та сфери його застосування.
4. Що є теорією масового обслуговування?
5. Перерахуйте компоненти структурної схеми СМО.
6. Назвіть класифікаційні ознаки СМО.
7. Моделі систем масового обслуговування.
8. Показники ефективності роботи СМО.
9. У чому полягає суть форми подання моделюючого алгоритму у вигляді детальної схеми?
10. В чому полягає суть форми подання моделюючого алгоритму у вигляді логічної схеми?
11. В чому полягає суть форми подання моделюючого алгоритму у вигляді схеми програми?

# 7. АНАЛІТИЧНЕ МОДЕЛЮВАННЯ СИСТЕМ

## 7.1. Аналітичні моделі

*Аналітична модель* – це математичний опис структури та процесу функціонування системи, а також методика визначення показників її ефективності. Така модель дозволяє швидко і з високою точністю характеризувати поведінку системи.

*Найбільш ефективними і наочними є аналітичні моделі при описі функціонування СМО*. СМО зазвичай розуміють як динамічну систему, призначену для ефективного обслуговування випадкового потоку заявок (вимог на обслуговування) при обмеженнях на ресурси системи. Іншими словами, можна сказати, що особливістю масового обслуговування є наявність потоку однорідних (ідентичних) подій (вимог, заявок), що зазнають обслуговування. Крім того, для повного опису СМО і постановки завдання дослідження необхідно визначити структуру системи та дисципліну (правила) обслуговування, а також показники якості обслуговування, тобто певні числові показники, за значенням яких можна було б судити про якість функціонування досліджуваної СМО. Розв’язанням перерахованих задач займається теорія масового обслуговування, що є розділом теорії ймовірностей. Наведено декілька задач і прикладів, які належать до проблематики теорії масового обслуговування.

1. У комп’ютер, що використовується для керування технологічним процесом, час від часу надходять сигнали від датчиків, пов’язаних із керованим об’єктом. Кожен сигнал потребує обробки протягом певного випадкового часу (залежить від змісту сигналу). Таким чином, роботу комп’ютера можна розглядати як операцію масового обслуговування, що складається з елементарних складових – обробки окремих сигналів. Потрібно розв’язати задачу: чи здатний комп’ютер із цим обсягом пам’яті і швидкодією впоратися з обробкою всіх сигналів, що надходять?

2. У вантажний морський порт, який має декілька причалів, прибувають судна на розвантаження-навантаження у випадкові моменти часу. Час розвантаження-навантаження (обслуговування) суден коливається в певному інтервалі залежно від типу судна і вантажу. Якщо всі причали зайняті, то прибуле судно стає на

рейді та чекає своєї черги. Кожен корабель має розклад прибуття, але через непередбачені обставини він, як правило, порушується. Тому вважається, що прибуття судна є випадковою подією з якимось математичним очікуванням, дисперсією і т. д.

Запитання: чому дорівнює середня тривалість розвантаження-навантаження?

3. У складальному цеху здійснюється складання виробів різних видів. У випадку нестачі хоча б одного виду деталей виробництво зупиняється. Надлишкові деталі надходять у бункери певної місткості. На процес надходження деталей, як і на час складання виробу, впливають випадкові чинники. Яка ймовірність простою виробничої лінії? Чому дорівнює ймовірність переповнення бункерів? Елементарною операцією в цьому випадку є складання одного виробу з готового комплекту деталей.

Шуканими характеристиками СМО вважаються показники ефективності, які розуміють як кількісні показники, що частково або повністю характеризують рівень виконання системою масового обслуговування покладених на неї функцій. Наприклад, це такі характеристики: імовірність обслуговування заявки (це ймовірність того, що заявка, яка надішла в СМО, буде обслужена), імовірність відмови в обслуговуванні, середній час очікування заявки початку обслуговування, середній час перебування заявки в СМО, середня довжина черги тощо.

На перший погляд перераховані характеристики свідчать тільки про факт їх обчислення. Але насправді це далеко не так.

Нехай обчислена величина ймовірності обслуговування заявки дорівнює 0,65. Це свідчить про те, що зі 100 заявок, які надійшли на обслуговування, тільки 65 з них буде обслужено. А це, у свою чергу, означає, що 35 заявок отримає відмову, тобто третина з тих, які надійшли. Кожна третя заявка (кожен третій клієнт) отримає відмову в обслуговуванні. А це втрачений прибуток, втрачений клієнт, втрачена можливість розширення сфери послуг.

Або інша характеристика: середня кількість зайнятих каналів. Нехай при кількості обслуговуючих каналів, яка дорівнює трьом, середня кількість зайнятих дорівнює 1,75. Це свідчить про те, що принаймні один канал постійно простоює і, можливо, потрібно задуматися про доцільність його існування. Або протилежний випадок: при трьох каналах середня кількість зайнятих дорівнює 2,98. Це може свідчити про те, що ймовірність відмови в обслуговуванні велика (знову втрати потенційних клієнтів) і, можливо, необхідно задуматися про збільшення кількості обслуговуючих каналів або про збільшення інтенсивності обслуговування заявок каналами (наприклад, купівлю нового обладнання).

Основне завдання, що стоїть перед теорією масового обслуговування, полягає у встановленні залежності між характером потоку заявок, кількістю каналів, їх продуктивністю, правилами роботи СМО й ефективністю обслуговування. Розв’язання задач аналізу СМО пов’язане із залученням різних методів оптимізації: лінійного або нелінійного програмування, динамічного програмування, теорії ігор, методу імітаційного моделювання і т. д.

Математичний аналіз роботи СМО значно спрощується, якщо випадковий процес, що протікає в системі, є марківським. У цьому випадку роботу СМО досить просто вдається описати за допомогою диференціальних рівнянь, а в граничному випадку – лінійних алгебраїчних, і відобразити в явному вигляді основні характеристики ефективності обслуговування через параметри СМО та потоку заявок.

Як характеристики ефективності обслуговування можуть застосовуватися різні величини і функції, наприклад:

- середня кількість заявок, що обслуговуються за одиницю часу;
- середній відсоток заявок, які отримують відмову і залишають систему не обслуженими;
- імовірність того, що заявка, що надійшла, негайно буде прийнята до обслуговування;
- середній час очікування в черзі;
- закон розподілу часу очікування;
- середня кількість заявок, що знаходяться в черзі, і т. д.

## 7.2. Багатоканальна система масового обслуговування з відмовами

Система масового обслуговування, описувана схемою $M|M|n|m = 0$, є найпростішою з усіх СМО – це багатоканальна система з відмовами. Звичайно, одноканальна СМО виду $M|M|1|m = 0$, будучи окремим випадком – багатоканальною, має більш прості характеристики, але вони можуть бути отримані з характеристик багатоканальної СМО при $n = 1$.

У систему надходить найпростіший, тобто пуассонівський, потік заявок на обслуговування з інтенсивністю $\lambda$, де $\lambda$ – кількість заявок, що надходять на обслуговування за одиницю часу. Інтервали між моментами надходження заявок передбачаються випадковими і розподіленими за показовим законом $f(t) = \lambda e^{-\lambda t}, t > 0$.

Заявка, що застала канали зайнятими, залишає систему необслуженою. Обслуговування заявки продовжується протягом випадкового часу, розподіленого за показовим законом з інтенсивністю обслуговування $T_{obs}$

$$\mu = {}^{1}/_{T_{obs}}\,;\ f(\tau) = \mu e^{\mu\tau}, \tau > 0.$$

З цього розподілу часу обслуговування випливає, що потік обслуговування є найпростішим.

Основними характеристиками такої СМО є:

- абсолютна пропускна здатність СМО – $A$;
- відносна пропускна здатність СМО – $q$;
- ймовірність відмови – $P_{vid}$.

Для знаходження цих характеристик СМО розглядається як фізична система $S$, що знаходиться по черзі в одному зі $n + 1$ станів, а саме:

$S_0$ – всі канали вільні;

$S_1$ – зайнятий один канал, інші $n - 1$ вільні;

...

$S_k$ – зайняті $k$ каналів, інші $n - k$ вільні;

...

$S_n$ – зайняті всі $n$ каналів.

Розмічений граф станів досліджуваної системи наведений на рис. 7.1.

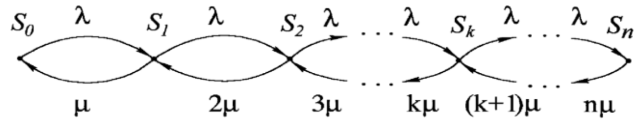


Рис. 7.1. Граф станів багатоканальної СМО з відмовами

Усього в графі $n + 1$ станів. Стрілки направо переводять систему зі стану $S_j$ в інший стан, який визначається початком обслуговування заявки, що надійшла черговим $(j + 1)$-м каналом. При цьому щільність перехідної ймовірності з одного стану в інший (подальший) визначається інтенсивністю λ надходження заявок у систему, тобто кожна заявка, що надходить, переводить СМО зі стану $S_j$ в стан $S_j + 1$ з інтенсивністю λ.

Як розмічаються переходи системи справа наліво? Якщо система знаходиться в стані $S_1$, то, як тільки закінчиться обслуговування заявки, що займає цей канал, система перейде в стан $S_0$. Цей перехід буде визначатися інтенсивністю μ обслуговування каналів. Якщо система знаходиться в стані $S_2$, тобто в системі обслуговуванням заявок зайнято два канали, то інтенсивність переходу з $S_2$ у $S_1$ збільшується вдвічі (2μ). Очевидно, якщо обслуговуванням зайнято $k$ каналів, то інтенсивність цього обслуговування буде в $k$ разів вища, ніж при обслуговуванні одним каналом ($k$μ).

Отже, процес, який протікає в цій СМО, є окремим випадком процесу «розмноження і загибелі», що дозволяє записати вирази для розрахунку ймовірностей станів

$$P = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu^2 2!} + \frac{\lambda^3}{\mu^3 3!} + \cdots \frac{\lambda^n}{\mu^n n!}\right)^{-1},$$

$$P_1 = \frac{\lambda}{\mu} P_0, P_2 = \frac{\lambda^2}{\mu^2 2!} P_0, P_3 = \frac{\lambda^3}{\mu^3 3!} P_0, \ldots, P_n = \frac{\lambda^n}{\mu^n n!} P_0.$$

Вводячи величину $\rho = \lambda/\mu$, яку називають зведеною інтенсивністю потоку заявок, тобто середню кількість заявок, що надходять в СМО за середній час обслуговування однієї заявки, можна отримати формули для розрахунку граничних ймовірностей станів (при) $k\ = \overline{1,n}$:

$$P_0 = \left(1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \frac{\rho^3}{3!} + \cdots \frac{\rho^n}{n!}\right)^{-1}.$$

Або в більш компактному записі:

$$P_0 = \left(\sum_{\alpha=0}^{n} \frac{\rho^n}{\alpha!}\right)^{-1}, \quad P_k = \frac{\rho^k}{k!} P_0, k = \overline{1,n}. \tag{7.1}$$

Формули (7.1) називаються формулами Ерланга. Вони виражають граничні ймовірності всіх станів системи залежно від параметрів $n, \lambda$ та μ.

Користуючись ними, можна знайти характеристики ефективності СМО: $P_{vid}$, $q$ і $A$. Обчислення $P_{vid}$ зводиться до визначення ймовірності зайнятості всіх $n$ каналів системи, тобто до визначення ймовірності $n$-го стану:

$$P_{vid} = P_n = \frac{\rho^n}{n!} P_0. \tag{7.2}$$

Відносна пропускна здатність $q$ – це ймовірність того, що заявка буде прийнята до обслуговування. Оскільки обслуговування і відмова – це повна група несумісних подій, то

$$q = 1 - P_{vid} = 1 - P_n\ . \tag{7.3}$$

Абсолютна пропускна здатність $A$ – це середня кількість заявок, обслужених системою за одиницю часу:

$$A = \lambda q = \lambda(1 - P_n). \quad (7.4)$$

Крім перерахованих характеристик (7.2)–(7.4), можна визначити ще цілий ряд важливих величин. Середня кількість зайнятих каналів $k$ може бути обчислена як математичне сподівання $M(k)$ випадкової величини $k$. Дискретна випадкова величина $k = \{0, 1, 2, \dots, n\}$ (кількість зайнятих каналів) настає з імовірністю $P_i$.

Тоді математичне сподівання випадкової величини

$$M(k) = 0 \cdot P_0 + 1 \cdot P_1 + 2 \cdot P_2 + 3 \cdot P_3 + \cdots + k \cdot P_k + \dots + n \cdot P_n$$

і, якщо вже відомі всі, $P_i(i = 0, n)$, то $k$ легко обчислюється.

Але середню кількість зайнятих каналів $\bar{k}$ можна виразити і через абсолютну пропускну здатність $A$, де $A$ є середньою кількістю заявок, що обслуговуються за одиницю часу. Один зайнятий канал обслуговує в середньому за одиницю часу μ заявок, а середня кількість зайнятих каналів

$$\bar{k} = \frac{A}{\mu} = \rho(1 - P_n). \quad (7.5)$$

**Приклад 1.** Задана чотириканальна СМО з відмовами ($M \mid M \mid 4 \mid 0$), $\lambda = 2\ (\text{с}^{-1})$, $\mu = 1(\text{с}^{-1})$. Обчислити $P_i(i = 0, 1, 2, 3, 4)$, $P_{vidm}$, $q$, $A, \bar{k}$.

Для розв'язання цієї задачі (і надалі в подальших прикладах) необхідно послідовно реалізовувати чотири етапи:

1) побудова графа станів;
2) розмітка графа;
3) обчислення $P_i$;
4) обчислення показників ефективності (у цій задачі – $P_{vidm}, q, A, \bar{k}$).

Для правильної побудови графа станів необхідно перерахувати можливі стани:

$S_0$ – всі канали вільні;

$S_1$ – 1 канал зайнятий, інші 3 вільні;

$S_2$ – 2 канали зайняті, інші 2 вільні;

$S_3$ – 3 канали зайняті, 1 вільний;

$S_4$ – всі канали зайняті.

Розмічений граф, у якому $n + 1 = 5$ станів, показаний на рис. 7.2. Це свідчить про те, що вже реалізовано перші два етапи.

Наведена інтенсивність обслуговування $\rho = \frac{\lambda}{\mu} = 2$.

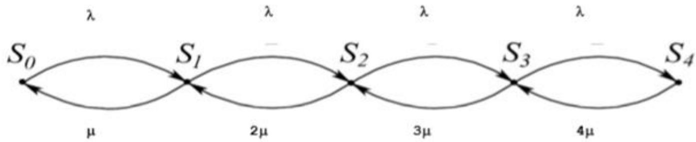


Рис. 7.2. Граф станів системи прикладу 7.1

$$P_0 = \frac{1}{7},\ P_1 = \frac{2}{7},\ P_2 = \frac{2}{7},\qquad P_3 = \frac{4}{21}, P_4 = \frac{2}{21}, P_{vidm} = P_4 = 0{,}095,$$
$$q = 0{,}905, A = \lambda q = 1{,}81, \bar{k} = \rho(1 - P_4) = 1{,}81.$$

Аналіз результатів показує, що приблизно 90 % заявок буде обслуговуватися і, відповідно, тільки 10 % отримають відмову. Це виходить за рахунок порівняно високої інтенсивності обслуговування заявок μ.

Унаслідок цього в середньому обслуговуванням заявок буде зайнято всього приблизно 2 канали, $\bar{k} = 1{,}81$, а відповідно, 2 канали – вільні. Ці результати наводять на думку про те, що, можливо, вигідно було б скоротити кількість обслуговуючих каналів. Отже, виникає питання про оптимізацію структури системи (кількості обслуговуючих каналів).

У виразах (7.2)...(7.5) основним параметром є наведена інтенсивність обслуговування $\rho = \lambda/\mu$. Тому будь-яка СМО розглянутого типу з $n$ каналами працює однаково добре (або однаково погано) при будь-яких λ і μ, для яких незмінною є величина ρ. Наприклад, якщо СМО з якоюсь кількістю каналів і з продуктивністю $\mu = 5$ зазнає впливу потоку заявок інтенсивності $\lambda = 5$, то така ж СМО з продуктивністю $\mu = 0{,}1$, що взаємодіє з потоком $\lambda = 0{,}1$, має ті самі величини $P_0, P_k, P_{vidm}, q, \bar{k}$, та тільки $A$ змінюється пропорційно λ.

*Як залежать показники ефективності СМО від n при* $\rho = \text{const}$?

При збільшенні $n$ ймовірності $P_0$ і $P_k$, при $k = \overline{1, n}$, зменшуються, оскільки
для, $n = 1$ $P_0 = (1 + \rho)^{-1}$,
для, $n = 2$ $P_0 = (1 + \rho + \rho^2/2!)^{-1}$,
для $n = 3$ $P_0 = (1 + \rho + \rho^2/2! + \rho^3/3!)^{-1}$ і т. д.,
а $P_k$ відповідно до (7.1) вони пропорційні $P_0$. Тобто, при збільшенні кількості станів $n$ системи $S$, здійснюється перерозподіл імовірностей $n$ станів при загальному обмеженні $P_0 + \sum_{k=1}^{n} P_k = 1$.

Таким чином, зростання $n$ при $\rho =$ const приводить до поліпшення всіх характеристик СМО. А саме при $n \to \infty$: $P_{vidm} \to 0, q \to 1, A \to \lambda,\ k \to \rho$.

При заданні верхніх меж для $P_{vidm}$ або $q$ можна вибрати конкретне значення $n$. Нехай, наприклад, $\rho = 1$ і $P_{vidm} \le\ 0{,}03$.. Тоді $P_0/n! \le\ 0{,}03$. Звідси знаходимо $n!\ /P_0 \geq 33$, тобто

$$!\left(1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}\right) > 33.$$

Остання нерівність починає виконуватися при $n\ =\ 4$.

*Як залежать показники ефективності СМО від* ρ *при заданому* $n$*?*

Зі збільшенням ρ заявки надходять в середньому частіше, ніж здійснюється їх обслуговування, зменшується ймовірність $P_0$, сума ймовірностей $P_k$ $(k\ = 1, n)$ відповідно збільшується, оскільки $n$

$$P_0+\sum_{k=1}^{n} P_k\ = 1.$$

Таким чином, при зростанні ρ значення $P_{vidm} \to\ 0, q \to 0,\ \bar{k} \to\ n.$

Тому необхідно зменшувати ρ за рахунок збільшення μ для поліпшення показників ефективності СМО, яка розглядається.

Використовуючи розмічений граф станів, наведений на рис. 7.3, і методики, описані в попередніх розрділах, можна отримати характеристики СМО описуваною схемою $M \mid M \mid 1 \mid 0$.

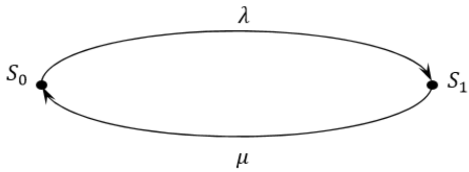


Рис. 7.3. Граф станів одноканальної СМО

Система диференціальних рівнянь Колмогорова може бути записана як:

$$\frac{dP_0(t)}{dt} = \mu P_1(t) - \lambda P_0(t), \frac{dP_1(t)}{dt} = \lambda P_0(t) - \mu P_1(t), P_0(t) + P_1(t) = 1.$$

Цю систему можна розв'язати при таких початкових умовах: у момент часу $t = 0$ система перебувала в стані $S_0$, тобто в початковий момент часу канал вільний, тоді $P_0(0) = 1$, $P_1(0) = 0$.

Виражаючи одну змінну через іншу, з системи рівнянь Колмогорова можна отримати диференціальне рівняння для визначення ймовірності $P_0(t)$ стану $S_0$

$$\frac{dP_0\,(t)}{dt} = -(\lambda + \mu)P_0\,(t) + \mu.$$

Це лінійне рівняння першого порядку. З огляду на початкову умову $P_0(t) = 1$, можна отримати його розв'язок у такому вигляді:

$$P_0(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\mu+\lambda)t}.$$

Тоді

$$P_1(t) = 1 - P_0(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\mu+\lambda)t}.$$

Графіки залежності $P_0(t)$ і $P_1(t)$ від $t$ наведено на рис. 7.4 для різних співвідношень λ і μ: а) для $\lambda > \mu$, б) для $\lambda < \mu$, в) для $\lambda = \mu$. Крива $P_0(t)$визначає ймовірність, яка зменшується, через те що канал вільний: при $t = 0$ $P_0 = 1$, , і при $t \to \infty$

$$P_0(t) \to \frac{\mu}{\lambda+\mu}. \tag{7.6}$$

Крива $P_1(t)$ доповнює криву $P_0(t)$ до одиниці, причому $P_1(0) = 0$ і при $t \to \infty$.

$$P_1(t) \to \frac{\lambda}{\lambda+\mu}. \tag{7.7}$$

Співвідношення (7.6) і (7.7) є граничними можливостями станів системи $M \mid M \mid 1 \mid 0$.

Для розглянутої одноканальної СМО з відмовами показники ефективності розраховуються за такими виразами:

$$q = P_0 = \frac{\mu}{\lambda+\mu},\;\; A = \lambda q,\;\; P_{0tk} = 1 - q = \frac{\lambda}{\lambda+\mu}. \tag{7.8}$$

Абсолютна пропускна здатність *A* має розмірність $1/s$ і визначає інтенсивність потоку, що утворився на вході каналу обслуговування з урахуванням частоти надходження в канал заявок та інтенсивності їх обслуговування μ.

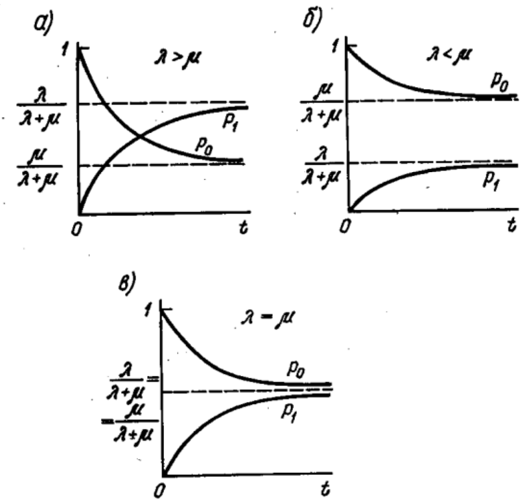


Рис. 7.4. Залежності $P_0(t)$ і $P_1(t)$

Якщо порівняти формули (7.1)–(7.4) із формулами (7.6)–(7.8), то можна побачити, що при $n = 1$ перші вирази зводяться до других, що природно, оскільки СМО виду $M \mid M \mid 1 \mid 0$ є окремим випадком системи $M \mid M \mid n \mid 0$.

**Приклад 2.** Система обслуговує кожну заявку протягом 30 хвилин. Заявки приходять в середньому 1 раз за 40 хвилин, якщо система зайнята (вона одноканального виду $M \mid M \mid 1 \mid 0$)). Визначити характеристики цієї СМО.

Інтенсивність надходження заявок в систему

$$\lambda = 60 \text{ хв} / 40 \text{ хв} = 1{,}5.$$

Інтенсивність обслуговування заявок

$$M = 60 \text{ хв} / 30 \text{ хв} = 2.$$

Імовірність того, що заявка, що надійшла в СМО, буде обслужена

$$P_0 = \frac{\mu}{\lambda + \mu} = \frac{2}{3{,}5} = 0{,}57,$$

тобто 57 % заявок буде обслуговуватися. А ймовірність відмови $P_1 = 1 - P_0 = 1 - 0{,}57 = 0{,}43$, тобто 43 % заявок отримають відмову.

Можна ще визначити частку часу роботи СМО, яку називають коефіцієнтом завантаження системи

$$k_{zag} = \frac{A}{\mu} = \frac{\lambda}{\mu} P_0 = 0{,}429.$$

На закінчення цього підрозділу проаналізуємо вираз (7.8). Як взаємодіє СМО з потоком заявок? Так, при $\lambda >> \mu$ $P_{vidm} \to 1$, $q \to 0$. При $\lambda << \mu$ $P_{vidm} \to 0$, $q \to 1, A \to \lambda$. Очевидно, що не можна знайти будь-які бажані співвідношення між $\lambda$ і $\mu$ для оптимізації показників ефективності даної СМО, але вибрати вигідні значення $\mu$ при будь-якому $\lambda = \text{const}$ можна, задаючись якими-небудь межами збільшення $A$ і $q$ чи зменшення $P_{vidm}$. Наприклад, для отримання $P_{vidm} \leq 0{,}1$ необхідно, щоб $\mu \geq 9\lambda$, а для $q \geq 0{,}8 - \mu \geq 4\lambda$ і т. д.

### 7.3. Модель одноканальної системи масового обслуговування з чергою

Розглянемо найпростішу задачу: досліджувана СМО одноканальна ($n = 1$) з обмеженою кількістю $m$ місць у черзі; потік заявок – стаціонарний пуассонівський з інтенсивністю $\lambda$, час (тривалість) обслуговування розподілено за показовим законом із параметром $\mu$. Заявка, що надійшла в момент, коли канал зайнятий, стає в чергу, якщо її довжина (кількість заявок у черзі) не перевищує $m$. В іншому випадку заявка отримує відмову і покидає систему необслуженою, тобто розглядається СМО виду $M \mid M \mid 1 \mid m$.

Система в довільний момент часу може виявитися в одному із станів: $S_0$ (канал вільний, черги немає), $S_1$ (канал зайнятий, черги немає), $S_2$ (канал зайнятий, у черзі одна заявка), …, $S_k$ (канал зайнятий, $k - 1$ заявка в черзі), …, $S_{m+1}$ (канал зайнятий, $m$ заявок у черзі).

Розмічений граф станів системи наведено на рис. 7.5. У графі всього $m + 2$ стани. Система переходить зі стану $S_i$ у стан $S_{i+1}$під впливом потоку заявок з інтенсивністю $\lambda$, і, навпаки, зі стану $S_{i+1}$ вона переходить у стан $S_i$ під впливом потоку обслужених заявок (потоку обслуговування) з інтенсивністю $\mu$.

Як знайти граничні ймовірності станів системи, тобто ймовірності станів для установленого (стаціонарного) режиму функціонування СМО? У цьому випадку, як зазначено у рівняннях Колмогорова $dP_i/dt = 0$, оскільки $P_i$ для всіх станів є постійними величинами.

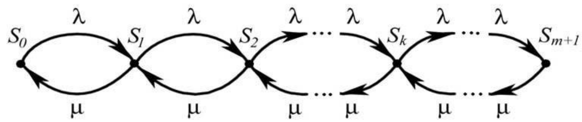


Рис. 7.5. Граф станів одноканальної СМО з чергою

Можна нагадати правило складання цих рівнянь:

– для всіх потоків (для всіх стрілок розміченого графа станів) інтенсивність потоку множиться на ймовірність того стану, з якого виходить потік (стрілка);

– у лівій частині рівності підсумовуються добутки, відповідні вхідним потокам;

– у правій частині рівності підсумовуються добутки вихідних потоків.

Або, користуючись загальним розв’язком для схеми «розмноження і загибелі», можна записати вирази для граничних ймовірностей станів

$$P_0 = (1 + \rho + \rho^2 + \rho^3 + \cdots + \rho^{m+1})^{-1}, P_1 = \rho P_0, P_2 = \rho^2 P_0, \ldots\ldots,$$

$$P_{m+1} = \rho^{m+1} P_0.$$

Або в більш компактному записі

$$P_0 = \left(\sum_{\alpha=0}^{m+1} \rho^{\alpha}\right)^{-1}, \qquad P_k = \rho^k P_0, \quad k = \overline{1, m+1}. \qquad (7.9)$$

У виразі для $P_0$ формул (7.9) підсумовуються члени геометричної прогресії $1, \rho, \rho^2, \ldots, \rho^{m+1}$. Використовуючи формулу суми членів прогресії, можна записати:

$$P_0 = \frac{1-\rho}{1-\rho^{m+2}}.$$

Або, що те саме

$$P_0 = (1-\rho)(1-\rho^{m+2})^{-1},$$

тоді

$$P_k = \rho^k (1-\rho)(1-\rho^{m+2})^{-1}. \qquad (7.10)$$

Вирази для $P_0$ і $P_k$ правдиві для всіх ρ, відмітних від 1, якщо ж $\rho = 1$, то для всіх $i$:

$$P_i = \frac{1}{m+2}\,. \tag{7.11}$$

Отже, відбувається перехід до одноканальної СМО з відмовами. Тепер можна $m = 0$ послідовно визначити основні характеристики ефективності СМО: $P_{vidm}$, $q, A, r$ – середня кількість заявок, що знаходяться в черзі, $k$ – середня кількість заявок, пов'язаних із системою (що стоять у черзі та обслуговуються), $t_{och}$ – середній час очікування заявки в черзі і $t_{sist}$ – середній час перебування заявки в системі.

Імовірність відмови в обслуговуванні (канал зайнятий, і всі місця в черзі зайняті):

$$P_{vidm} = P_{m+1} = \rho^{m+1} P_0 = \rho^{m+1}(1-\rho)(\,1-\rho^{m+2})^{-1}, \tag{7.12}$$

оскільки $P_{m+1}$ – імовірність того, що канал зайнятий і всі $m$ місць у черзі зайняті.

Відносна пропускна здатність системи (імовірність того, що заявка, яка надійшла в СМО, буде обслужена)

$$q = 1 - P_{vidm} = \rho^{m+1} P_0. \tag{7.13}$$

Абсолютна пропускна здатність (кількість заявок, що обслуговуються за одиницю часу):

$$A\; = \lambda q. \tag{7.14}$$

Середня кількість заявок, що знаходяться в черзі, – $\bar{r}$. Для визначення цієї величини необхідно розглянути випадкову величину $R$ – кількість заявок у черзі, закон розподілу якої такий: з імовірністю $P_2$ у черзі знаходиться одна заявка ($R = 1$), з імовірністю $P_3$ – дві заявки ($R = 2$), з імовірністю $P_{m+1}$ – $m$ заявок ($R =\; m$). Тоді математичне очікування цієї випадкової величини дорівнює середній кількості заявок, що знаходяться в черзі

$$\bar{r} = M(R) = 1 \cdot P_2 + 2 \cdot P_3 + \ldots + m\,P_{m+1}\; =$$

$$=\; \rho^2\, P_0\,(1 +\; 2\rho + \ldots + (k\; - 1)\rho^{k-2} + \ldots + m\rho^{m+1}).$$

В отриманому виразі в дужках знаходиться сума деяких доданків. Потрібно вивести формулу для цієї суми, яка являє собою похідну за ρ таким виразом:

$$\sum_{k=1}^{m} \rho^k = \rho + \rho^2 + \ldots + \rho^k + \ldots + \rho^m,$$

а вже для отриманого виразу можна використовувати формулу суми геометричної прогресії

$$\sum_{k=1}^{m} \rho^k = \frac{\rho - \rho^{m+1}}{1 - \rho}.$$

Диференціюючи за ρ праву і ліву частини цієї рівності, отримаємо

$$1 + 2\rho + \ldots + (k - 1)\rho^{k-2} + \ldots + m\rho^{m-1} = \frac{1 - \rho^m(m + 1 - m\rho)}{(1 - \rho)^2}.$$

У результаті вираз для середньої кількості заявок, що знаходяться в черзі, записується в такому вигляді:

$$\bar{r} = \frac{\rho^2[1 - \rho^m(m+1-m\rho)]}{(1-\rho^{m+2})(1-\rho)}. \quad (7.15)$$

Вираз (7.15) правдивий для $\rho \neq 1$, якщо ж $\rho = 1$, то, використовуючи (7.11) і формулу для суми ряду, можна отримати

$$\bar{r} = \frac{m(m+1)}{2(m+2)}. \quad (7.16)$$

Позначимо середню кількість заявок, пов'язаних із системою (що стоять у черзі і обслуговуються) $\bar{k}$. Розглянемо випадкову величину, де $K = R + \Omega$, де $K$ – кількість заявок, пов'язаних із системою; $R$ – кількість заявок у черзі, $\Omega$ – кількість заявок, що знаходяться на обслуговуванні. Враховуючи те, що математичне очікування від суми випадкових величин є сумою математичних очікувань, отримаємо

$$\overline{k} = M[K] = M[R] + M[\Omega] = \bar{r} + \overline{\omega},$$

де $\bar{r}$ обчислюється за формулою (7.15), а вираз для $\overline{\omega}$ необхідно вивести.

У досліджуваній системі використовується один канал, тому випадкова величина $\Omega$ набуває лише два значення: 0 – з імовірністю

$$P_0 = (1 - \rho)(1 - \rho^{m+2})^{-1},$$

коли канал вільний, і 1 – з імовірністю $1 - P_0$, коли канал зайнятий.

Тому

$$\overline{\omega} = M[\Omega] = 0 \cdot P_0 + 1 \cdot (1 - P_0) = (\rho - \rho^{m+2})(1 - \rho^{m+2})^{-1}.$$

Середній час очікування заявки в черзі $\bar{t}_{och}$. Якщо врахувати, що середній час обслуговування однієї заявки дорівнює $1/\mu$, то, розглядаючи випадкову величину $T_{och}$ – час очікування в черзі, можна визначити закон розподілу цієї випадкової величини: з імовірністю $P_1$ заявка прийде в систему під час обслуговування попередньої заявки, але перед нею не буде черги, і вона буде чекати початку обслуговування протягом часу $T_{och} = 1/\mu$; з імовірністю $P_2$ в черзі перед заявкою, яка розглядається, стоятиме ще одна, і час очікування$T_{och} = 2/\mu$; з імовірністю $P_k$ заявка, що надійшла, застане в системі $k$ заявок і буде чекати $T_{och} = k/\mu$ одиниць часу. Тому

$$\bar{t}_{och} = M[T_{och}] = P_1 \frac{1}{\mu} + P_2 \frac{2}{\mu} + \cdots + P_m \frac{m}{\mu}.$$

.

Підставляючи в отриманий вираз $P_k(k = \overline{1,m})$ і перетворюючи його, отримаємо

$$\bar{t}_{och} = M[T_{och}] = \frac{\rho[1-\rho^m(m+1-m\rho)]}{\mu(1-\rho^{m+2})(1-\rho)}. \tag{7.17}$$

Вираз (7.17) з урахуванням (7.15) можна записати у спрощеному вигляді

$$\bar{t}_{och} = \frac{\bar{r}}{\rho\mu} = \frac{\bar{r}}{\lambda}, \tag{7.18}$$

тобто середній час очікування заявки початку обслуговування дорівнює середній кількості заявок у черзі, поділеній на інтенсивність вхідного потоку заявок.

Середній час перебування заявки в системі $\bar{t}_{sist}$. Для визначення $\bar{t}_{sist}$ уведемо випадкову величину $T_{sist}$ – час перебування заявки в СМО

$$T_{sist} = T_{och} + \theta,$$

де $T_{och}$ – час очікування заявки в черзі; $\theta$ – випадкова величина часу обслуговування. Вона дорівнює часу обслуговування $T_{obsl}$, якщо заявка обслуговується, і нулю, якщо вона не обслуговується (отримує відмову). Таким чином,

$$\bar{t}_{sist} = M[T_{sist}] = M[T_{och}] + M[\theta] = \bar{t}_{och} + q\bar{t}_{obsl} = \bar{t}_{och} + q/\mu,$$

або інакше, з урахуванням виразу (7.17):

$$\bar{t}_{sist} = \frac{\bar{r}}{\lambda} + \frac{q}{\mu}.$$

**Приклад 3.** У СМО потік заявок є найпростішим з інтенсивністю $\lambda = 2$ заявки за хвилину. Працює один обслуговуючий прилад з інтенсивністю обслуговування $\mu = 2$ заявки за хвилину. Визначити характеристики СМО за умови, що довжина черги на вході системи обмежена і $m = 5$. Вихідні дані: розглядається СМО $\lambda = \mu = 2$, $\rho = 1$. Граф системи містить $m + 2 = 7$ станів.

*Розв'язання.* згідно з виразом для $\rho = 1$:

$$P_0 = P_1 = P_2 = \cdots = P_6 = \frac{1}{7} \approx 0{,}143.$$

Імовірність відмови в обслуговуванні – $P_{vidm} = P_6$ =0,143.

Імовірність обслуговування $= 1 - P_{vidm} = 0{,}857$.

Продуктивність системи $= \lambda q = 2 \cdot 0{,}857 = 1{,}7$ заявок / хв.

Інші характеристики

$$\bar{r} = \frac{m(m+1)}{2(m+2)} = \frac{5 \cdot 6}{2 \cdot 7} \approx 2{,}1 \text{ заявок},$$

$$\bar{t}_{och} = \frac{\bar{r}}{\lambda} = \frac{2{,}1}{2} = 1{,}05 \text{ хв},$$

$$\bar{t}_{obsl} = \frac{q}{\mu} = \frac{0{,}857}{2} = 0{,}428 \text{ хв},$$

$$\bar{t}_{sist} = \frac{\bar{r}}{\lambda} + \frac{q}{\mu} = 1{,}05 + 0{,}428 \approx 1{,}5 \text{ хв}.$$

**Приклад 4.** Як і в попередньому прикладі 3, розглядається СМО виду $M \mid M \mid 1 \mid 5$, $\lambda = 6$, заявок / хв, $\bar{t}_{obs} = 20$ с. Визначити характеристики СМО.

Інтенсивність обслуговування $\mu = \frac{1}{\bar{t}_{obs}} = \frac{60}{20} = 3$ заявки / хв.

Наведена інтенсивність обслуговування $\rho = \frac{\lambda}{\mu} = \frac{6}{3} = 2.$

Імовірність нульового стану – це частка часу простою обслуговуючого каналу $P_0 = (1 - \rho)/(1 - \rho^{m+2}) \approx 0{,}008.$

Імовірність відмови в обслуговуванні

$$P_{\text{мшвь}} = \rho^{m+1} P_0 = 2^6 \cdot 0{,}008 = 0{,}51.$$

Імовірність обслуговування

$$q = 1 - P_{otk} = 0{,}49.$$

Продуктивність СМО

$$A = \lambda q = 6 \cdot 0{,}49 = 2{,}94 \text{ заявки /хв}.$$

Середня кількість заявок в черзі

$$\bar{r} = \frac{\rho^2[1 - \rho^m(m+1-m\rho)]}{(1-\rho^{m+2})(1-\rho)} = \frac{2^2[1-2^5(5+1-5\cdot 2)]}{(1-2^7)(1-2)} = 4{,}03.$$

Середня тривалість очікування заявки в черзі

$$\bar{t}_{och} = \frac{\bar{r}}{\lambda} = \frac{4.03}{6} = 0.67 \text{ хв.}$$

Середня тривалість обслуговування

$$\bar{t}_{obsl} = \frac{q}{\mu} = \frac{0{,}49}{2} = 0{,}245 \text{ хв.}$$

Загальний час перебування заявки в системі

$$\bar{t}_{sist} = \frac{\bar{r}}{\lambda} + \frac{q}{\mu} = 0{,}67 \ + \ 0{,}245 \ \approx \ 0{,}91 \text{ хв.}$$

Аналіз показників ефективності двох СМО у цьому прикладі і попередньому показує, що системи відрізняються тільки інтенсивністю вхідного потоку заявок: у другій МО $\lambda$ в три рази більше, ніж у першій, а це означає, що наведена інтенсивність обслуговування $\rho$ стала в другому випадку в два рази більшою. Ця обставина в другій СМО збільшує $P_{vidm}$ приблизно в 3,6 раза, а $q$ збільшує майже у два рази. У другій СМО половина заявок йдуть із системи не обслуженими ($P_{vidm} = 0{,}51$). Ці обставини вимагають кардинальної зміни характеристик другої СМО: збільшення кількості обслуговуючих каналів або збільшення інтенсивності обслуговування, або зміни того чи іншого, або зняття обмеження на кількість місць у черзі.

Підтвердженням того, що необхідні деякі структурні зміни в СМО прикладу 4, служить елементарний розрахунок: нехай СМО працює за добу 12 годин, прибуток системи від обслуговування однієї заявки – 100 одиниць, $\mu = \ 3$ заявки / хв, $P_{otk} = 0{,}51$.

Тоді за 12 годин буде втрачено

$$12 \cdot 60 \cdot \mu \cdot P_{otk} = 720 \cdot \ 3 \cdot \ 0{,}51 \approx 102 \ \text{ заявки.}$$

Втрата виручки становитиме 1102·100 = 110 тис. одиниць, що, звичайно, небажано.

Тепер можна розглянути роботу одноканальної СМО ($n = 1$) з очікуванням для випадку, коли кількість місць у черзі не обмежена ($m = \ \infty$). Процес функціонування такої системи описується шляхом граничного переходу від системи $M \mid M \mid 1 \mid m$ при ($m = \infty$), тобто буде розглядатися система $M \mid M \mid 1 \mid \infty$

(рис. 7.6). Цей граничний режим може існувати тільки при_$\rho < 1$ $(\mu > \lambda)$, оскільки в першій формулі (7.18) наведена сума нескінченної кількості членів геометричної прогресії. При $\rho > 1$ черга зростає до нескінченності.

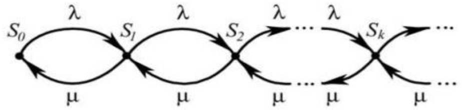


Рис. 7.6. Граф станів одноканальної СМО з необмеженою чергою

Тому надалі вважають, що $\rho < 1$, тоді з формул (7.18) при $m \to \infty$ отримують, що сума прогресії

$$1 + \rho + \rho^2 + \cdots + \rho^k + \cdots = \frac{1}{1-\rho},$$

звідки

$$P_0 = 1 - \rho,\ P_1 = \rho(1 - \rho), P_2 = \rho^2(1 - \rho), \ldots,\ P_k = \rho^k(1 - \rho).$$

Імовірності $P_0, P_1, P_2, \ldots, P_k$ утворюють геометричну прогресію зі знаменником $\rho$. Максимальна з них $P_0$ – імовірність того, що канал буде взагалі вільний. Це свідчить про те, що якою б не була навантаженою система з чергою, якщо вона тільки взагалі справляється з потоком заявок $(\rho < 1)$, найімовірніша кількість заявок у системі дорівнюватиме нулю.

Основні показники ефективності для системи $M \mid M \mid 1 \mid \infty$ визначаються за такими формулами (7.19), отриманими шляхом граничного переходу при $m \to \infty$ з відповідних формул для системи $M \mid M \mid 1 \mid m$.. Оскільки при відсутності обмежень за довжиною черги кожна заявка, що надійшла в систему, буде обслужена, то

$$P_{vidm} = 0, q = 1, A = \lambda q = \lambda, \bar{r} = \frac{\rho^2}{1-\rho}, \bar{k} = \bar{r} + \rho = \frac{\rho}{1-\rho}, \bar{t}_{och} = \frac{1}{\mu}\frac{\rho^2}{\lambda(1-\rho)},$$

$$\bar{t}_{sist} = \frac{1}{\mu}\frac{1}{1-\rho}. \qquad (7.19).$$

**Приклад 5.** Розглядається одноканальна СМО, в яку надходить найпростіший потік заявок з інтенсивністю $\lambda =$ 2 заявки/год. Обслуговування проводиться із середнім часом $\bar{t}_{obs} = 20min$. _У системі є внутрішній бункер із двома місцями, на яких можуть чекати обслуговування заявки, що надходять. Якщо обидва місця зайняті, то заявки змушені чекати на зовнішньому накопичувачі. Потрібно знайти для стаціонарного режиму характеристики системи; середню кількість заявок, що очікують на зовнішньому накопичувачі, обчислити середньодобову величину штрафу за простій у зовнішньому накопичувачі, якщо за 1 годину очікування однієї заявки система сплачує $a$ од.

*Розв'язання.* За умовою задачі маємо: $\mu = 3$ заявки / год, $\lambda = 2$ заявки / год,

$$\rho = \frac{2}{3}, \qquad \rho < 1, \qquad q = 1.$$

Середня кількість заявок у системі

$$\bar{k} = \frac{\rho}{1-\rho} = 2 \text{ заявки.}$$

Середня тривалість перебування заявки в системі

$$\bar{t}_{sist} = \frac{1}{\mu}\frac{1}{1-\rho} = 1 \text{ година.}$$

Середня кількість заявок у черзі

$$\bar{r} = \frac{\rho^2}{1-\rho} = \frac{4/9}{1/3} = \frac{4}{3} \text{ заявки.}$$

Середня тривалість перебування заявки в черзі

$$\bar{t}_{och} = \frac{\rho^2}{\lambda(1-\rho)} = \frac{2}{3} \text{ години.}$$

Позначимо $L_{vn.nak}$ – довжину черги в зовнішньому накопичувачі. Тоді для обчислення середньої довжини черги необхідно враховувати, що стан $S_4$ – це, коли канал зайнятий, дві заявки знаходяться у внутрішньому бункері, надійшла ще одна заявка, що розміщується в зовнішньому накопичувачі, і т. д.

$$\bar{L}_{vn.nak} = \frac{\rho^4}{1-\rho} = 3 \cdot \frac{16}{81} = \frac{16}{27} \text{заявки.}$$

Середня тривалість перебування заявки в зовнішньому накопичувачі

$$\bar{T}_{vn.nak} = \frac{\bar{L}_{vn.nak}}{\lambda} = \frac{16/27}{2} = \frac{8}{27} \text{ годин.}$$

І, нарешті, сума штрафу

$$S_{\text{штрафу}} = \lambda_{\text{доба}} \cdot T_{vn.ocher} \cdot a = 2 \cdot 24 \frac{8}{27} a \approx 14{,}2 \cdot a.$$

### 7.4. Багатоканальна система масового обслуговування з чергою

Схема СМО, яка розглядається, має вигляд: $M \mid M \mid n \mid m$, тобто кількість каналів обмежена і дорівнює $n$ $(n > 1)$, кількість місць у черзі теж обмежена і дорівнює $m$ $(m > 1)$. На вхід системи надходить потік заявок з інтенсивністю λ, а інтенсивність обслуговування кожного каналу μ.

Для правильної побудови графа станів необхідно перерахувати можливі стани СМО:

$S_0$ – усі канали вільні, черги немає;

$S_1$ – зайнятий один канал, черги немає;

$S_2$ – зайняті два канали, черги немає;

...

$S_n$ – зайняті всі $n$ каналів, черги немає;

$S_{n+1}$ – зайняті всі канали і одна заявка перебуває в черзі;

$S_{n+2}$ – канали зайняті, і дві заявки в черзі;

...

$S_{n+m}$ – зайняті всі $n$ каналів, і всі $m$ місць у черзі.

Розмічений граф, у якому $n + m + 1$ стан, поданий на рис. 7.7. Як був розмічений цей граф?

Над усіма верхніми стрілками залишена інтенсивність вхідного потоку заявок λ. З такою інтенсивністю вхідні заявки переводять систему з будь-якого стану в сусідній правий. Тепер щодо нижніх стрілок. Якщо СМО знаходиться в стані $S_1$, то зайнятий один канал, який обслуговує свою заявку з інтенсивністю μ. Тому стрілкою в графі на рис. 3.18, що з'єднує стани $S_1$ і $S_0$, поставлена інтенсивність поверх обслуговування μ. Якщо ж система знаходиться в стані $S_2$, то зайняті два канали і кожен із них обслуговує свою заявку з інтенсивністю $\mu$, тому сумарна інтенсивність буде 2μ і над стрілкою, що з'єднує стани $S_2$ і $S_1$, позначено інтенсивність обслуговування 2μ і т. д.: 3μ, 4μ,… Коли ж СМО знаходиться в стані $S_n$, то зайняті всі $n$ каналів; кожен із них обслуговує свою заявку з інтенсивністю μ, і тому сумарна інтенсивність є максимальною та дорівнює $n$μ. Вона поставлена над стрілкою, що з'єднує стани $S_n$ і $S_{n-1}$. Далі інтенсивність обслуговування не зростатиме, оскільки будуть займатися тільки вільні місця черги, тому під усіма іншими нижніми стрілками у графі поставлена інтенсивність $n\mu$.

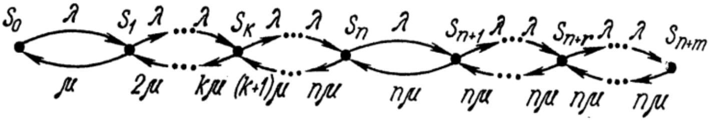


Рис. 7.7. Граф багатоканальної СМО з обмеженою чергою

Користуючись уведеним позначенням $\rho = \lambda/\mu$ (це наведена інтенсивність $\mu$ обслуговування) і готовими формулами процесу «розмноження і загибелі», можна записати

$$P = \left(1 + \rho + \frac{\rho^2}{2!} + \frac{\rho^3}{3!} + \cdots \frac{\rho^n}{n!} + \frac{\rho^{n+1}}{nn!} + \frac{\rho^{n+2}}{n^2 n!} + \cdots + \frac{\rho^{n+m}}{n^m n!}\right)^{-1},$$

$$P_1 = \rho P_0, P_2 = \frac{\rho^2}{2!} P_0, P_3 = \frac{\rho^3}{3!} P_0, \dots, P_n = \frac{\rho^n}{n!} P_0,$$

$$P_{n+1} = \frac{\rho^{n+1}}{n \cdot n!} P_0, P_{n+2} = \frac{\rho^{n+1}}{n^2 \cdot n!} P_0, \dots, P_{n+m} = \frac{m}{n^m \cdot n!} P_0.$$

Наведені вирази розрахунку ймовірностей станів СМО можна записати компактно

$$P_a = \frac{\rho^k}{a!} P_0, \;\; a = \overline{1, n}, \;\; P_{n+\beta} = \frac{\rho^{n+\beta}}{n^\beta n!} P_0, \;\; \beta = \overline{1, m},$$

$$P_0 = \left(\sum_{\alpha=1}^{n} \frac{\rho^a}{a!} + \frac{\rho^n}{n!} \sum_{\beta=0}^{m} \frac{\rho^\beta}{\beta!}\right)^{-1}. \tag{7.20}$$

Останні $m$ доданків у виразі для $P_0$ (7.20) є членами геометричної прогресії зі знаменником тому $\rho/n$, отже

$$P_0 = \left(\sum_{\alpha=0}^{n} \frac{\rho^a}{a!} + \frac{\rho^n}{n!} \frac{\frac{\rho}{n} - (\rho/n)^{m+1}}{1 - \rho/n}\right)^{-1}. \tag{7.21}$$

Визначимо характеристики ефективності цієї СМО. Коли заявка, що надійшла в систему, отримує відмову? Коли все зайнято: усі $n$ каналів і всі $m$ місць у черзі

$$P_{vidm} = P_{n+m} = \frac{\rho^{n+m}}{n^m n!} P_0, \;\; q = 1 - P_{otk} = 1 - \frac{\rho^{n+m}}{n^m n!} P_0, \tag{7.22}$$

$$A = \lambda q = \lambda\left(1 - \frac{\rho^{n+m}}{n^m n!} P_0\right). \tag{7.23}$$

Імовірність утворення черги дорівнює сумі ймовірностей через те всі канали зайняті і в системі буде знаходитися $n, n+1, n+2, \dots, n+m-1$ заявок

$$P_{och} = \sum_{k=n}^{n+m+1} P_k = \frac{\rho^{\mathrm{n}}}{n!} \frac{1-\left(\frac{\rho}{\mathrm{n}}\right)^m}{1-\frac{\rho}{\mathrm{n}}} P_0.$$

Для знаходження середнього часу очікування заявки в черзі $\bar{t}_{och}$ потрібно розглянути випадкову величину $G$, що набуває значення $\frac{r+1}{n\mu}$ з ймовірностями $P_{n+r}$, для всіх $r$, що змінюються від 0 до $m-1$:

$$\bar{t}_{och} = \frac{1}{n\mu} P_n + \frac{2}{n\mu} P_{n+1} + \cdots + \frac{m}{n\mu} P_{n+m+1} = \frac{\rho^n P_0}{nm!\mu}\left(1 + \frac{2\rho}{n} + \frac{3\rho^2}{n^2} + \cdots + \frac{m\rho^{m-1}}{n^{m-1}}\right). \quad (7.24)$$

Для знаходження середньої кількості заявок, що очікують початку обслуговування $\bar{r}$, розглядається дискретна випадкова величина $R$, що набуває значення $r$ з імовірністю $P_{n+r}$ для всіх $r$, що змінюються від 1 до $m$:

$$\bar{r} = M(R) = 1 \cdot P_{n+1} + 2 \cdot P_{n+2} + \cdots + mP_{n+m} = \sum_{r=1}^{m} r \cdot P_{n+r} = \frac{\rho^{n+1} P_0}{nm!}\left(1 + \frac{2\rho}{n} + \right.$$
$$\left. + \frac{3\rho^2}{n^2} + \cdots + \frac{m\rho^{m-1}}{n^{m-1}}\right). \quad (7.25)$$

Вираз (7.24) з урахуванням (7.25) можна переписати

$$\bar{t}_{och} = \bar{r}/\lambda. \quad (7.26)$$

У виразах (7.24) і (7.25) у дужках знаходиться похідна за $\rho/n$ суми членів геометричної прогресії

$$\sum_{i=1}^{m} \left(\frac{\rho}{n}\right)^i = \frac{\frac{\rho}{n} - \left(\frac{\rho}{n}\right)^{m+1}}{1 - \frac{\rho}{n}},$$

диференціюючи яку за $\rho/n$ і роблячи деякі перетворення, отримують

$$\bar{r} = \frac{\rho P_0}{n \cdot n!}\left(\frac{1-(m+1)\left(\frac{\rho}{\mathrm{n}}\right)^m + m\left(\frac{\rho}{n}\right)^{m+1}}{\left(1-\frac{\rho}{n}\right)^2}\right). \quad (7.27)$$

Аналогічно можна отримати вираз для $\bar{t}_{och}$.

Середня кількість зайнятих каналів $\bar{z}$ визначається з урахуванням того, що кожен зайнятий канал обслуговує в середньому μ заявок за одиницю часу, а вся СМО обслуговує $A$ заявок, тоді

$$\bar{z} = \frac{A}{\mu} = \rho\left(1 - \frac{\rho^{n+m}}{n^m \cdot n!} P_0\right). \quad (7.28)$$

Середня кількість заявок $\bar{k}$, пов'язаних з системою, визначається з урахуванням виразів (7.27) і (7.28):

$$\bar{k} = \bar{r} + \bar{z}.$$

Для знаходження середнього часу перебування заявки в системі можна, використовуючи вираз (7.26), записати

$$\bar{t}_{sist} = \bar{t}_{och} + \bar{t}_{obs} = \frac{\bar{r}}{\lambda} + \frac{q}{\lambda}. \quad (7.29)$$

Таким чином, розглянуто основні показники ефективності $n$-канальної СМО з очікуванням, причому довжина черги обмежена і дорівнює $m$.

***Приклад 6.*** У СМО в середньому через кожні 30 хв надходять заявки на обслуговування. Середня тривалість обслуговування однієї заявки становить 1,5 год. Обробку заявок здійснюють два канали. У системі є накопичувальний бункер, у якому можуть знаходитися в черзі в очікуванні обслуговування не більше 4 заявок. Визначити основні характеристики СМО.

*Розв'язання.* Розглядається СМО виду $M \mid M \mid 2 \mid 4$, тобто $n = 2$, $m = 4$, $\lambda = 2$ заявки / год, μ = 2/3 заявки / год, $\rho = \frac{\lambda}{\mu} = 3, \frac{\rho}{n} = 3/2$..

Відповідно до виразу (3.42) імовірність нульового стану обчислюється так:

$$P_0 = \left(1 + 3 + \frac{3^{\,2}}{2!} + \frac{3}{2!}\frac{\frac{3}{2} - \left(\frac{3}{2}\right)^5}{1 - 3/2}\right)^{-1} = 0{,}158.$$

Імовірності інших станів відповідно до перших двох формул виразів (7.29): $P_1 = 3 \cdot 0{,}0158 = 0{,}0474$; $P_2 = \frac{3^2}{2} \cdot 0{,}0158 = 0{,}0711$; $P_3 = 0{,}107$; $P_4 = 0{,}160$; $P_5 = 0{,}240$; $P_6 = 0{,}360$.

Можна перевірити правильність проведених обчислень, оскільки сума $P_i$ повинна дорівнювати одиниці або з урахуванням округлень – приблизно одиниці.

Імовірність того, яка заявка, що надійшла в СМО, отримає відмову (усі канали зайняті і всі місця в черзі теж зайняті):

$$P_{vidm} = P_6 = 0{,}36.$$

Імовірність утворення черги

$$P_{och} = \sum_{k=n}^{n+m+1} P_k = \frac{\rho^n}{n!} \frac{1-\left(\frac{\rho}{n}\right)^m}{1-\frac{\rho}{n}} P_0 = \frac{3^2}{2} \frac{1-\left(\frac{3}{2}\right)^4}{1-\frac{3}{2}} 0{,}0158 \approx 0{,}58,$$

але цю величину (ймовірність утворення черги) можна обчислити, просто підсумувавши вже отримані значення ймовірностей $P_2$, $P_3$, $P_4$ і $P_5$ .

Імовірність того, що заявка, яка надійшла в СМО, буде обслужена, обчислюється за формулою (7.22):

$$q = 1 - P_{\text{отр}} = 0{,}64.$$

Продуктивність СМО (7.23): $A = 2 \cdot 0{,}64 = 1{,}28$ заявок / год.

Середня кількість заявок, що знаходяться в черзі, (7.25): $\bar{r} \approx 2{,}6$ заявки.

Середня тривалість очікування заявки в черзі (7.26):

$$\bar{t}_{och} = \frac{2{,}6}{2} = 1{,}3 \text{ год.}$$

Середня тривалість обслуговування заявки:

$$\bar{t}_{obs} = 0{,}96 \text{ год} \approx 58 \text{ хв.}$$

Середня тривалість перебування заявки в системі (7.29):

$$\bar{t}_{sist} = 1{,}3 + 0{,}96 = 2{,}26 \text{ год.}$$

Середня кількість зайнятих каналів (7.28):

$$\bar{z} = \frac{1{,}28}{2/3} = 1{,}92 \text{ каналів.}$$

Коефіцієнт зайнятості каналів

$$Z_{zan} = \frac{\bar{z}}{n} = 1{,}92/2 = 0{,}96.$$

Отримані характеристики СМО $M \mid M \mid 2 \mid 4$ свідчать про те, що частка простою каналів дуже мала і становить всього 1,58 % робочого часу, а ймовірність відмови в обслуговуванні дуже велика: 36 % заявок отримають відмову. Коефіцієнт завантаження каналів – 0,96, а відносна пропускна здатність маленька (тільки 64 % заявок буде обслужено). З усього цього можна зробити висновок: СМО не справляється з потоком заявок, і необхідно збільшити кількість каналів або інтенсивність обслуговування та, можливо, збільшити кількість місць у черзі.

За аналогією з підходом, використовуваним вище, можна перейти до системи з необмеженою чергою. Нехай у СМО, що складається з $n$ каналів, надходить найпростіший потік заявок з інтенсивністю $\lambda$, час обслуговування однієї заявки підпорядкований показовому закону розподілу з параметром $\mu$.

Розмічений граф станів таких систем $M \mid M \mid n \mid \infty$ має нескінченну кількість станів (рис. 7.8).

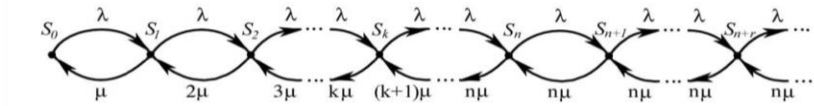


Рис. 7.8. Граф багатоканальної СМО з необмеженою чергою.

Можна перерахувати можливі стани в цьому графі:

$S_0$ – всі канали вільні, черги немає;

$S_1$ – зайнятий один канал, черги немає;

$S_2$ – зайняті два канали, черги немає;

...

$S_n$ – зайняті всі $n$ каналів, черги немає;

$S_{n+1}$ – зайняті всі канали, в черзі одна заявка;

$S_{n+2}$ – зайняті всі канали, в черзі дві заявки;

...

$S_{n+i}$ – зайняті всі канали, в черзі заявок; $i$

...

Використовуючи формули (7.20) і (7.21), отримують імовірності станів для системи $M \mid M \mid n \mid \infty$ , яка розглядається, здійснюючи граничний перехід ($m \to \infty$). Іншим підходом може бути використання формул процесу «розмноження і загибелі». У будь-якому разі вирази $M \mid M \mid n \mid \infty$ для ймовірностей станів системи записані нижче:

$$P_1 = \rho \cdot P_0, P_2 = \frac{\rho^2}{2!} P_0, \dots, P_n = \frac{\rho^n}{n!} P_0,$$

$$P_{n+1} = \frac{\rho^{n+1}}{n \cdot n!} P_0, P_{n+2} = \frac{\rho^{n+2}}{n^2 \cdot n!} P_0, \dots, P_{n+i} = \frac{\rho^{n+i}}{n^i \cdot n!} P_0, \dots,$$

$$P_0 = \left(1 + \rho + \frac{\rho^2}{2!} + \dots + \frac{\rho^{n+1}}{n! \cdot n(n-\rho)}\right)^{-1}.$$

Отримані вирази можна записати більш компактно

$$P_\alpha = \frac{\rho^\alpha}{\alpha!} P_0,\ \alpha = \overline{1, n},\ P_{n+1} = \frac{\rho^{n+1}}{nn!} P_0, \dots, P_{n+r} = \frac{\rho^{n+r}}{n^r n!} P_0, \dots$$

$$P_0 = \left( \sum_{\alpha=0}^{n} \frac{\rho^a}{a!} + \frac{\rho^{n+1}}{n(n-\rho)} \right)^{-1}.$$

Вираз для $P_0$ у формулах (7.29) отримано в припущенні, що $\rho/n < 1$,, оскільки тільки при виконанні цієї умови існує усталений режим, а при $\rho/n \geq 1$ черга нескінченно зростає.

Визначимо основні залежності для обчислення параметрів ефективності СМО, яка розглядається.

Імовірність відмови в обслуговуванні $P_{vidm} = 0$, оскільки кількість місць у черзі не є обмеженою і при зайнятих каналах заявка, що надійшла, завжди буде поставлена в чергу та коли-небудь буде обслужена.

Імовірність утворення черги обчислюється як ймовірність того, що всі канали зайняті

$$P_{ocher} = \frac{\rho^{n+1}}{n!(n-\rho)} P_0.$$

Імовірність обслуговування $q = 1$, а абсолютна пропускна здатність відповідно $A = \lambda$.

Середня довжина черги $\bar{r}$ знаходиться з (3.49) при $n \to \infty$:

$$\bar{r} = \frac{\rho^{n+1}}{n \cdot n!(1-\rho/n)^2} P_0 = \frac{n}{n-\rho} P_{ocher}. \quad (7.30)$$

З (7.24) і (7.26) або з (7.30) знаходиться середня тривалість очікування заявки початку обслуговування

$$\bar{t}_{och} = \frac{\bar{r}}{\lambda} = \frac{\rho^n}{n \cdot \mu \cdot n! \left(1 - \frac{\rho}{n}\right)^2} P_0.$$

Середня кількість зайнятих каналів $\bar{z} = \frac{A}{\mu} = \rho$. Як і раніше, середня кількість $\mu$ заявок, пов'язаних із системою, $\bar{k} = \bar{r} + \bar{z}$, а середній час перебування заявок у СМО

$$\bar{t}_{sist} = \bar{t}_{ochid} + 1/\mu.$$

***Приклад 7.*** Розглядається СМО виду $M \mid M \mid 3 \mid \infty$ , тобто це триканальна система з необмеженою кількістю місць у черзі. Інтенсивність вхідного потоку заявок становить 5 заявок за одиницю часу. Інтенсивність обслуговування заявок кожним каналом – 2 заявки за одиницю часу.

Визначити характеристики СМО.

Таким чином, $n = 3, \lambda = 5, \mu = 2, \rho = \frac{\lambda}{\mu} = 2{,}5, \rho/n = 2{,}5/3 < 1$.

Відповідно до формул (7.29): $P_0 \approx 0{,}045$; $P_1 \approx 0{,}112$; $P_2 \approx 0{,}141$;

$P_3 \approx 0{,}117$; $P_4 \approx 0{,}098$; $P_5 \approx 0{,}081$; $P_6 \approx 0{,}068$; $P_7 \approx 0{,}057$; $P_8 \approx 0{,}047$; $P_9 \approx 0{,}039; \ldots$

Аналіз результатів розрахунку ймовірностей станів показує, що ймовірність того, що обслуговуючі канали вільні ($P_0$), значно менша величин $P_1$, $P_2$, або $P_3$ (тобто ймовірностей зайнятості одного, двох або трьох каналів). Потім, зі зростанням черги ймовірності станів зменшуються і при п'яти зайнятих місцях у черзі $P_0$ і $P_8$ майже рівні. Це свідчить про те, що в даній СМО малоймовірні як простоювання всіх обслуговуючих каналів, так і черга з п'яти і більше заявок.

Імовірність виникнення черги

$$P_{ocher} = \frac{\rho^{n+1}}{n!(n-\rho)} P_0 \approx 0{,}586.$$

Цю величину можна обчислити ще й шляхом вирахування з одиниці суми ймовірностей тих станів, у яких черга відсутня

$$P_{ocher} = 1 - (P_0 + P_1 + P_2 + P_3) = 1 - (0{,}045 + 0{,}112 + 0{,}141 + 0{,}117) \approx 0{,}586.$$

Імовірність обслуговування, як уже говорилося, завжди дорівнює одиниці ($q = 1$), а продуктивність цієї СМО $A = \lambda\, q = 5$ (заявок за одиницю часу).

Інші характеристики:

$$\bar{r} = \frac{n}{n-\rho} P_{ocher} \approx 3{,}5;\ \bar{t}_{och} = \frac{\bar{r}}{\lambda} = 0{,}7;\ \bar{t}_{obs} = \frac{1}{\mu} = 0{,}5,$$
$$\bar{t}_{sist} = \bar{t}_{ochid} + \bar{t}_{obs} = 1{,}2;\ \bar{k} = \bar{r} + \bar{z} = 3{,}5 + 2{,}5 = 6,$$

Таким чином, частка простою каналу дорівнює 4,5 % від тривалості інтервалу функціонування СМО; ймовірність виникнення черги досить велика – 0,586, але довжина черги не велика (3,5 заявки) і час очікування обслуговування не великий (0,7 одиниць часу). Роботу цієї системи слід визнати задовільною.

## 7.5. Системи масового обслуговування з обмеженим часом очікування

Розглядається канальна СМО з очікуванням і необмеженою кількістю місць у черзі. Однак час очікування заявки початку обслуговування в цій системі обмежений певним випадковим терміном із середнім значенням $\bar{t}_{och}$ (тобто заявка вважається нетерплячою). Якщо до закінчення цього часу заявка не буде прийнята до обслуговування, то вона залишає чергу і йде із системи необслуженою. У цьому випадку можна сказати, що на кожну заявку, яка стоїть у черзі, діє «потік відходів» інтенсивності $\nu = \frac{1}{\bar{t}_{och}}$.

Якщо цей потік пуассонівський, то процес, що протікає в СМО, є марківським. Розмічений граф станів такої системи наведений на рис. 7.9.

При цьому над стрілками, спрямованими справа наліво від $S_n$ до $S_0$, може бути проставлена сумарна інтенсивність потоку обслуговування зайнятих каналів.

У решти стрілок справа наліво проставлені сумарні інтенсивності потоку обслуговувань усіх $n$ каналів, $n\mu$, а також відповідна інтенсивність потоку відходів із черги. Це знову схема «розмноження і загибелі». Тому, користуючись формулами з параграфа 8.4 і позначаючи додатково $\beta = \frac{\nu}{\mu}$, можна записати

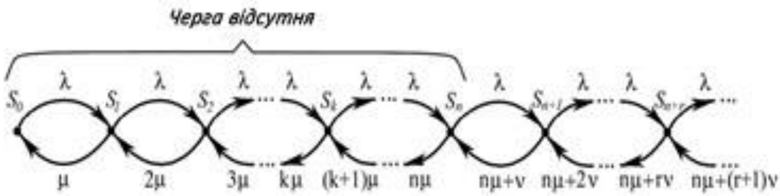


Рис. 7.9. Граф станів СМО з обмеженим часом очікування

$$P_0 = \left\{1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \cdots + \frac{\rho^n}{n!} + \frac{\rho^n}{n!}\left[\frac{\rho}{n+\beta} + \frac{\rho^2}{(n+\beta)(n+2\beta)} + \cdots + \right.\right.$$

$$\left.\left. + \frac{\rho^r}{(n+\beta)(n+2\beta)\ldots(n+r\beta)+\cdots}\right]\right\}^{-1} = \left\{\sum_{\alpha=0}^{n} \frac{\rho^\alpha}{\alpha!} + \frac{\rho^n}{n!}\sum_{r=1}^{\infty} \frac{\rho^r}{\prod_{s=1}^{r}(n+s\beta)}\right\}^{-1},$$

$$P_\alpha = \frac{\rho^\alpha}{\alpha!} P_0, \alpha = \overline{1, n},$$

$$P_{n+1} = \frac{\rho^n}{n!}\frac{\rho}{n+\beta}P_0, P_{n+2} = \frac{\rho^n}{n!}\frac{\rho^2}{(n+\beta)(n+2\beta)}P_0, \dots, P_{n+r} = \frac{\rho^n}{n!}\frac{\rho^r}{\prod_{s=1}^{r}(n+s\beta)}P_0. \tag{7.31}$$

У формулах (7.31) у виразі для $P_0$ записана сума нескінченного ряду, яка не є прогресією. За необхідності вона може бути обчислена наближено.

Якщо у формулах (7.31) перейти до межі при $\gamma \to 0$ (або $\beta \to 0$), то при $\rho < n$ виходять вирази (7.20) і (7.21), тобто «нетерплячі» заявки стають «терплячими».

Тепер можна визначити основні характеристики цієї системи. Поняття «ймовірність відмови», мабуть, не має сенсу, оскільки кожна заявка, що надійшла в систему, ставиться в чергу. Але заявка може і не дочекатися обслуговування, пішовши з черги після закінчення часу очікування.

Середня кількість заявок у черзі

$$\bar{r} = M\,[R] = \sum_{r=1}^{\infty} rP_{n+r} = 1 \cdot P_{n+1} + 2 \cdot P_{n+2} + \cdots$$

Абсолютна пропускна здатність СМО $A$ (середня кількість обслужених заявок за одиницю часу) легко визначається, якщо врахувати, що з середньої кількості заявок у черзі $\bar{r}$ під впливом «потоку відходу» буде йти, не дочекавшись обслуговування, $\gamma\bar{r}$ заявок за одиницю часу. Тому

$$A = \lambda - \gamma\,\bar{r}, \quad q = \frac{A}{\lambda} = 1 - \frac{\gamma}{\lambda}\bar{r}.$$

Можна знайти середню кількість зайнятих каналів

$$\bar{z} = \frac{A}{\mu} = \rho - \beta\bar{r},$$

а звідси

$$\bar{r} = \frac{\rho - \bar{z}}{\beta}.$$

Що стосується середньої кількості зайнятих каналів, то її можна знайти як математичне очікування величини $z$, що набуває значення $0, 1, 2, \dots, n$ з імовірностями

$$P_0, P_1, P_2, \dots, P_{n-1}\ (1 - (P_0 + P_1 + P_2 + \dots + P_{n-1})),$$

$$\bar{z} = 1p + 2p + \ldots + n\left(1 - \sum_{k=0}^{n-1} P_k\right).$$

Формули для середнього часу очікування в черзі і середнього часу перебування заявки в системі виводяться досить складно.

## Контрольні запитання

1. Що є аналітичною моделлю?
2. В чому полягає відмінність аналітичних та імітаційних методів моделювання?
3. Для якого випадкового процесу, що протікає в системі масового обслуговування, можливо задіяти аналітичне моделювання?
4. Які характеристики ефективності обслуговування можуть застосовуватися при аналітичному моделюванні СМО?
5. Загальна характеристика моделі багатоканальної системи масового обслуговування з відмовами.
6. Загальна характеристика моделі одноканальної системи масового обслуговування з чергою.
7. Загальна характеристика моделі багатоканальної системи масового обслуговування з чергою.

# ВИСНОВОК

У навчальниму посібнику викладено методології моделювання різних процесів у системах і використання комп'ютерних технологій або реальних пристроїв. Для наукового дослідження системи вдаються до деяких припущень, які стосуються її функціонування. Опис процесу функціонування у вигляді математичних або логічних залежностей формує модель, що дає уявлення про поведінку системи. Набір процедур формування моделі та її дослідження створює технологію моделювання.

Процес проектування інформаційних систем, що реалізують нову інформаційну технологію, безперервно удосконалюється. У центрі уваги інженерів-системотехніків виявляються дедалі більш складні системи, що ускладнює використання фізичних моделей і підвищує значущість математичних моделей та машинного моделювання систем. Машинне моделювання стало ефективним інструментом дослідження і проектування складних систем. Актуальність математичних моделей безперервно зростає через їх гнучкість, адекватність реальним процесам, невисоку вартість реалізації на базі сучасних ЕОМ. Великі можливості надаються користувачеві, тобто фахівцеві з моделювання систем, засобами обчислювальної техніки. Особливо ефективним є застосування моделювання на ранніх етапах проектування автоматизованих систем, коли ціна помилкових рішень найбільш значна. Сучасні обчислювальні засоби дозволили істотно збільшити складність використовуваних моделей при вивченні систем, з'явилася можливість побудови комбінованих, аналітикоімітаціонних моделей, що враховують все різноманіття чинників, що мають місце в реальних системах, тобто використання моделей, більш адекватних досліджуваним явищам.

Тому в навчальному посібнику особливу увагу приділено методологічним аспектам моделювання, типовим математичним схемами, методам і засобам їх реалізації в процесі моделювання. В цілому в навчальному посібнику зроблена спроба системного підходу до викладу наукових основ моделювання систем. Перспективним і значущим для теорії і практики системного моделювання є подальший розвиток наукових основ моделювання з орієнтацією на нові інформаційні технології в наукових дослідженнях, проектуванні, управлінні і навчанні. Однією з основних сьогодні є проблема інформатизації освіти в умовах переходу до інформаційного суспільства. У цих умовах інформатизацію можна розглядати як процес переходу від індустріального суспільства до інформаційного. Відбувається перерозподіл трудових ресурсів з матеріального виробництва в сферу інформації. У ряді країн сумарні витрати на комп’ютерну техніку, телекомунікації, електроніку перевищують витрати на енергетику, тобто настало століття інформатизації. Для інформаційного суспільства характерно повне задоволення інформаційних потреб при завершенні формування єдиного інформаційного середовища, яке визначає інформаційну культуру суспільства і кожної людини. Інформаційна культура не обмежується системою знань в сфері інформаційних процесів і повинна включати активно-перетворювальний аспект ставлення до світу. Досягненнями інформатики, в якій як у науковому напрямі можна виділити 3 рівні:

- фізичний – програмно-апаратні засоби обчислювальної техніки і техніки зв’язку, тобто засоби телематики (телекомунікацій та інформатики);
- логічний – інформаційні технології;
- прикладний – призначені для користувача інформаційні системи.

Виходячи з цього, можна провести аналіз стану і тенденцій розвитку процесу інформатизації та визначити місце комп’ютерного (машинного) моделювання в цьому багатогранному процесі. Моделювання взагалі і комп’ютерне імітаційне моделювання, зокрема, набувають застосування в таких інформаційних технологіях:

1) комп’ютерні програми, які включають у себе електронні підручники, тренажери, лабораторні практикуми, тестові системи, ділові ігри, що використовують машинні імітатори;

2) системи на базі мультимедіа технологій, побудовані з використанням персональних ЕОМ, відеотехніки, накопичувачів на оптичних дисках, включаючи системи віртуальної реальності;

3) інтелектуальні і навчальні експертні системи, котрі широко використовують імітаційний підхід у різних предметних галузях;

4) розподіл бази даних за галузями знань, електронні бібліотеки, розподілені і централізовані видавничі системи;

5) засоби телекомунікації, що включають в себе електронну пошту, телеконференції, локальні і регіональні мережі зв’язку, цифрові мережі інтегрального обслуговування і т. д. з використанням нових інтелектуальних мережевих технологій;

6) геоінформаційні системи, що базуються на технології, яка об'єднує комп'ютерну картографію та системи керування базами даних, і реалізують технологію створення багатошарової електронної карти, опорний шар якої описує географію території, а кожен з інших шарів – один з аспектів стану території;

7) технології захисту інформації, включаючи захист від несанкціонованого доступу до ПЕОМ, захист від перехоплення в мережах і т. п.

Важливо відзначити, що, кажучи про сукупність засобів інформатизації, слід мати на увазі не тільки засоби обчислювальної техніки та деяку «суму інформаційних технологій», але також і суму суспільних знань та умінь з використання зазначених засобів, що може бути визначено як рівень суспільного навчання. Очевидно, що як одна предметна галузь, у тому числі і машинне моделювання, не може переступити через деякі об'єктивні стадії такого громадського навчання. Поряд із психологічною необхідно забезпечити і професійну підготовленість користувачів. Особливо це актуально в сфері використання ЕОМ для цілей імітації широкого класу систем. Таким чином, інформатизація є об'єктивною реальністю.

Перехід до інформаційного суспільства передбачає формування інформаційної культури як зводу правил поведінки в інформаційному суспільстві, в комунікаційному середовищі, людино-машинних системах, що вписуються в світову гуманістичну культуру людства. Входження користувача в світову мережу (наприклад, Internet) дозволяє отримувати величезні потоки інформації, в тому числі і для цілей моделювання.

Комп'ютерні реалізації моделей дозволяють отримувати числові оцінки характеристик систем, а за допомогою отриманих даних можна розрахувати і реальні характеристики. Результати комп'ютерних експериментів дозволяють розвіяти існуюче враження про моделювання як вправу в програмуванні, нехай і дуже складному.

## СПИСОК ЛІТЕРАТУРИ


1. Советов Б. Я. Моделирование систем : учебник для ВУЗов / Б. Я. Советов, С. А. Яковлєв. – Москва : Высшая школа . – 343 с.

2. Моделирование информационных систем : учеб. пособие / под ред. О. И. Шелухина. – Москва : Радиотехника, 2005. – 368 с.

3. Борисов Ю. П. Математическое моделирование радиосистем : учеб. пособие / Ю. П. Борисов. – Москва : Советское радио, 1976. – 296 с.

4. Быков В. В. Цифровое моделирование в статистической радиотехнике / В. В. Быков. – Москва : Советское радио, 1971. – 328 с.

5. Волков Л. Н. Системы цифровой радиосвязи: базовые методы и характеристики : учеб. пособие / Л. Н. Волков, М. С. Немировский, Ю. С. Шинаков. – Москва : Эко-Трендз, 2005. – 392 с.

6. Кловский Д. Д. Модели непрерывных каналов связи на основе стохастических дифференциальных уравнений / Д. Д, Кловский, В. Я. Конторович, С. М. Широков. – Москва : Радио и связь, 1984. – 248 с.

7. Комашинский В. И. Системы подвижной радиосвязи с пакетной передачей информации. Основы моделирования / В. И. Комашинский, А. В. Максимов. – Москва : Горячая линия. – Телеком, 2007. – 176 с.

8. Крамущенко В. И. Многоканальные системы передачи информации : конспект лекций / В. И. Крамущенко, Л. Я. Новосельцев, В. Н. Смирнов. – Ленинград : ЛЭТИ, 1983. – 48 с.

9. Мандельброт Б. Фрактальная геометрия природы / Б. Мандельброт ; пер. с англ. А. Р. Логунова; под ред. А. Д. Морозова. – Москва: Институт компьютерных исследований, 2002. – 656 с.

10. Полляк Ю. Г. Статистическое машинное моделирование средств связи / Ю. Г. Полляк, В. А. Филимонов. – Москва : Радио и связь, 1988. – 176 с. (Серия СТС – Вып. 30).

11. Теория электрической связи : учеб. пособие / А. Г. Зюко, Д. Д. Кловский, В. И. Коржик, В. Д. Назаров ; под ред. Д. Д. Кловского. – Москва : Радио и связь, 1999. – 432 с.

12. Хорафас Д. Н. Системы и моделирование / Д. Н. Хорафас; пер. с англ; под ред. И. Н. Коваленко. – Москва : Мир, 1967. – 418 с.

13. Шалыгин А. С. Прикладные методы статистического моделирования / А. С. Шалыгин, Ю. И. Палагин. – Ленинград : Машиностроение, 1986. – 320 с.

14. Шеннон Р. Имитационное моделирование систем – искусство и наука / Р. Шеннон ; пер. с англ.: под ред. Е. К. Масловского. – Москва : Мир, 1978. – 418 с.

15. Голяницкий И. А. Математические модели и методы в радиосвязи / И. А. Голяницкий ; под ред. Ю. А. Громакова. – Москва : Эко-Трендз, 2005. – 440 с.

16. Величко В. В. Телекоммуникационные системы и сети:Т. 3 : Мультисервисные сети / В. В. Величко, Е. А. Субботин, В. П. Шувалов, А. Ф. Ярославцев. – Москва : Горячая линия–Телеком, 2005. – 592 с.: ил.

17. Лужковський А. Г. Теорія масового обслуговування в телекомунікаціях / А.Г. Лужковський. – Одеса : ОНАЗ ім. О.С. Попова, 2010. – 112 с.: іл.

## ДЛЯ НОТАТОК

Навчальне видання

**ОБОД** Іван Іванович
**ЗАВОЛОДЬКО** Ганна Едвардівна
**СВИД** Ірина Вікторівна

**МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ СИСТЕМ**
Навчальний посібник
для студентів спеціальностей
«Комп'ютерна інженерія»,
«Комп'ютерні науки та інформаційні технології»

Відповідальний за випуск проф. Серков О. А.
Роботу до видання рекомендував проф. Заполовський М. Й.
Редактор О. В. Козюк


План 2019 р.п. 52

Підписано до друку 27.08.2019 Формат 60x84/16. Папір офсетний
Гарнітура Times New Roman Цифровий друк. Ум. друк. арк.
Наклад 100 прим. Зам. № 2794
Ціна договірна

Видавець та виготовлювач ТОВ «Друкарня Мадрид»
61024, м. Харків, вул. Максиміліанівська, 11
Тел.: (057) 756-53-25
Свідоцтво суб'єкта видавничої справи:
Серія ДК № 4399 від 27.08.12 р.
www.madrid.in.ua e-mail@info@madrid.in.ua