

**Н.С. ЛОБОДА, Є.Д. ГОПЧЕНКО**  
**Стохастичні моделі у гідрологічних**  
**розрахунках**

2006

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**  
**Одеський державний екологічний університет**

**Н.С. ЛОБОДА, Є.Д. ГОПЧЕНКО**  
**Стохастичні моделі у гідрологічних**  
**розрахунках**

*Навчальний посібник*

Рекомендовано Міністерством освіти і науки  
України як навчальний посібник для студентів  
вищих навчальних закладів

(Лист 14/18.2-368 від 15.02.2006 р.)

**О д е с а**  
**“Екологія”**  
**2006**

ББК 26.22  
Л68  
УДК 556.16.01

В навчальному посібнику викладені теоретичні та практичні основи стохастичного моделювання характеристик стоку у природних та порушених водогосподарською діяльністю умовах. Розглянуті питання визначення статистичних параметрів стоку за даними спостережень, установлення параметрів умовного закону розподілу. Представлені стохастичні моделі стоку у виді простого ланцюга Маркова. Посібник розрахований на студентів, магістрів, аспірантів гідрологічних та гідромеліоративних спеціальностей вищих навчальних закладів.

Рецензенти:

д.г.н., проф. *О.О.Светлічний*,  
д.г.н. *Б.В.Кіндюк*  
к.г.н. *Г.М.Андреевська*

Гриф надано Міністерством освіти і науки України  
(Лист № 14/18.2-368 від 15.02.2006 р.).

ISBN

© Н.С.Лобода, 2006  
© Є.Д.Гопченко, 2006  
© Одеський державний  
екологічний університет, 2006

## ЗМІСТ

ВСТУП.....	6
РОЗДІЛ 1. СТАТИСТИЧНІ МЕТОДИ В ГІДРОЛОГІЧНИХ РОЗРАХУНКАХ .....	10
1.1.Обґрунтування можливостей застосування методів математичної статистики до визначення характеристик стоку річок .....	10
1.2 Основні положення теорії ймовірностей .....	13
1.2.1 Поняття ймовірності. Випадкова подія .....	13
1.2.2 Випадкові величини та закони їх розподілу .....	13
1.2.2.1 Інтегральний закон розподілу безперервних випадкових величин .....	15
1.2.2.2 Диференціальний закон розподілу безперервної випадкової величини .....	17
1.2.2.3 Характеристики положення кривих розподілу .....	19
1.3 Статистичні характеристики стокових величин, методи їх визначення та особливості практичного застосування .....	22
1.3.1 Метод моментів. Статистичні моменти як характеристики кривої розподілу .....	23
1.3.2 Точність розрахунку статистичних параметрів за вибірковими даними .....	28
1.3.2.1 Вимоги до оцінок статистичних параметрів .....	28
1.3.2.2 Поняття про систематичні та випадкові похибки визначення статистичних параметрів .....	30
1.3.2.3 Метод Монте-Карло .....	30
1.3.2.4 Метод статистичних випробувань .....	36
1.3.3 Метод найбільшої правдоподібності .....	40
1.3.3.1 Приклад застосування методу найбільшої правдоподібності до визначення статистичних параметрів нормального закону розподілу .....	41
1.3.3.2 Точність визначення статистичних параметрів, обчислених за методом найбільшої правдоподібності .....	45
1.3.4 Графо-аналітичний метод Г. А. Алексєєва .....	47
1.3.5 Теоретичні криві розподілу, їх особливості та межі застосування .....	51
1.3.5.1 Нормальний закон розподілу випадкових величин .....	54
1.3.5.2 Розподіл Пірсона III .....	57
1.3.5.3 Логаріфмічно-нормальний закон розподілу .....	59
1.3.5.4 Крива трипараметричного гама-розподілу С. Н. Крицького і М. Ф. Менкеля .....	60
1.3.6 Клітчатка ймовірностей, їх призначення .....	62

1.3.7 Розрахунки характеристик стоку заданої забезпеченості за теоретичними законами розподілу .....	63
1.4 Приклади розрахунків .....	64
<b>РОЗДІЛ 2. ОСНОВИ ТЕОРІЇ КОРЕЛЯЦІЇ .....</b>	<b>71</b>
2.1 Системи випадкових величин. Закони розподілу випадкових величин .....	71
2.2 Умовні закони розподілу .....	76
2.3 Залежні і незалежні випадкові величини, кореляційний момент, коефіцієнт кореляції .....	79
2.4 Нормальний закон розподілу для системи двох випадкових величин .....	82
2.5 Рівняння парної лінійної регресії, оцінка їх параметрів за вибірковими даними .....	85
2.5.1 Оцінка параметрів рівняння .....	86
2.5.2 Оцінка вірогідності лінійної регресійної моделі .....	88
2.5.3 Визначення довірчих інтервалів для коефіцієнтів рівняння лінійної регресії .....	89
2.5.4 Визначення довірчого інтервалу для коефіцієнта кореляції в рівнянні лінійної регресії .....	93
2.5.5 Перевірка гіпотези про статистичну значущість коефіцієнта кореляції і коефіцієнтів рівняння регресії .....	95
<b>РОЗДІЛ 3. ЗАГАЛЬНІ УЯВЛЕННЯ ПРО ТЕОРІЮ ВИПАДКОВИХ ПРОЦЕСІВ .....</b>	<b>97</b>
3.1 Поняття про випадкову функцію .....	97
3.2 Закон розподілу випадкової функції .....	98
3.3 Характеристики випадкових функцій та їх визначення .....	99
3.4 Поняття про стаціонарність випадкового процесу. Ергодичність стаціонарних випадкових процесів .....	107
3.5 Автокореляційна функція, коефіцієнт автокореляції .....	114
3.6 Приклад розрахунків .....	118
<b>РОЗДІЛ 4. СТОХАСТИЧНІ МОДЕЛІ В ГІДРОЛОГІЧНИХ РОЗРАХУНКАХ .....</b>	<b>121</b>
<b>РОЗДІЛ 5. МОДЕЛЮВАННЯ РІЧНОГО СТОКУ В УМОВАХ ВОДОГОСПОДАРСЬКИХ ПЕРЕТВОРЕНЬ .....</b>	<b>131</b>

5.1 Огляд методів оцінки річного стоку в порушених господарською діяльністю умовах .....	132
5.2 Математичне моделювання побутового річного стоку на водозборах річок зони недостатнього зволоження .....	136
5.2.1 Обґрунтування стохастичної моделі, що описує вплив штучних водоймищ на річний стік .....	138
5.2.2 Обґрунтування стохастичної моделі, яка описує вплив зрошувальних меліорацій на річний стік .....	144
5.2.3 Оцінка характеристик річного стоку в умовах господарської діяльності для річок зони недостатнього зволоження (на основі результатів стохастичного моделювання) .....	149
5.3 Математичне моделювання і розрахунки річного побутового стоку на водозборах річок зони надмірного зволоження .....	157
5.3.1 Огляд методів кількісної оцінки впливу осушення боліт на річний стік річок .....	157
5.3.2. Обґрунтування стохастичної моделі, яка описує вплив осушувальних меліорацій на річний стік .....	161
5.4 Приклади .....	172
5.4.1 Стохастичне моделювання рядів річного стоку .....	172
5.4.2 Модель II (моделювання сукупності випадкових величин, зв'язаних кореляцією між суміжними членами рядів) .....	176
5.4.3 Модель побутового стоку при наявності на водозборі штучних водойм .....	180
Предметний показчик .....	184
Іменний показчик .....	186
Список літератури .....	187
Додаток .....	190

## ВСТУП

Значне накопичення вихідних матеріалів спостережень за стоком річок, що відбулося в останні десятиліття, а також інтенсивне впровадження різних типів ЕОМ у практику гідрологічних розрахунків дали новий імпульс до використання математичних прийомів обробки, аналізу й узагальнення гідрологічної інформації. Чисельні експерименти, виконувані з математичними моделями, відкривають нові перспективи в дослідженнях складних ситуацій, можливих на гідрологічних об'єктах, особливо в умовах водогосподарських перетворень.

У залежності від ступеня невизначеності процесу математичні моделі поділяються на детерміністичні (генетичні) і стохастичні (ймовірнісні). Детерміністичні моделі базуються на фізичних законах. При заданих вхідних граничних умовах вони дають єдине рішення на виході. Математичні моделі, в описі яких використовуються випадкові величини, називаються ймовірнісними або *стохастичними*. Моделі цього класу засновані на застосуванні розподілів імовірностей різних характеристик і враховують імовірнісну природу досліджуваних процесів.

На початкових етапах розвитку гідрологічних розрахунків спостережені ряди річкового стоку розглядалися як сукупність незалежних випадкових величин. Початок широкого використання апарату теорії ймовірностей і математичної статистики зв'язаний з роботами А.Хазена, який запропонував для опису статистичного розподілу рядів двопараметричний нормальний розподіл (криву Гаусса). Як відомо, цей розподіл симетричний, має область припустимих значень в інтервалі від  $-\infty$  до  $+\infty$  і характеризується двома параметрами: математичним сподіванням і дисперсією. Надалі А.Форстер установив, що внаслідок взаємозалежності стокоформуючих факторів ряди стоку звичайно не є симетричними, і тому рекомендував використовувати в розрахунках стоку асиметричний розподіл Пірсона III типу. Параметрами цього розподілу є математичне сподівання  $m_x$ , середньоквадратичне відхилення  $\sigma_x$  чи коефіцієнт варіації  $C_v$ , коефіцієнт асиметрії  $C_s$ . Уже на початку широкого застосування кривої Пірсона III була звернена увага на наступний недолік: при  $C_s < 2C_v$  крива розподілу переходить в область від'ємних значень стоку. Відзначена властивість розглянутого розподілу приводить до одержання від'ємних значень стоку при екстраполяції кривих забезпеченостей у її нижній частині за межі спостережених даних.

У гідрологічній літературі наводиться ціла низка робіт, спрямованих на трансформацію широко відомих в математиці розподілів (нормального і

Пірсона III) з метою забезпечити можливість їхнього застосування до розрахунків стоку.

Найбільш ефективними в цій галузі досліджень стали перетворення Фішера, С.Н.Крицького і М.Ф.Менкеля. Напрямок, зв'язаний з логарифмічним перетворенням нормального закону розподілу, спирається на математичний аналіз, виконаний А.Фішером і реалізований у гідрології Слейдомом.

Основна ідея цього перетворення полягає в тому, що логарифмічно нормальний розподіл може мати тільки випадкова величина, яка приймає лише позитивні значення (від'ємні числа логарифми не містять). При заміні в рівнянні нормального розподілу вихідної величини  $x$  на  $\lg x$  новий розподіл виявляється асиметричним, починається з нуля і не обмежений зверху. У такий спосіб виконуються всі вимоги, які стосуються до теоретичних кривих, що описують імовірнісний розподіл стоку.

Перетворення С.Н.Крицького і М.Ф.Менкеля зв'язані з розподілом Пірсона III за умови, що  $C_s = 2C_v$ , а вихідна змінна  $x$  замінюється зв'язаною з нею величиною  $z = ax^b$ , де  $a$  і  $b$  - параметри, що залежать від величин коефіцієнтів варіації й асиметрії емпіричного ряду вихідної величини  $x$ . Новий розподіл одержав назву трипараметричного гамма-розподілу і знайшов широке застосування в практиці вітчизняної гідрології.

Подальшим кроком до ускладнення стохастичних моделей стоку стала пропозиція Юхимовича враховувати внутрішньорядні зв'язки в багаторічних коливаннях стоку. Підставою для таких досліджень служила встановлена за вибірковими даними стійкість чергування угруповань років підвищеної і зниженої водності в багаторічних коливаннях річного стоку (циклічність). Циклічність процесу річкового стоку деякі автори пояснюють тим, що при великих атмосферних опадах вода акумулюється в ґрунті і в наступні роки повертається до річок у вигляді ґрунтового стоку, підтримуючи підвищену водність. У випадку послідовності сухих років спостерігається зворотне явище. Д.Я.Раткович висунув гіпотезу, відповідно до якої однією з основних причин циклічності річного стоку є зкорельованість у рядах річного шару і випару.

Питання про вплив сонячної циклічності на коливання клімату і зв'язаного з ним стоку не знаходять визнання в деяких фахівців. У принципі всі зміни клімату й погоди є сонячнообумовленими, однак вплив сонячної активності є опосередкованим, у формуванні стоку бере участь різноманітна система природних процесів. Тому можливість існування прямих зв'язків між коливаннями стоку і сонячної активності малоімовірна.

Проте, наявність періодичної складової в змінах сонячної активності дала змогу ряду авторів висунути гіпотезу про глобальне походження



періодичних складових у річковому стоці. Періодична складова (тренд), яка виділяється в багаторічних коливаннях річного стоку, носить детермінований характер і рекомендується для прогнозу стоку на майбутнє. Однак пошуки закономірностей виникнення і розпаду циклів, спроби їхньої систематизації не привели до одержання оптимістичних результатів. Оскільки угруповання років підвищених і знижених значень стоку характерні для будь-якої стаціонарної послідовності незалежних випадкових величин, деякі учені вважають, що немає потреби в урахуванні періодичної складової через її статистичну незначущість. Саме питання про вірогідність висновків, отриманих по коротких рядах спостережень, поставив під сумнів доцільність використання методів гармонійного і спектрального аналізів в процесі виділення прихованих періодичностей річного стоку.

Таким чином, у цей час більшістю дослідників прийнята гіпотеза про те, що процес річного стоку носить імовірнісний характер і періодичної складової не містить. Сучасна стохастична модель стоку описує його як нестационарний випадковий процес, використовуючи для цього апарат теорії випадкових функцій. У більшості випадків у гідрології приймається гіпотеза про стаціонарність процесу. Строго кажучи, гіпотеза про стаціонарність не може бути перевірена експериментально, оскільки для цього необхідно мати ансамблі реалізацій стоку. У цьому випадку нестационарний процес уявляється таким, що складається з дискретних квазістаціонарних ділянок, на кожній з яких процес вважається стаціонарним і ергодичним. Іншими словами, поняття строгої стаціонарності заміняється поняттям «локальної» стаціонарності, тобто вважається, що стохастичні характеристики аналізованої реалізації не змінюються за період спостережень. У гідрологічних розрахунках такий інтервал складає 100-150 років. Доказом ергодичності процесу на цьому проміжку часу є убування кореляційних функцій з ростом аргументу.

У цьому випадку циклічність стаціонарного ергодичного процесу описується лінійними авторегресивними моделями стоку, вперше введеними в область гідрологічних розрахунків С.Н.Крицьким і М.Ф.Менкелем. Такі моделі підлягають закону великих чисел і центральній граничній теоремі. Вони дуже зручні для опису часових рядів і дозволяють відтворювати будь-яку кореляційну функцію, починаючи від простого однокрокового зв'язку (простий ланцюг) до складного багатокрокового зв'язку (складний ланцюг). Іншими словами, у практику побудови стохастичних моделей річного стоку були введені поняття теорії Марковських процесів.

Марковськими називають випадкові процеси  $\xi(t)$ , розвиток яких після будь-якого заданого моменту часу  $t = t_0$  залежить тільки від значення  $\xi(t_0)$  і не залежить від попереднього розвитку. Для опису Марковських

процесів з дискретним часом звичайно використовують поняття «ланцюг Маркова». Якщо значення випадкового процесу  $\xi(t)$  в момент часу  $t_n$  залежить лише від значення  $\xi(t)$  в попередній момент часу  $t_{n-1}$  і не залежить від значень  $\xi(t)$  у більш ранні моменти часу  $t_0, t_1, \dots, t_{n-2}$ , то такий ланцюг Маркова ідентифікується як простий.

Через те, що питання про визнання істотності далеких зв'язків у рядах річного стоку залишаються відкритими, простий ланцюг Маркова розглядається як цілком задовільне наближення для опису багаторічних коливань стоку. Таким чином, розгляд двох суміжних перетинів випадкового процесу багаторічних коливань стоку може бути зведений до розгляду спільного розподілу системи двох випадкових величин  $x_i$  і  $x_{i+1}$  (надалі позначимо  $x_i = x, x_{i+1} = y$ ). Для імовірнісного опису Марковських процесів, представлених простими ланцюгами, необхідно знати спільний двовимірний закон  $F(x, y)$  чи розподіл  $f(x, y)$ .

При підготовці цього навчального посібника був використаний досвід викладання методів статистичного аналізу гідрологічних рядів Санкт-Петербурзького державного гідрометеорологічного університету, інституту водних проблем Російської академії наук, Російського гідрологічного інституту та багатолітній особистий педагогічний досвід авторів в Одеському державному екологічному університеті.

Автори вважають своїм приємним обов'язком подякувати інженерам кафедри Арестовій Ользі Василівні і Романенко Олені Юріївні та зав. редакційним відділом університету Соколенко Олені Денисівні за корисну співпрацю при підготовці рукопису посібника.

## РОЗДІЛ 1

### СТАТИСТИЧНІ МЕТОДИ В ГІДРОЛОГІЧНИХ РОЗРАХУНКАХ

#### 1.1 Обґрунтування можливостей застосування методів математичної статистики до визначення характеристик стоку річок

Математична статистика – це прикладна математична дисципліна, яка споріднена теорії ймовірностей. Задача математичної статистики полягає у тому, щоб на основі властивостей деякої підмножини (вибірки) зробити висновки про властивості усієї множини в цілому. Уся множина в цілому має назву генеральної сукупності.

Часові ряди однорідних за своїм походженням гідрометеорологічних величин можна розглядати як вибірку з генеральної сукупності (статистичний ряд). При цьому до статистичного ряду ставиться вимога щодо незв'язності її членів. Це означає, що статистична залежність між ними має бути відсутньою. Ця вимога не завжди виконується і для урахування існуючих внутрішньорядних зв'язків при оцінках статистичних характеристик у гідрологічних розрахунках використовують основні положення теорії випадкових процесів [11,12,19,24].

***Чи можливо застосовувати методи математичної статистики до величин стоку? Це ж саме питання можна сформулювати і таким чином: чи є допустимим розглядати величини стоку як випадкові?***

Теоретичним обґрунтуванням можливості застосування статистичних методів до розрахунків величин стоку є ***закон великих чисел та центральна гранична теорема***, які ще мають назву граничних теорем теорії ймовірностей [25].

У вузькому розумінні під законом великих чисел розуміють комплекс теорем, для кожної з яких установлюється факт наближення середніх характеристик до деяких сталих. Фізичний зміст закону великих чисел зводиться до такого: індивідуальні особливості кожного випадкового явища практично не впливають на середній результат, а випадкові відхилення нівелюються. Закон великих чисел є законом сталості середніх величин. Іншими словами, закон великих чисел дозволяє зробити такі висновки: при достатньо великій кількості спостережень отримані за вибірками характеристики стоку достатні для опису генеральної сукупності у цілому. Відомі теореми, які відносяться до закону великих чисел, - це теореми Чебишева, Бернуллі та інш.

Виходячи з того, що більшість осереднених характеристик стоку набуває сталості при довжині рядів у 40-50 років і не змінюється при подальшому зростанні довжини рядів, говорить про те, що стокові ряди підлягають закону великих чисел.

Під центральною граничною теоремою розуміють групу теорем, зміст яких полягає у такому: подія, що відбувається у результаті підсумовування або визначення добутку великої кількості незалежних або малозалежних подій є випадковою і підлягає нормальному закону розподілу. Якщо величини стоку можна розглядати як результат додавання випадкових незалежних подій, то легко прийти до ствердження, що стік є випадковою величиною, яка підлягає нормальному закону розподілу. Але аналіз вихідної інформації по стоку показує, що статистичний розподіл величин стоку у більшості випадків не може бути описаний за нормальним законом розподілу. Справа у тому, що події, які приводять до формування величин стоку, залежні одна від одної, тобто між стокоформуєчими факторами у більшості випадків завжди існує зв'язок. В результаті виникає також зв'язок між попередніми та наступними членами рядів стоку. Структура таких часових рядів стоку може бути описаною за допомогою так званих автокореляційних функцій.

На відміну від метеорологічних рядів існує багато особливостей при застосуванні методів математичної статистики у гідрологічних розрахунках. Ці особливості можна представити наступним чином.

1. Першою особливістю застосування методів математичної статистики у гідрологічних розрахунках є обмежена за часом інформація. Ряди стоку короткі, їх середня довжина становить у середньому по Україні менше 40 років. У зв'язку з цим особлива увага у гідрологічних розрахунках приділяється методам оцінок статистичних параметрів за даними спостережень. Окрім того, у гідрологічних розрахунках через нестачу даних не використовуються згруповані ряди. Особливої ваги набувають питання приведення коротких гідрологічних рядів та їх статистичних параметрів до тривалого періоду спостережень, екстраполяції емпіричних кривих розподілу стоку за межі вимірювання.

2. Ряди стоку можуть бути статистично неоднорідними через вплив водогосподарських заходів, які трансформують вихідні ряди. Великого значення набувають методи прийняття чи відхилення гіпотез про статистичну однорідність рядів, наявність “викидів” та ін. Постійний вплив водогосподарських заходів на формування процесів стоку приводить до появи тенденцій монотонної зміни осереднених характеристик стоку. Наприклад, середнє арифметичне значення може зменшуватись у часі. Для описання таких рядів на основі методів математичної статистики звертаються до теорії нестационарних гідрологічних процесів.

3. Значення стоку завжди додатні. Область визначення стокових величин знаходиться у межах від 0 до нескінченності, у той час як випадкова величина, розподілена за нормальним законом розподілу ймовірностей, змінюється від  $-\infty$  до  $+\infty$ . Ряди стоку характеризуються додатною асиметрією, на відміну від нормально розподілених випадкових величин, для яких асиметрія дорівнює нулю. Виражена відмінність розподілів стокових величин від нормального закону спонукає до пошуку інших теоретичних законів розподілу випадкових величин, які б задовільно описували емпіричні закони розподілу. Найчастіше у гідрологічних розрахунках як теоретичні використовуються закони розподілу Пірсона III і трипараметричний гама-розподіл С.М. Крицького та М.Ф. Менкеля.

4. Як вже відмічалось, між суміжними членами рядів стоку може існувати зв'язок, який характеризується коефіцієнтом автокореляції  $r(1)$ . Для описання таких рядів у гідрологічних розрахунках використовують теорію випадкових функцій. Наявність внутрішньорядної зв'язності, яка порушує принцип випадковості, зменшує об'єм незалежної інформації, тому у гідрологічних розрахунках розроблені методики врахування впливу внутрішньорядної кореляції на статистичні оцінки, отримані відомими у математичній статистиці методами.

5. Задачі гідрологічних розрахунків по визначенню характеристик стоку є ретроспективно-описувальними та прогностичними. На початку необхідно виявити закономірності формування явища, побудувати математичну модель і вже за математичною моделлю передбачувати майбутній режим річок та тимчасових водотоків. Таке передбачення у гідрологічних розрахунках надається в ймовірнісній формі.

Математичні моделі можна класифікувати за ступенем їхньої невизначеності – на детерміністичні і стохастичні. Детерміністична модель не містить у собі випадкових компонент, а стохастична – їх поєднує. У існуючих стохастичних моделях ряди стоку розглядаються або як *послідовність незалежних випадкових величин*, або за стохастичу модель береться та чи інша модифікація *простого ланцюга Маркова*, де стік розглядається як послідовність залежних випадкових величин. Ступінь залежності характеризується коефіцієнтом автокореляції, а як математичний апарат виступає закон двовимірного розподілу випадкових величин [22-24].

## 1.2 Основні положення теорії ймовірностей

### 1.2.1 Поняття ймовірності. Випадкова подія

Як зазначалося, процес, відображенням якого є гідрологічні послідовності вимірювань, може бути описаний як імовірнісний або стохастичний. Перед тим, як перейти до визначення ймовірності, введемо поняття випадкової події.

Випадкова подія – факт, який в результаті експерименту (випробування) може відбутись, а може й не відбутись. Випадковою подією можуть бути такі характеристики:

*A* – число сонячних днів за тиждень;

*B* – тривалість пересихання річок за період спостережень;

*C* – кількість дощових паводків на протязі року.

Розглядаючи ці події, приходимо до висновку, що кожна з них має ту або іншу можливість своєї появи: одна – більшу, друга – меншу. Для того, щоб кількісно порівняти між собою події за ступенем можливості їх появи, необхідно з кожною подією зв'язати відповідне число, яке тим більше, чим більш імовірна подія. Таке число має назву ймовірності події. Тобто, **ймовірність події – чисельна міра ступеня об'єктивної можливості здійснення цієї події.**

Для співвідношення різних подій за ступенем можливості їхнього здійснення встановлюють одиницю вимірювання. Імовірність достовірної події (**достовірна подія – подія, яка в результаті експерименту обов'язково має відбутися**) приймається рівною одиниці, ймовірність недостовірної події (**недостовірна подія** - така подія, яка в даному досліді ніяк не може відбутися) дорівнює 0.

Наприклад, ймовірність такої події як наявність протягом тижня не більше семи сонячних днів дорівнює одиниці. У свою чергу, наявність в календарному тижні більше семи сонячних днів має ймовірність, що дорівнює нулю.

Імовірності інших подій знаходяться у інтервалі від 0 до 1. Сума ймовірностей розглядуваної випадкової події завжди дорівнює 1.

### 1.2.2 Випадкові величини та закони їх розподілу

Якщо в результаті випробувань розглядається не сама подія, а її кількісна характеристика, то в теорії ймовірностей використовується **поняття** випадкової величини. **Випадковою називається величина, яка внаслідок випробування набуває того чи іншого значення, наперед невідомо, якого саме.** Якщо можливі значення випадкової величини можна

перерахувати, то випадкова величина відноситься до дискретної, тобто випадкова величина описується кінцевим числом значень. Приклади дискретних випадкових величин: 1) число випадків пересихання річки за літньо-меженний період; 2) кількість дощових паводків у році; 3) число днів у році з витратами води вище або менше будь-якого рівня. Дискретна випадкова величина  $X$ , з можливими значеннями якої  $x_1, x_2, x_3, \dots, x_N$  з імовірнісної точки зору буде повністю описана, якщо кожному значенню випадкової величини поставити у відповідність значення ймовірності їх появи у випробуванні. Наприклад, у результаті  $n$  випробувань випадкова величина приймає  $n$  значень:  $x_1, x_2, x_3, \dots, x_N$  з відповідними ймовірностями  $p_1, p_2, p_3, \dots, p_N$ . Законом розподілу випадкової дискретної величини називається співвідношення, яке установлює зв'язок між можливими значеннями випадкової величини та їх імовірностями. Такий закон розподілу називається таблицею розподілу, наприклад,

$$x_i: \quad x_1 \quad x_2 \quad x_3 \quad \dots \quad x_N$$

$$p_i: \quad p_1 \quad p_2 \quad p_3 \quad \dots \quad p_N$$

$$\text{Сума ймовірностей } \sum_{i=1}^N p_i = 1.$$

Якщо таблицю розподілу представити у графічному вигляді (рис.1.1), то ми отримаємо багатокутник розподілу.

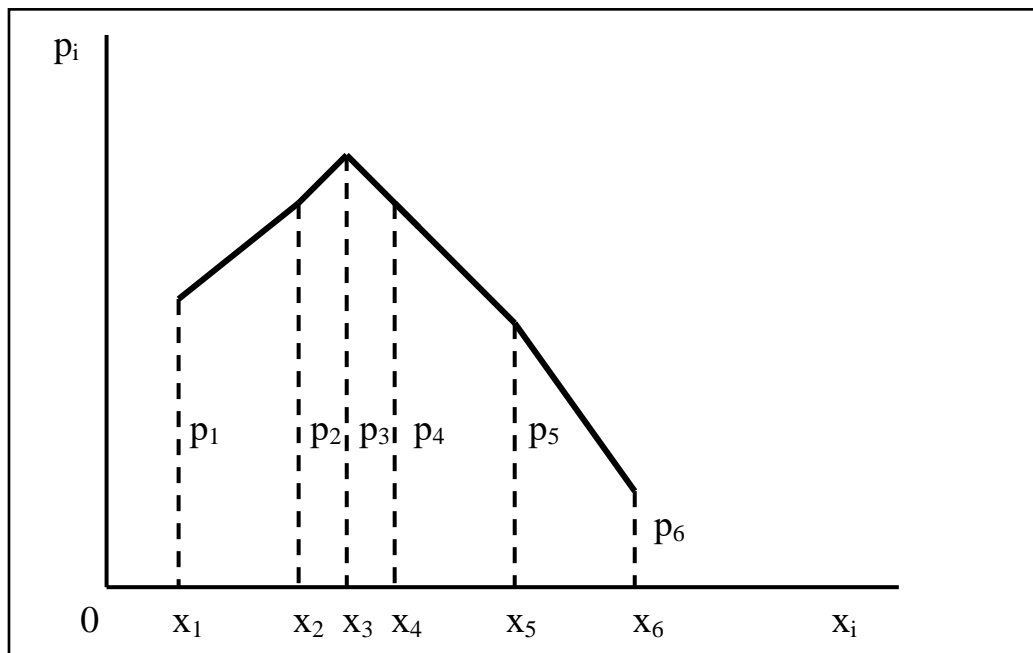


Рисунок 1.1 – Багатокутник розподілу

Закон розподілу, представлений у вигляді таблиці або багатокутника, є вичерпною ймовірнісною характеристикою дискретної випадкової величини.

Визначити ймовірність появи кожного значення випадкової дискретної величини у результаті проведення  $n$  випробувань або спроб можливо, якщо представити їх як відношення числа випадків, коли розглядувана величина спостерігалася, до загальної кількості випробувань

$$p_i = \frac{m}{N}, \quad (1.1)$$

де  $p_i$  - ймовірність появи випадкової величини;

$i$  - номер випробування;

$m$  - число випадків, коли величина спостерігалася у результаті проведення випробувань;

$N$  - загальна кількість величин, які описують дискретну випадкову величину  $X$ .

Поряд з дискретними величинами існують і безперервні. Випадкові величини, які заповнюють деякий проміжок на осі  $x$ , називають безперервними. Кожне окреме значення безперервної випадкової величини не володіє ніякою, відмінною від нуля, ймовірністю. Для того, щоб задати закон розподілу безперервної випадкової величини, вводиться поняття функції розподілу  $F(x) = p(X < x)$ , де  $F(x)$  - ймовірність події  $X=x$ , а ймовірність того, що  $X < x$ , де  $x$  - деяка поточна змінна.

### **1.2.2.1 Інтегральний закон розподілу безперервних випадкових величин**

Для описування закону розподілу безперервної випадкової величини використовується інтегральна функція розподілу ймовірної величини  $F(x) = P(X < x)$ . Функція розподілу  $F(x)$  має такі властивості:

1. Вона монотонно неспадна, тобто при  $x_2 > x_1$ ,  $F(x_2) > F(x_1)$  (рис.1.2).
2. На мінус нескінченності, коли  $x \rightarrow -\infty$ , то  $F(-\infty) = 0$
3. На плюс нескінченності, коли  $x \rightarrow +\infty$ , то  $F(+\infty) = 1$

У гідрологічних розрахунках здебільшого використовується не інтегральна функція розподілу  $F(x)$ , а функція забезпеченості  $P(x)$ , яка пов'язана з інтегральною функцією співвідношенням

$$P(x) = 1 - F(x). \quad (1.2)$$



Забезпеченість випадкової величини  $x$  – це ймовірність того, що випадкова величина  $X$  більше деякого заданого значення  $x$ , тобто  $P(x)=p(X>x)$ .

Для дискретних випадкових величин функція забезпеченості може бути представленою у вигляді

$$F(x)=p(x<X)=\sum_{x_i<x}p_i(X=x_i), \quad (1.3)$$

де нерівність  $x_i < x$  указує на те, що підсумовування відбувалося не для усіх  $x_i$ , а лише для тих, які менше  $x$ .

Таким чином, функція розподілу дискретної випадкової величини  $x$  може бути представленою у вигляді ступінчастої функції (рис.1.3). Інтегральна функція, яка представлена у графічному вигляді, має назву інтегральної кривої розподілу (рис.1.2, рис.1.3).

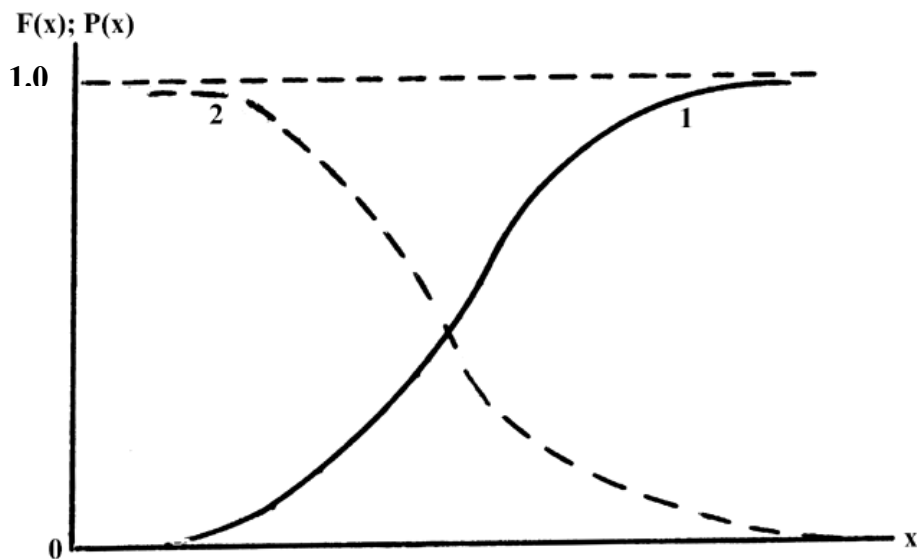


Рисунок 1.2 – Інтегральна функція розподілу  $F(x)$  випадкової величини  $x$  (1) та функція забезпеченості  $P(x)$  (2)

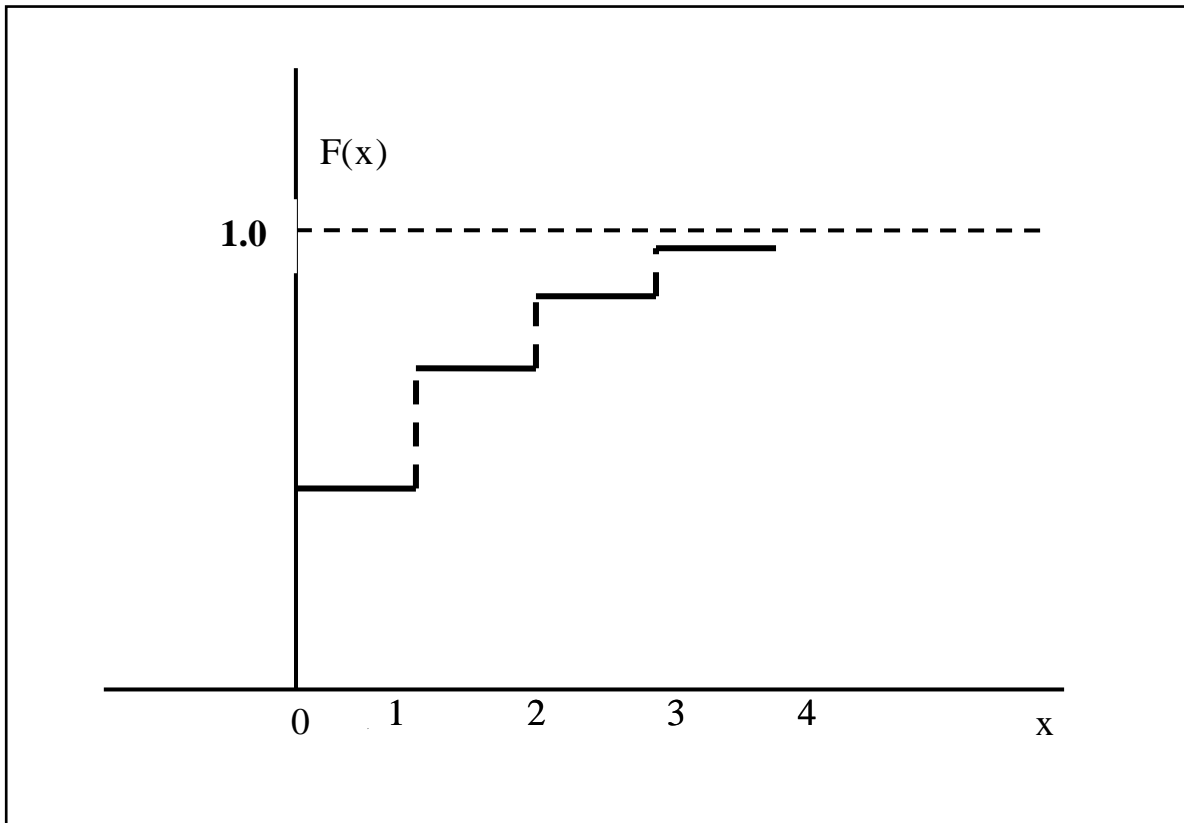


Рисунок 1.3 - Інтегральна функція розподілу  $F(x)$  випадкової величини  $x$ , визначена за дискретними даними

### 1.2.2.2 Диференціальний закон розподілу безперервної випадкової величини

Розглянемо безперервну випадкову величину  $X$ , яка має властивість монотонності. Імовірність попадання безперервної випадкової величини на ділянці від  $x$  до  $x + \Delta x$  на осі  $x$  становить

$$p(x < X < x + \Delta x) = F(x + \Delta x) - F(x), \quad (1.4)$$

де права частина є приріст функції розподілу на ділянку довжиною  $\Delta x$ . Якщо (1.4) розділити на  $\Delta x$ , то отримаємо середню ймовірність, яка приходить на одиницю довжини цієї ділянки. При наближенні  $\Delta x$  до нуля перейдемо до похідної від функції розподілу в точці  $x$

$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = f'(x), \quad (1.5)$$

звідки  $F'(x) = f(x)$ .

Похідна  $F'(x)$  носить назву щільності розподілу безперервної випадкової величини  $f(x)$  і є диференціальним законом розподілу. Ця формула існує тільки для безперервних рядів випадкових величин.

Згідно з виразом (1.5) можна вирішити протилежну задачу, тобто виразити функцію розподілу через щільність. Якщо

$$F(x) = p(X < x) = p(-\infty < X < x), \quad (1.6)$$

то

$$F(x) = \int_{-\infty}^{+x} f(x) dx \quad (1.7)$$

Диференційована функція розподілу безперервної випадкової величини  $x$  має наступні властивості:

1. Щільність розподілу завжди невід'ємна, тобто  $f(x) \geq 0$ ,
2. Інтеграл від функції щільності ймовірностей у межах від  $-\infty$  до  $\infty$  дорівнює одиниці

$$\int_{-\infty}^{+\infty} f(x) dx = 1. \quad (1.8)$$

3. Щільність ймовірності – це величина розмірна і її розмірність зворотна розмірності величини  $x$ .

З геометричної точки зору перераховані властивості вказують на те, що крива розподілу функції щільності ймовірності лежить не нижче осі абсцис та що повна площа, обмежена кривою  $f(x)$  і віссю абсцис, дорівнює 1 (рис.1.4).

Між інтегральною, диференційованою функціями та функцією забезпеченості існують певні співвідношення, які можна представити таким чином

$$p(\alpha < X < \beta) = F(\alpha) - F(\beta) = F(\beta) - F(\alpha), \quad (1.9)$$

$$p(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx, \quad (1.10)$$

$$P(x) = 1 - \int_{-\infty}^x f(x) dx = \int_x^{\infty} f(x) dx . \quad (1.11)$$

З (1.9) – (1.11) витікає, що знаючи закон розподілу однієї з перерахованих характеристик, можна перейти до закону розподілу іншої характеристики.

*У загальному випадку закон розподілу випадкової величини – це співвідношення між випадковою величиною та її ймовірнісною характеристикою. Як ймовірнісні характеристики можуть бути використані ймовірність, інтегральна функція розподілу, щільність розподілу, забезпеченість.*

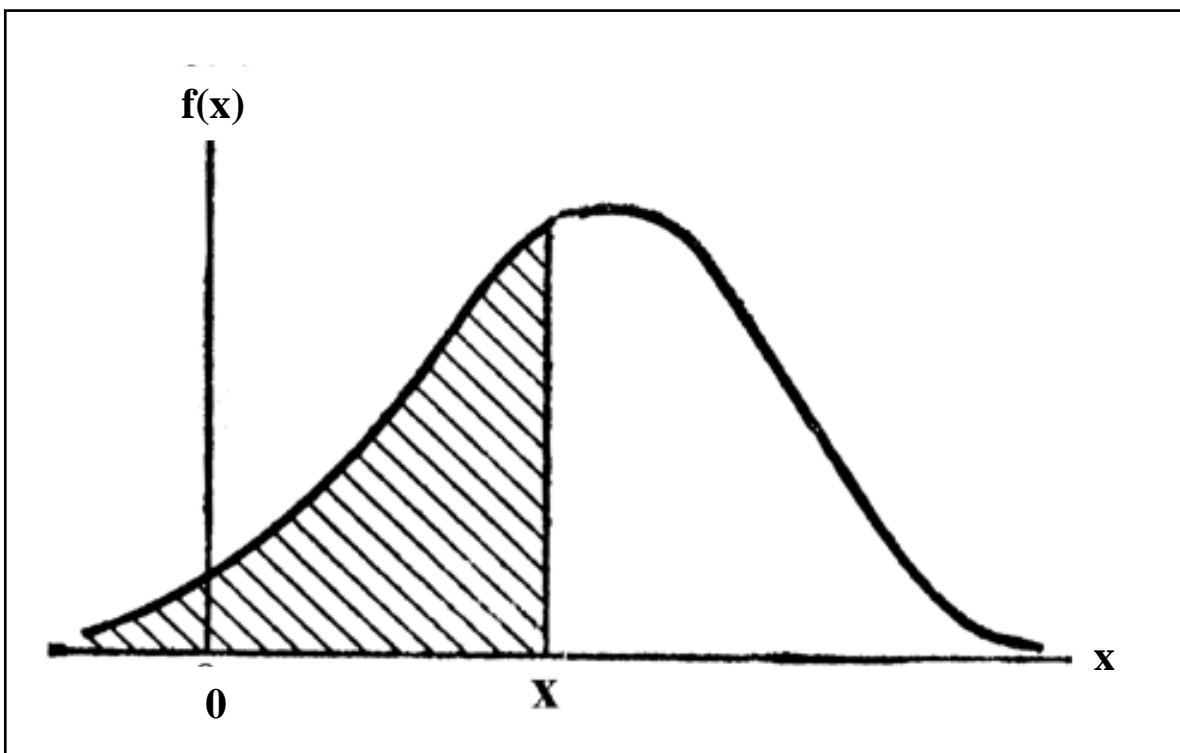


Рисунок 1.4 – Крива розподілу  $f(x)$  випадкової величини  $x$ .  
(Площа під кривою, що лежить лівіше перетину  $x$ , є функцією розподілу  $F(x)$ . Загальна площа під кривою розподілу дорівнює одиниці.)

### 1.2.2.3 Характеристики положення кривих розподілу

При виконанні гідрологічних розрахунків однією з основних задач є описування ймовірнісних властивостей випадкової величини на основі кривих розподілу. Для уяви про характер розподілу достатньо знати лише

числові параметри кривих розподілу: математичне сподівання випадкової величини, моду, медіану, варіацію, асиметрію, ексцес випадкових величин та інше.

**Математичне сподівання** випадкової величини  $m_x$  є центром статистичного розподілу, відносно якого групуються члени сукупності та являє собою середнє зважене по ймовірності значення випадкової величини  $X$ , тобто

$$m_x = \frac{x_1 p_1 + x_2 p_2 + \dots + x_N p_N}{p_1 + p_2 + \dots + p_N} = \frac{\sum_{i=1}^N x_i p_i}{\sum_{i=1}^N p_i}. \quad (1.12)$$

Враховуючи, що для дискретних випадкових величин  $\sum_{i=1}^N p_i = 1$ , з (1.12)

отримаємо

$$m_x = \sum_{i=1}^N x_i p_i. \quad (1.13)$$

Математичне сподівання для безперервної випадкової величини може бути представлене у вигляді

$$m_x = \int_{+\infty}^{-\infty} x f(x) dx, \quad (1.14)$$

де  $f(x)dx$ - елемент імовірності, який несе те саме смислове навантаження, що й  $p_i$  для дискретних випадкових величин.

Оскільки на практиці фахівці використовують обмежені у часі вибірки – статистичні ряди - , для яких період спостережень  $n$  значно менший за довжину генеральної сукупності  $N$ , значення ймовірності появи заданого значення випадкової величини замінюється на відносну частоту

$$p_i^* = \frac{m_i}{n}.$$

Слід зазначити, що статистичні параметри, розраховані за даними вибірок, являють собою лише наближені значення параметрів генеральної сукупності і називаються “оцінками” таких параметрів.

Оцінкою або емпіричним еквівалентом математичного сподівання, розрахованого за вибіркою, є середнє арифметичне значення  $\bar{x}$

$$m_x^* = \bar{x} = \sum_{i=1}^n x_i \frac{m_i}{n}, \quad (1.15)$$

де  $n$  - загальне число незалежних випробувань(спостережень);

$m_1, m_2, \dots, m_n$  - частоти випадкових величин  $x_1, x_2, x_3, \dots, x_n$ ;

$\frac{m_i}{n}$  - відносні частоти або емпіричні ймовірності  $p_i^*$ .

Якщо  $n$  прагне до  $N$ , то  $p_i^*$  прагне до свого теоретичного значення  $p_i$ , а  $\bar{x}$  - до математичного сподівання  $m_x$ .

Якщо частоти  $m_1, m_2, \dots, m_n$  дорівнюють одиниці, тобто кожне значення випадкової величини повторюється в досліді один раз, то

$m_1 = m_2 = \dots = m_n = \frac{1}{n}$ . Відповідно,

$$m_x^* = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1.16)$$

Середнє арифметичне значення має ту саму розмірність що й величина, по котрій воно розраховано.

**Модою випадкової величини**  $m_0$  називають найбільш імовірне її значення. На графіку - це найбільша ордината кривої розподілу. Розрізняють одно- та багатомодальні розподілення. Розподіл гідрологічних величин є одномодальним. Для безперервної випадкової величини мода - це випадкова величина з найбільшою щільністю ймовірності, для дискретної - з найбільшою ймовірністю.

**Медіана випадкової величини**  $m_e$  - це значення випадкової величини, яке ділить площу під кривою розподілу порівну, тобто

$$\int_{-\infty}^{m_e} f(x) dx = \int_{m_e}^{+\infty} f(x) dx. \quad (1.17)$$

Для дискретних випадкових величин береться правило, що коли число членів ряду  $x_i$  непарне та дорівнює  $2m+1$ , то медіана ряду буде дорівнювати значенню  $x_{m+1}$ , яке є середнім членом ранжированого ряду, тобто  $m_e = x_{m+1}$ . Якщо число членів ряду парне, тобто дорівнює  $2m$ , то за медіану береться середнє значення між центральними величинами ранжированого ряду.

### 1.3 Статистичні характеристики стокових величин, методи їх визначення та особливості практичного застосування

Основною задачею гідрологічних розрахунків є визначення величин стоку із заданими ймовірнісними характеристиками. Найчастіше використовується така ймовірнісна характеристика як забезпеченість (ймовірність перевищення заданої величини). Для того, щоб отримати величини стоку заданої забезпеченості, необхідно підібрати теоретичний закон розподілу, який найточніше описує емпіричний розподіл стокових величин, та визначити статистичні параметри, необхідні для практичного застосування цього закону. *Під статистичними параметрами розуміють параметри законів розподілу, які описують основні властивості статистичного розподілу випадкової величини.*

Величини стоку завжди позитивні, а верхня границя їх значень не обмежена, отже, областю визначення теоретичного закону розподілу величин стоку повинна бути область  $(0, \infty)$ .

Як правило, ряди спостережень за стоком здебільшого короткі, тому кількість статистичних параметрів, які використовуються, не повинна бути великою.

Світовий досвід показує, що при розрахунках стоку найбільш оптимальними для практичного застосування є такі теоретичні закони розподілу, для описування яких достатньо двох або трьох статистичних параметрів – математичного сподівання, дисперсії, а також коефіцієнтів варіації та асиметрії.

Статистичні параметри стоку визначаються по рядах спостережень фазовооднорідних стокових величин. Наприклад, статистичні параметри можна розраховувати по рядах максимального стоку весняного водопілля, річного стоку, максимального стоку дощових паводків та ін. У сучасній гідрології, як правило, використовуються три найбільш поширені методи визначення статистичних параметрів – моментів, найбільшої правдоподібності, графо-аналітичний. Перший з цих методів не пов'язаний ні з одним із законів розподілу випадкової величини.

*Статистичні параметри, визначені по рядах спостережень, відрізняються від відповідних параметрів генеральної сукупності і називаються оцінками статистичних параметрів.*

При відсутності чи недостатності даних спостережень для визначення статистичних параметрів розроблено додаткові спеціальні методи, які викладаються у наступних розділах.

### 1.3.1 Метод моментів. Статистичні моменти як характеристики кривих розподілу

Поняття моментів перенесено в статистику із розділу фізики “Механіка”, де момент являє собою добуток сили на плече. Плече – відстань від точки, у якій прикладено силу, до точки опори. Значення дискретної випадкової величини розглядається як матеріальна точка на числовій осі з масою, пропорційною ймовірності появи цієї випадкової величини. Якщо плечем є відстань від нуля числової осі до матеріальної точки, то такі статистичні моменти називаються початковими. Коли ж для визначення статистичного моменту береться відстань від математичного сподівання до розглядуваної матеріальної точки, то статистичний момент отримує назву центрального.

Для описування властивостей кривих розподілу широко використовують початкові та центральні моменти. Початкові моменти  $s$ -го порядку дискретної величини  $X$  являють собою суму

$$\alpha_s = \sum_{i=1}^N x_i^s p_i. \quad (1.18)$$

Для безперервної випадкової величини сума записується через інтеграл

$$\alpha_s = \int_{-\infty}^{\infty} x^s f(x) dx. \quad (1.19)$$

Таким чином, для дискретних випадкових величин оцінка початкових моментів  $s$ -того порядку розраховується за формулою

$$\hat{\alpha}_s = \sum_{i=1}^n x_i^s p_i^*, \text{ де } p_i^* = \frac{1}{n}. \quad (1.20)$$

Перший початковий момент ( $s=1$ ) представляє собою математичне сподівання

$$\alpha_1 = m_x = \sum_{i=1}^N x_i p_i, \quad (1.21)$$



а його оцінка – середнє арифметичне значення  $\bar{x}$

$$\hat{\alpha}_1 = \hat{m}_x = \bar{x} = \sum_{i=1}^n x_i p_i^* = \frac{\sum_{i=1}^n x_i}{n}. \quad (1.22)$$

Центральний момент  $s$ -го порядку  $\beta_s$  дискретних випадкових величин описується наступним виразом

$$\beta_s = \sum_{i=1}^N (x_i - m_x)^s p_i; \quad (1.23)$$

відповідно для безперервних випадкових величин -

$$\beta_s = \int_{-\infty}^{\infty} (x - m_x)^s f(x) dx. \quad (1.24)$$

Оцінка центрального моменту  $s$ -го порядку виконується в такий спосіб

$$\hat{\beta}_s = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^s, \quad \text{де } p_i^* = \frac{1}{n}. \quad (1.25)$$

При  $s=1$

$$\beta_s = \sum_{i=1}^n (x_i - m_x) p_i = \sum_{i=1}^n x_i p_i - m_x \sum_{i=1}^n p_i = m_x - m_x = 0, \quad (1.26)$$

тобто

$$\hat{\beta}_s = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) = \frac{1}{n} \sum_{i=1}^n x_i - \frac{1}{n} \sum_{i=1}^n \bar{x} = 0. \quad (1.27)$$

Таким чином, перший центральний момент дорівнює нулю.

При  $s=2$

$$\beta_2 = \sum_{i=1}^N (x_i - m_x)^2 p_i \quad (1.28)$$

або

$$\hat{\beta}_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (1.29)$$

Другий центральний момент характеризує розсіювання випадкової величини відносно її центра розподілу і носить назву дисперсії  $D_x$

$$\beta_2 = D_x = \sigma_x^2. \quad (1.30)$$

Квадратний корінь із дисперсії називається середнім квадратичним відхиленням  $\sigma_x$ . Оцінка середнього квадратичного відхилення розраховується наступним чином

$$\hat{\sigma}_x = \sqrt{D_x} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}. \quad (1.31)$$

Оцінки другого центрального моменту мають від'ємні зміщення (заниження). Для їх усунення в (1.31) вводиться поправочний коефіцієнт  $\sqrt{n/(n-1)}$ , з урахуванням якого (1.31) набирає вигляду

$$\hat{\sigma}_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}. \quad (1.32)$$

З метою порівняння мінливості різномасштабних випадкових величин використовують безрозмірну характеристику  $C_v = \sigma_x / m_x$ , яка має назву коефіцієнта варіації і за даними спостережень оцінюється таким чином

$$\hat{C}_v = \frac{\hat{\sigma}_x}{\bar{x}}. \quad (1.33)$$

У більш загальному вигляді оцінка коефіцієнта варіації знаходиться за виразом

$$\widehat{C}_v = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\bar{x}^2 (n-1)}} = \sqrt{\frac{\sum_{i=1}^n (k_i - 1)^2}{n-1}}, \quad (1.34)$$

де  $k_i = x_i / \bar{x}$  - модульний коефіцієнт.

При  $s=3$  з використанням (1.23) та (1.25) отримаємо вирази для розрахунків третього центрального моменту для генеральної сукупності та вибірки

$$\beta_3 = \sum_{i=1}^N (x_i - m_x)^3 p_i, \quad (1.35)$$

$$\widehat{\beta}_3 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3. \quad (1.36)$$

Третій центральний момент характеризує несиметричність (асиметричність) розподілу випадкової величини відносно математичного сподівання і може бути як від'ємним, так і позитивним.

Нормування  $\beta_s$  по  $\sigma_x^3$  дозволяє отримати безрозмірний параметр статистичного розподілу, названий коефіцієнтом асиметрії  $C_s$

$$C_s = \frac{\beta_3}{\sigma_x^3}, \quad (1.37)$$

який при розрахунках за вибірками представляється у вигляді

$$\widehat{C}_s = \widehat{\beta}_3 / \widehat{\sigma}_x^3 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n \widehat{\sigma}_x^3}. \quad (1.38)$$

Вираз (1.38) при використанні модульних коефіцієнтів набуває вигляду

$$\widehat{C}_s = \frac{\sum_{i=1}^n (k_i - 1)^3}{n \widehat{C}_v^3}. \quad (1.39)$$

Як і оцінка параметра  $C_v$ , оцінка коефіцієнта асиметрії за виразом (1.39) є зміщеною відносно відповідного параметра генеральної сукупності. Від’ємна зміщеність може бути усунена шляхом введення поправкового коефіцієнта  $n^2 / (n-1)(n-2)$ , запропонованого Є.Г. Блохіновим.

Таким чином, кінцевий вигляд формули для розрахунків коефіцієнта асиметрії має наступний вигляд

$$\hat{C}_s = \frac{n}{(n-1)(n-2)} \frac{\sum_{i=1}^n (k_i - 1)^3}{\hat{C}_v^3}. \quad (1.40)$$

Для випадкової величини, яка підлягає нормальному закону розподілу,  $C_s = 0$ , тобто крива щільності ймовірностей симетрична відносно математичного сподівання.

Четвертий центральний момент ( $s=4$ ), покладений в основу характеристики гостровершинності кривої розподілу випадкової величини,  $E$  – ексцес. Формула для визначення четвертого центрального моменту за вибіркою набуває вигляду

$$\hat{\beta}_4 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4. \quad (1.41)$$

Четвертий центральний момент характеризує гостровершинність кривої розподілу у порівнянні з кривою нормального розподілу. Для випадкової величини з нормальним законом розподілу співвідношення  $\beta_4 / \sigma_x^4$  завжди дорівнює 3. Таким чином, нормування  $\beta_4$  по  $\sigma_x^4$  дозволяє отримати безрозмірний параметр, названий ексцесом

$$E = \frac{\beta_4}{\sigma_x^4} - 3 \quad (1.42)$$

або

$$\hat{E} = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n\hat{\sigma}_x^4} - 3. \quad (1.43)$$

Якщо  $E > 0$ , то крива розподілу витягнута відносно нормального закону розподілу, для якого  $E = 0$ . Коли ж  $E < 0$ , крива розподілу приплюснута по відношенню до кривої нормального розподілу (рис.1.5).

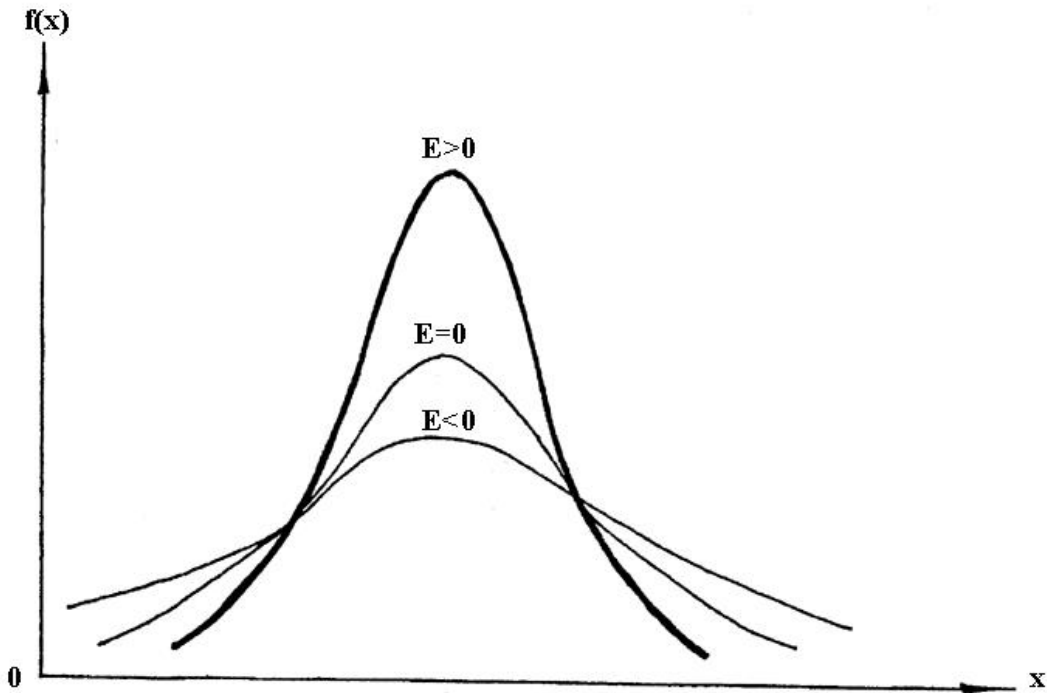


Рисунок 1.5 - Криві розподілу з різними значеннями ексцесу  $E$

### **1.3.2 Точність розрахунків статистичних параметрів за вибірковими даними**

#### ***1.3.2.1 Вимоги до оцінок статистичних параметрів***

Статистичні параметри, розраховані за вибірками, називаються вибірковими оцінками статистичних параметрів. Значення вибіркових оцінок відрізняються від значень відповідних параметрів генеральної сукупності.

Для того, щоб по оцінках достатньо вірогідно характеризувати параметри генеральної сукупності, ці оцінки повинні задовольняти вимоги незсуненості, ефективності та умотивованості.

Припустимо, що з генеральної сукупності отримано  $m$  вибірок, за якими отримано  $m$  оцінок параметрів.

Оцінка статистичного параметра називається незсуненою, якщо її математичне сподівання дорівнює параметру генеральної сукупності

$$M[\hat{\theta}_m] = \theta, \quad (1.44)$$

де  $M$  - позначка математичного сподівання;

$\hat{\theta}_m$  - оцінка статистичного параметра;

$\theta$  - значення параметра генеральної сукупності.

Вимога незсушеності гарантує відсутність систематичних похибок при оцінках параметрів.

Незсунена оцінка, яка має найменшу дисперсію серед усіх можливих незсунених оцінок параметра, розрахованих по вибірках одного і того ж об'єму, називається ефективною, тобто повинна виконуватись умова

$$D[\hat{\theta}_m] \rightarrow 0 \text{ при } m \rightarrow \infty. \quad (1.45)$$

Оцінка параметра називається умотивованою, якщо вона по ймовірності збігається до параметра генеральної сукупності

$$\lim_{n \rightarrow \infty} (P\{|\hat{\theta}_m - \theta| < \varepsilon\}) = 1, \quad (1.46)$$

де  $\varepsilon$  - мале позитивне число;

$\hat{\theta}_m$  - вибіркова оцінка параметра;

$\theta$  - значення параметра генеральної сукупності.

У методі моментів вимогам незсушеності, ефективності та умотивованості відповідає тільки середнє арифметичне значення.

Як було зазначено вище, для усунення систематичних похибок при розрахунку параметрів  $\sigma_x$ ,  $C_v$ ,  $C_s$  вводилися відповідні поправкові коефіцієнти  $\sqrt{n/(n-1)}$  - при обчисленні середньоквадратичного відхилення та коефіцієнта варіації, а також  $n^2/(n-1)(n-2)$  - при обчисленні коефіцієнта асиметрії  $C_s$ .

У зв'язку зі зміщенням параметрів, метод моментів рекомендовано вживати при  $C_s < 0.5$ .

Особливістю методу моментів є той факт, що при розрахунках за цим методом збільшується внесок значних відхилень від центра розподілу.

### ***1.3.2.2 Поняття про систематичні та випадкові похибки визначення статистичних параметрів***

Похибки вибіркового параметрів можна поділити на систематичні та випадкові. Систематичні похибки виникають за рахунок факторів, які однаково впливають при багатократному повторенні вимірювань. Випадкові похибки виникають в результаті впливу комплексу факторів, кожний з яких урахувати неможливо, що приводить при кожному вимірюванні до різних похибок як за числовим значенням, так і за знаком. Систематичні похибки не можуть бути вилучені або зменшені при багатократному вимірюванні. У той же час випадкові похибки зменшуються при збільшенні числа вимірювань однієї і тієї ж величини, через те, що середнє арифметичне таких вимірювань у середньому завжди ближче до значення генеральної сукупності. Систематичні похибки, коли вони відомі, можуть бути усунені або шляхом зміни методу вимірювань, або шляхом введення у розрахункові формули відповідних поправкових коефіцієнтів. Випадкові ж похибки можна лише оцінити. При цьому постає питання про те, як саме виконується оцінка випадкових похибок.

### ***1.3.2.3 Метод Монте-Карло***

#### ***Загальне представлення про метод***

Метод Монте-Карло – це чисельний метод вирішення математичних задач за допомогою моделювання випадкових величин.

*Походження методу Монте-Карло.* Датою народження методу Монте-Карло прийнято вважати 1949 рік, коли з'явилася стаття за назвою «The Monte Carlo method». Творцями цього методу вважають американських математиків Дж. Неймана і С. Улама.

Цікаво, що теоретична основа методу була відома вже давно. Більш того, деякі задачі статистики розраховувалися іноді за допомогою випадкових вибірок, тобто фактично за методом Монте-Карло. Однак до появи електронно-обчислювальних машин (ЕОМ) цей метод не міг знайти скільки-небудь широкого застосування, тому що моделювати випадкові величини – дуже трудомістка робота. Таким чином, виникнення методу Монте-Карло як універсального чисельного методу стало можливим тільки завдяки появі ЕОМ.

Сама назва «Монте-Карло» походить від міста Монте-Карло в князівстві Монако, знаменитого своїми ігорними будинками. Справа в тім, що одним з найпростіших механічних приладів для одержання випадкових величин є ... рулетка.

*Приклад.* Для того, щоб було більш зрозуміло, про що піде мова, розглянемо дуже простий приклад. Припустимо, що потрібно обчислити площу плоскої фігури  $S$ . Це може бути зовсім довільна фігура з криволінійною межею, задана графічно чи аналітично. Нехай це буде фігура, зображена на рис. 1.6. Припустимо, що вона вся розташована усередині одиничного квадрата.

Виберемо в квадраті  $N$  випадкових точок. Позначимо через  $N'$  число точок, які потрапили при цьому всередину  $S$ . Геометрично очевидно, що площа  $S$  приблизно дорівнює відношенню  $N'/N$ . Чим більше буде  $N$ , тим більше буде точність цієї оцінки.

У прикладі, зображеному на рис. 1.6, представлено  $N = 40$  точок. З них  $N' = 12$  точок виявилися всередині  $S$ . Відношення  $N'/N = 12/40 = 0.30$ , у той час як площа  $S$  дорівнює 0,35.

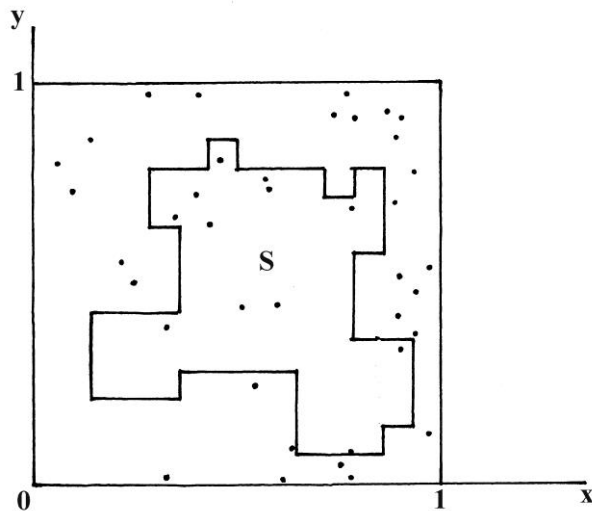


Рисунок 1.6 – Фігура  $S$

*Дві особливості методу Монте-Карло.* Перша особливість методу – проста структура обчислювального алгоритму. Як правило, складається програма для здійснення одного випадкового випробування (у прикладі потрібно вибрати випадкову точку в квадраті і перевірити, чи належить вона  $S$ ). Потім це випробування повторюється  $N$  раз, причому кожне випробування не залежить від всіх інших і результати всіх випробувань осереднюються. Тому іноді метод Монте-Карло називають *методом статистичних випробувань*.

Друга особливість методу: похибка обчислень, як правило, пропорційна  $\sqrt{D/N}$ , де  $D$  - деяка постійна, а  $N$  - число випробувань. З цієї формули видно, що для того, щоб зменшити помилку в 10 разів



(інакше кажучи, щоб одержати у відповіді ще один вірний десятковий знак), потрібно збільшити  $N$  ( тобто обсяг вибірки) у 100 разів.

Ясно, що домогтися високої точності таким шляхом неможливо. Тому звичайно говорять, що метод Монте-Карло особливо ефективний при вирішенні тих задач, які мають результат з точністю 5-10%.

Однак ту ж саму задачу можна вирішувати різними варіантами методу Монте-Карло, яким відповідають різні значення  $D$ . У багатьох задачах удається значно збільшити точність, вибравши спосіб розрахунку, якому відповідає значно менше значення  $D$ .

*Задачі, розв'язувані методом Монте-Карло.* По-перше, метод Монте-Карло дозволяє моделювати будь-який процес, на протікання якого впливають випадкові фактори. По-друге, для багатьох математичних задач, не зв'язаних з якими-небудь випадками, можна штучно придумати ймовірнісну модель (і навіть не одну), яка дозволяє вирішувати ці задачі.

Таким чином, можна говорити про метод Монте-Карло як про універсальний спосіб вирішення математичних задач.

### ***Одержання випадкових величин на ЕОМ***

Сама постановка питання «одержання випадкових величин на ЕОМ» іноді викликає здивування: адже все, що робить машина, повинно бути заздалегідь запрограмованим; відкіля ж може з'явитися випадковість?

У цьому питанні і справді є деякі труднощі, але вони відносяться скоріше до філософії, так що зупинятися на них немає необхідності.

Про всякий випадок помітимо тільки, що випадкові величини – це ідеальні математичні поняття. Питання про те, чи можна з їхньою допомогою описати яке-небудь явище природи, зважається досвідом. Такий опис завжди є наближеним. Більш того, випадкова величина, яка цілком задовільно описує якусь фізичну величину в одному колі явищ, може виявитися поганою характеристикою цієї ж величини при дослідженні інших явищ. Точно так дорога, яку на карті країни можна вважати прямою (ідеальною математичною прямою «без ширини»), постає смугою з вигинами на великомасштабному плані населеного пункту.

Звичайно розрізняють три способи одержання випадкових величин: таблиці випадкових чисел, генератори випадкових чисел і метод псевдовипадкових чисел.

*1 Таблиці випадкових чисел.* Проробимо наступний дослід. Напишемо на десятих однакових папірцях цифри 0, 1, 2, ..., 9... Покладемо ці папірці в шапку, перемішаємо і будемо витягати відтіля по одному папірцю, щораз повертаючи його назад і знову перемішуючи всі папірці. Отримані в такий спосіб цифри запишемо у вигляді таблиці.

Така таблиця називається *таблицею випадкових чисел*, хоча вірніше було б назвати її *таблицею випадкових цифр*. І в ході розрахунку, коли знадобиться значення випадкової величини з розподілом

$$\begin{pmatrix} 0 & 1 & 2 & \dots & 9 \\ 0,1 & 0,1 & 0,1 & \dots & 0,1 \end{pmatrix}, \quad (1.47)$$

будемо брати чергову цифру з цієї таблиці.

Найбільша з опублікованих таблиць випадкових чисел містить 1000000 цифр. Звичайно, вона складалася за допомогою більш сучасної техніки, чим шапка: була сконструйована спеціальна рулетка з використанням електроніки. Найпростіша схема такої рулетки приведена на рис.1.7 (обертювий диск різко зупиняється і вибирається та цифра, на яку вказує нерухома стрілка).

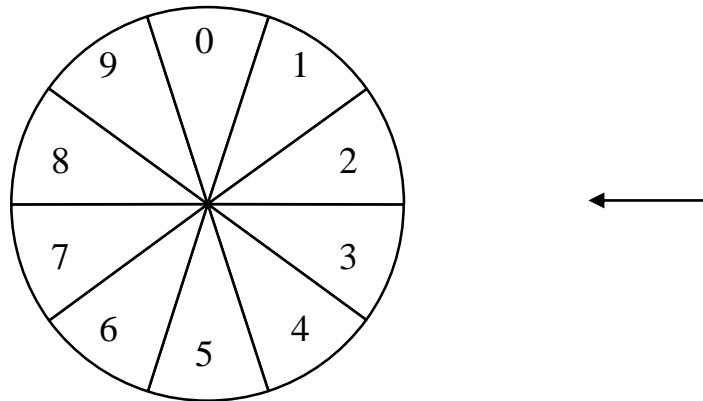


Рисунок 1.7 – Схема рулетки

Необхідно відмітити, що скласти таблицю випадкових чисел не так просто, як може здатися. Будь-який реальний фізичний прилад виробляє випадкові величини з розподілом, що трохи відрізняється від ідеального (1.47). До того ж у ході експерименту можливі помилки. Тому складені таблиці ретельно перевіряються за допомогою спеціальних статистичних тестів: чи не суперечать ті чи інші властивості групи чисел гіпотезі про те, що ці числа - значення випадкової величини (1.47).

Наведемо один з найпростіших тестів. Розглянемо таблицю, що містить  $N$  цифр. Нехай число нулів у цій таблиці  $\nu_0$ , число одиниць  $\nu_1$ , число двійок  $\nu_2$  і т.д. Обчислимо суму

$$\sum_{i=0}^9 (\nu_i - 0,1N)^2 . \quad (1.48)$$

Теорія ймовірностей дозволяє передбачити, у яких межах може лежати ця сума: вона не повинна бути занадто великою (тому що математичне сподівання кожного з  $v_i$  дорівнює  $0,1N$ ), але не повинна бути і занадто малою (тому що це означало б «занадто закономірний» розподіл значень).

Таблиці випадкових чисел (додаток А) використовуються тільки при розрахунках за методом Монте-Карло вручну.

2 *Генератори випадкових чисел.* Здавалося б, що згадувану рулетку можна приєднати до обчислювальної машини й у міру потреби виробляти випадкові числа. Однак будь-який механічний прилад буде занадто повільним у порівнянні з ЕОМ. Тому як генератори випадкових величин найчастіше використовують шуми в електронних лампах: якщо за деякий фіксований проміжок часу  $\Delta t$  рівень шуму перевищив заданий поріг парне число раз, то записується нуль, а якщо непарне число раз, то записується одиниця.

На перший погляд це дуже зручний спосіб. Нехай  $m$  таких генераторів працюють паралельно, працюють увесь час і засилають випадкові нулі й одиниці в усі двоїчні розряди спеціального осередку. Кожен такт – одне з  $m$ -розрядних чисел. У будь-який момент рахунку можна звернутися до цього осередку й узяти відтіля значення випадкової величини  $\gamma$ , рівномірно розподіленої в інтервалі  $(0, 1)$ . Звичайно, значення наближене, записане у формі  $m$ -розрядного двоїчного дроби  $0,\alpha_{(1)}\alpha_{(2)}\dots\alpha_{(m)}$ , де кожна з величин  $\alpha_{(t)}$  імітує випадкову величину з розподілом

$$\begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}.$$

Однак і цей метод не вільний від недоліків. По-перше, важко перевірити «якість» вироблюваних чисел. Перевірки приходиться робити періодично, тому що через які-небудь несправності може виникнути так званий «дрейф розподілу» (тобто нулі й одиниці в якому-небудь з розрядів стануть з'являтися неоднаково часто). По-друге, звичайно всі розрахунки на ЕОМ проводять двічі, щоб виключити можливість випадкового збою. Але відтворити ті ж випадкові числа неможливо, якщо їх по ходу рахунку не запам'ятовувати. А якщо їх запам'ятовувати, то ми знову приходимо до випадку таблиць.

Датчики такого типу, безсумнівно, виявляться корисними тоді, коли будуть будуватися спеціалізовані ЕОМ для вирішення задач за методом Монте-Карло. А для універсальних ЕОМ, на яких розрахунки за допомогою випадкових чисел проводяться лише зрідка, містити й

експлуатувати спеціальний пристрій просто не економічно. Краще використовувати так звані псевдовипадкові числа.

3 *Псевдовипадкові числа.* Оскільки «якість» використовуваних випадкових чисел перевіряється за допомогою спеціальних тестів, можна не цікавитися тим, як ці числа отримані: аби вони задовольняли прийнятій системі тестів. Можна навіть спробувати обчислювати їх за заданою формулою.

Числа, які розраховуються за допомогою якої-небудь формули, і значення випадкової величини, що імітують  $\gamma$ , називаються *псевдовипадковими* числами. Під словом «імітуючі» мається на увазі, що ці числа задовольняють низці тестів так, ніби вони були значеннями цієї випадкової величини.

Перший алгоритм для одержання псевдовипадкових чисел був запропонований Дж. Нейманом. Він називається *методом середини квадратів*. Пояснимо його на прикладі.

Нехай задане 4-значне число  $\gamma_0 = 0,9876$ . Зведемо його в квадрат. Одержимо 8-значне число  $\gamma_0^2 = 0,97535376$ . Виберемо чотири середні цифри цього числа і покладемо  $\gamma_1 = 0,5353$ .

Потім зведемо  $\gamma_1$  в квадрат ( $\gamma_1^2 = 0,28654609$ ) і знову витягнемо чотири середні цифри. Одержимо  $\gamma_2 = 0,6546$ .

Далі,  $\gamma_2^2 = 0,42850116$ ,  $\gamma_3 = 0,8501$ ;  $\gamma_3^2 = 0,72267001$ ,  $\gamma_4 = 0,2670$ ;  
 $\gamma_4^2 = 0,07128900$ ,  $\gamma_5 = 0,1289$  і т.д.

Перевага методу псевдовипадкових чисел досить очевидна.

По-перше, на одержання кожного числа затрачається усього декілька простих операцій, так що швидкість генерування випадкових чисел має той же порядок, що й швидкість роботи ЕОМ.

По-друге, програма займає усього декілька блоків пам'яті.

По-третє, кожне з чисел  $\gamma_k$  може бути легко відтворено.

По-четверте, потрібно лише один раз перевірити «якість» такої послідовності, потім її можна багато разів використовувати при розрахунку подібних задач.

Єдиний недолік методу – обмеженість «запасу» псевдовипадкових чисел. Однак існують способи, які дозволяють одержувати набагато більше чисел. Зокрема, можна змінювати початкові числа  $\gamma_0$ .

Переважає більшість розрахунків за методом Монте-Карло в цей час здійснюється з використанням псевдовипадкових чисел.

### 1.3.2.4 Метод статистичних випробувань

**Мірою точності обчислення статистичного параметра є середнє квадратичне відхилення або випадкова похибка оцінки цього параметра.**

Формули для обчислення випадкових похибок визначення статистичних параметрів отримуються на основі методу статистичних випробувань або методу Монте – Карло. Цей метод дозволяє моделювати процеси, на розвиток яких суттєво впливає випадковий фактор.

Моделювання виконується наступним чином.

1. Задається ряд випадкових чисел  $\delta_i$  (наприклад, за спеціально розробленою таблицею випадкових чисел, де числа змінюються в інтервалі від 0 до 1, тобто  $0 < \delta_i < 1$ ).
2. Отримані випадкові числа приймаються за забезпеченості  $\delta_i = P_i$ .
3. Величина стоку заданої забезпеченості визначається за вибраним законом розподілу, тобто

$$P(x) = \int_x^{\infty} f(x) dx. \quad (1.49)$$

З (1.49) або за таблицею вибраного розподілу встановлюється шукане  $x$ . При цьому статистичні параметри розподілу задаються рівними відповідним параметрам генеральної сукупності.

Розглянемо випадок, коли як теоретичний закон розподілу вибраний закон розподілу Пірсона III. Це трипараметричний закон розподілу випадкової величини, для практичного застосування якого необхідно задати три параметри – математичне сподівання  $m_x$ , коефіцієнт варіації  $C_v$ , коефіцієнт асиметрії  $C_s$ . Припустимо, що ці параметри задані у вигляді конкретних чисел.

Як відомо, закон Пірсона III представляється у вигляді чисел Форстера  $\Phi_p$

$$\Phi_p = \frac{x_p - m_x}{\sigma_x}, \quad (1.50)$$

звідки

$$x_p = m_x + \Phi_p \sigma_x, \quad (1.51)$$

де  $P$  - забезпеченість.

За виразом (1.51) можна генерувати ряди довжиною в декілька тисяч членів.

4. Після генерації довготривалого ряду величин стоку різної забезпеченості, який розглядається як генеральна сукупність, відбувається його розбиття на вибірки меншої довжини. Наприклад, якщо ряд довжиною 5000 членів розбити на вибірки довжиною в 100 елементів, то буде отримано 50 вибірок, а отже і 50 значень оцінок статистичних параметрів.

4. Середня квадратична похибка визначення середнього арифметичного значення при довжині виборки  $n = 100$  буде визначатися за формулою

$$\sigma_{\bar{x}, n=100} = \sqrt{\frac{\sum_{i=1}^m (\bar{x}_j - m_x)^2}{m-1}}, \quad (1.52)$$

$\sigma_{\bar{x}, n=100}$  - середня квадратична похибка визначення середнього арифметичного значення при довжині вибірки  $n = 100$ ;

$\bar{x}_j$  - середнє арифметичне, розраховане за  $j$ -тою вибіркою;

$m$  - число вибірок.

Середнє квадратичне відхилення коефіцієнтів варіації, визначене по 200 вибірках довжиною в 20 елементів, розраховується таким чином

$$\sigma_{C_v, n=20} = \sqrt{\frac{\sum_{i=1}^m (C_{v_j} - C_v)^2}{m-1}}, \quad (1.53)$$

де -  $C_{v, n=20}$  - середня квадратична похибка визначення коефіцієнта варіації при довжині вибірки  $n = 20$ ;

$C_{v_j}$  - значення коефіцієнта варіації, розраховане за  $j$ -тою вибіркою;

$m$  - число вибірок.

Середньоквадратичні відхилення статистичних параметрів при їхньому визначенні за вибірковими даними або так звані випадкові похибки можна розраховувати для вибірок різної довжини і у подальшому установити математичне співвідношення між похибками та довжиною вибірок  $\sigma_{C_v} = f(n)$ , або між похибками та середньоквадратичним відхиленням вихідних вибірок  $\sigma_{C_v} = f(C_v)$ . Саме методом статистичних випробувань і були отримані формули для розрахунків випадкових похибок визначення статистичних параметрів  $C_S$  та  $C_V$  за обмеженими у часі даними.

Що стосується оцінки математичного сподівання, то її середньоквадратичне відхилення розраховується за формулою, розробленою для величин, які підлягають нормальному закону розподілу, тобто формула може бути отримана аналітичним шляхом, дотримуючись при цьому припущення, що нормальний закон розподілу вибірових середніх зберігається і для вибірок, які відхиляються від нормального розподілу

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}. \quad (1.54)$$

Похибки  $\sigma_{C_v}$  коефіцієнтів варіації  $C_v$ , розрахованих за методом моментів, обчислюються за формулою СН – 435-72

$$\sigma_{C_v} = C_v \sqrt{(1 + C_v^2) / 2n} \quad (1.55)$$

або за формулою Є. Г. Блохінова

$$\sigma_{C_v} = \frac{C_v}{n + 4C_v^2} \sqrt{\frac{n}{2}(1 + C_v^2)}. \quad (1.56)$$

Середня квадратична похибка коефіцієнта асиметрії  $\sigma_{C_s}$  визначається за теоретичною формулою С.М.Крицького та М.Ф.Менкеля

$$\sigma_{C_s} = \sqrt{\frac{6}{n}(1 + 6C_v^2 + 5C_v^4)} \quad (1.57)$$

або формулою А.Ш. Резніковського

$$\sigma_{C_s} = \frac{1}{C_v} \sqrt{\frac{6}{n}}. \quad (1.58)$$

Для визначення якості розрахунків статистичних параметрів за даними спостережень необхідно ввести критерій якості, а саме – допустиму похибку розрахунків. Допустиму похибку краще представляти у частках або відсотках від значення параметра, точність визначення якого

оцінюється. Отже, і самі випадкові похибки зручніше представляти у відсотках, тобто

$$\varepsilon_{\bar{x}} = \frac{\sigma_{\bar{x}}}{\bar{x}} = \frac{\sigma_x}{\bar{x}\sqrt{n}} = \frac{C_v}{\sqrt{n}} \cdot 100\%. \quad (1.59)$$

Із виразу (1.59) видно, що відносна випадкова похибка визначення середнього арифметичного значення пропорційна коефіцієнту  $C_v$  та зворотна числу членів вибірки  $n$ . Із збільшенням числа спостережень відносна похибка  $\varepsilon_{\bar{x}}$  зменшується.

Для середніх арифметичних значень допустимою похибкою є 10%. Якщо фактична похибка розрахунку, визначена за формулою (1.59), перевищує 10%, то вихідний ряд вважається коротким, а визначений статистичний параметр (у даному випадку – середнє арифметичне) потребує уточнення. Уточнення може бути виконано із залученням даних по річці-аналогу більшої довжини. Методи уточнення параметрів розрахункового ряду з використанням даних по річці – аналогу називаються методами приведення статистичних параметрів коротких рядів до довгого періоду. Такі методи будуть розглянуті у наступних розділах.

Відносна випадкова похибка коефіцієнта варіації визначається таким чином

$$\varepsilon_{C_v} = \frac{\sigma_{C_v}}{C_v} \cdot 100\%. \quad (1.60)$$

Допустимою похибкою визначення коефіцієнта варіації є величина, що дорівнює 15%.

Відносна похибка визначення коефіцієнта асиметрії розраховується за виразом

$$\varepsilon_{C_s} = \frac{\sigma_{C_s}}{C_s} \cdot 100\% . \quad (1.61)$$

При практичному застосуванні формул (1.55) – (1.61) використовуються значення коефіцієнтів варіації, розрахованих за вибіркою.

З формул (1.54) – (1.59) можна знайти довжину рядів, при яких випадкова похибка дорівнює заданій величині. Так, при  $C_v = 0.2-1.0$  та заданій похибці  $\varepsilon_{C_v} = 15\%$  необхідні ряди довжиною 300- 350 років. Зазначимо, що гідрологічні ряди найчастіше мають довжину 40-50 років і лише окремі



з них досягають довжини в 100 років і більше. У зв'язку з великою похибкою визначення за даними спостережень коефіцієнт  $C_S$  нормують за співвідношенням  $C_S/C_V$  у межах окремих районів.

### 1.3.3 Метод найбільшої правдоподібності

Походження назви цього методу пов'язане із застосуванням функції правдоподібності до визначення статистичних параметрів трипараметричного гама-розподілу, запропонованого С.М.Крицьким та М.Ф.Менкелем. Сутність методу найбільшої правдоподібності полягає в тому, щоб застосувати функцію правдоподібності до вибраного теоретичного закону статистичного розподілу випадкової величини з метою оптимізації параметрів розподілу.

Функція правдоподібності - це ймовірність сумісної появи подій. У свою чергу ймовірність сумісної появи подій може розглядатися як добуток ймовірностей кожної з них.

Нехай задана вибірка спостережених значень випадкових величин  $x_1, x_2, \dots, x_n$ . Знайти функцію правдоподібності цієї вибірки - це означає обчислити ймовірність сумісної появи значень  $x_1, x_2, \dots, x_n$  в результаті проведення  $n$  експериментів.

Метод найбільшої правдоподібності відноситься до методів математичної статистики, у яких за оцінку невідомого значення параметру щільності ймовірності береться те його значення, при якому функція правдоподібності досягає свого максимуму для даної вибірки випадкових величин. Математичний вираз для функції правдоподібності з невідомим параметром  $a$  має такий вигляд:

$$L(x_1, x_2, x_3, \dots, x_n, a) = f(x_1, a) f(x_2, a) \dots f(x_n, a) \quad (1.62)$$

Функція правдоподібності задає ймовірність появи вибірки  $x_1, x_2, \dots, x_n$ . Чим більше значення функції  $L$ , тим більша ймовірність появи вибірки. Звідси і походить назва методу найбільшої правдоподібності.

Якщо в залежності від параметра  $a$  величина  $L$  значно змінюється, то можна з імовірністю, близькою до 1, отримати значення  $a$ . І, навпаки, якщо  $L$  не змінюється в залежності від  $a$ , то немає підстав робити висновки про  $a$  взагалі.

Сам параметр  $a$  знаходять з умови, що похідна від функції правдоподібності буде дорівнювати нулю. Відповідно до правил диференціального обчислення для знаходження оцінки найбільшої правдоподібності необхідно розв'язати рівняння

$$\frac{\partial L}{\partial a} = 0. \quad (1.63)$$

З метою спрощення визначення функції правдоподібності до рівняння (1.63) вводять логарифм і розглядають рівняння

$$\frac{1}{L} \frac{\partial L}{\partial a} = \frac{\partial(\ln a)}{\partial a} = 0. \quad (1.64)$$

### **1.3.3.1 Приклад застосування методу найбільшої правдоподібності до визначення статистичних параметрів нормального закону розподілу**

Розглянемо нормальний закон розподілу, який можна представити у вигляді

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}, \quad (1.65)$$

де  $f(x)$  - щільність імовірності;

$m_x$  - математичне сподівання;

$\sigma_x$  - середнє квадратичне відхилення.

У такому випадку функція правдоподібності має вигляд

$$L = f(x_1) f(x_2) f(x_3) \dots f(x_n). \quad (1.66)$$

Задача полягає у тому, щоб знайти статистичні параметри  $\hat{\sigma}_x^2, \bar{x}$ , використовуючи функцію найбільшої правдоподібності. Представимо ймовірність кожного члена вибірки, виражену через функцію  $f(x)$

$$f(x_1) = \frac{1}{\sqrt{2\pi\hat{\sigma}_x^2}} e^{-\frac{(x_1-\bar{x})^2}{2\hat{\sigma}_x^2}}; \quad (1.67)$$

$$f(x_2) = \frac{1}{\sqrt{2\pi\hat{\sigma}_x^2}} e^{-\frac{(x_2 - \bar{x})^2}{2\hat{\sigma}_x^2}} ; \quad (1.68)$$

$$f(x_n) = \frac{1}{\sqrt{2\pi\hat{\sigma}_x^2}} e^{-\frac{(x_n - \bar{x})^2}{2\hat{\sigma}_x^2}} . \quad (1.69)$$

У такому випадку функція правдоподібності набуде вигляду

$$L = f(x_1)f(x_2)f(x_3)\dots f(x_n) = \frac{1}{(2\pi\hat{\sigma}_x^2)^n} e^{-\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{2\hat{\sigma}_x^2}} \quad (1.70)$$

Представимо функцію у логарифмічному вигляді

$$\ln L = \ln 1 - \frac{n}{2} \ln 2\pi - \frac{n}{2} \ln(\hat{\sigma}_x^2) - \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{2(\hat{\sigma}_x^2)} \quad (1.71)$$

Для пошуку оптимального визначення статистичного параметра  $\bar{x}$  обчислюється похідна

$$\frac{\partial(\ln \bar{x})}{\partial \bar{x}} = 0. \quad (1.72)$$

Звідки отримуємо:

$$\frac{\partial \ln L}{\partial \bar{x}} = \frac{2}{2(\hat{\sigma}_x^2)} \cdot \sum_{i=1}^n (x_i - \bar{x}) = 0; \quad (1.73)$$

$$\sum_{i=1}^n (x_i - \bar{x}) = 0. \quad (1.74)$$

$$\sum_{i=1}^n x_i - \bar{x} \cdot n = 0. \quad (1.75)$$

Отже,

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}. \quad (1.76)$$

Аналогічно відбувається пошук оптимального виразу для статистичного параметра  $\hat{\sigma}_x^2$

$$\frac{\partial \ln L}{\partial (\hat{\sigma}_x^2)} = -\frac{n}{2} \frac{1}{(\hat{\sigma}_x^2)} - \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{2} \cdot (-1) \frac{1}{(\hat{\sigma}_x^2)^2} = 0; \quad (1.77)$$

$$\frac{n}{\hat{\sigma}_x^2} - \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\hat{\sigma}_x^4} = 0; \quad (1.78)$$

$$\frac{1}{\hat{\sigma}_x^2} \left( n - \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\hat{\sigma}_x^2} \right) = 0; \quad (1.79)$$

$$\hat{\sigma}_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}. \quad (1.80)$$

Таким чином, у результаті застосування функції правдоподібності до нормального закону розподілу були отримані математичні вирази для визначення статистичних параметрів  $\hat{\sigma}_x^2, \bar{x}$ , які збігаються з відповідними виразами методу моментів.

Метод найбільшої правдоподібності приводить до обґрунтованих оцінок з незначним зміщенням. Але вигляд розрахункових формул статистичних параметрів залежить від обраного закону розподілу випадкової величини. Є.Г.Блохінов застосував метод найбільшої правдоподібності до трипараметричного гама-розподілу С.М.Крицького та

М.Ф.Менкеля. Строге вирішення приводить до складних трансцендентних рівнянь. У зв'язку з цим був запропонований спрощений спосіб оцінки параметрів. У результаті отримані такі статистики

$$\hat{\lambda}_1 = \frac{\sum_{i=1}^n x_i}{n}; \quad (1.81)$$

$$\hat{\lambda}_2 = \frac{\sum_{i=1}^n \lg \frac{x_i}{\bar{x}}}{n}; \quad (1.82)$$

$$\hat{\lambda}_3 = \frac{\sum_{i=1}^n \frac{x_i}{\bar{x}} \lg \frac{x_i}{\bar{x}}}{n}. \quad (1.83)$$

Перша із статистик  $\hat{\lambda}_1$  дорівнює середньоарифметичному значенню випадкової величини  $X$ .

Дві другі ( $\lambda_2, \lambda_3$ ) функціонально зв'язані з коефіцієнтом варіації  $C_V$  та співвідношенням  $C_S/C_V$ .

Для переходу від  $\lambda_2$  та  $\lambda_3$  до  $C_V$  та  $C_S/C_V$  побудовані спеціальні номограми.

Деяка зміщеність параметрів  $\lambda_2$  та  $\lambda_3$  може бути усунена за рахунок поправкового множника  $\frac{n}{n-1}$ . Тоді

$$\hat{\lambda}_2 = \frac{\sum_{i=1}^n \lg k_i}{n-1}; \quad (1.84)$$

$$\hat{\lambda}_3 = \frac{\sum_{i=1}^n k_i \lg k_i}{n-1}, \quad (1.85)$$

де  $k_i$  - модульний коефіцієнт.

**Метод найбільшої правдоподібності рекомендується застосовувати для визначення статистичних параметрів при  $C_V > 0.5$ .**

При  $C_V \leq 0.5$  методи найбільшої правдоподібності та моментів практично дають однакові результати.

### 1.3.3.2 Точність визначення статистичних параметрів, обчислених за методом найбільшої правдоподібності

Середня квадратична похибка обчислення  $n$ -річних середніх арифметичних значень стокових рядів  $\sigma_{\bar{x}}$  може бути обчислена за тією ж формулою, що й в методі моментів

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}. \quad (1.86)$$

Середня квадратична похибка розрахунку коефіцієнта варіації  $C_V$ , отримана на основі статистичних випробувань, для умови  $C_S/C_V = 2$  становить:

$$\sigma_{C_V} = \frac{C_V}{\sqrt{2n}} \sqrt{\frac{3}{3 + C_V^2}}. \quad (1.87)$$

Якщо  $C_S / C_V \neq 2.0$ , стандарт вибіркової оцінки  $C_V$ , визначеної методом найбільшої правдоподібності, корегується за допомогою поправкового коефіцієнта

$$K_{C_V} = \frac{\varphi_{C_V} (\text{при } C_V = mC_V)}{\varphi_{C_V} (\text{при } C_S = 2C_V)}, \quad (1.88)$$

який знімається з графіка залежності  $K_{C_V}$  для вибірових оцінок  $C_V$  та співвідношення  $C_S/C_V$  (рис.1.8).

Середнє квадратичне відхилення для  $C_S/C_V$ , визначеного за методом найбільшої правдоподібності, установлюється за зв'язком  $\sigma_{C_S/C_V}$  і  $C_V$  (рис.1.9).

Обчислення відносних середньоквадратичних відхилень статистичних параметрів відбувається за формулами (1.59 – 1.61).

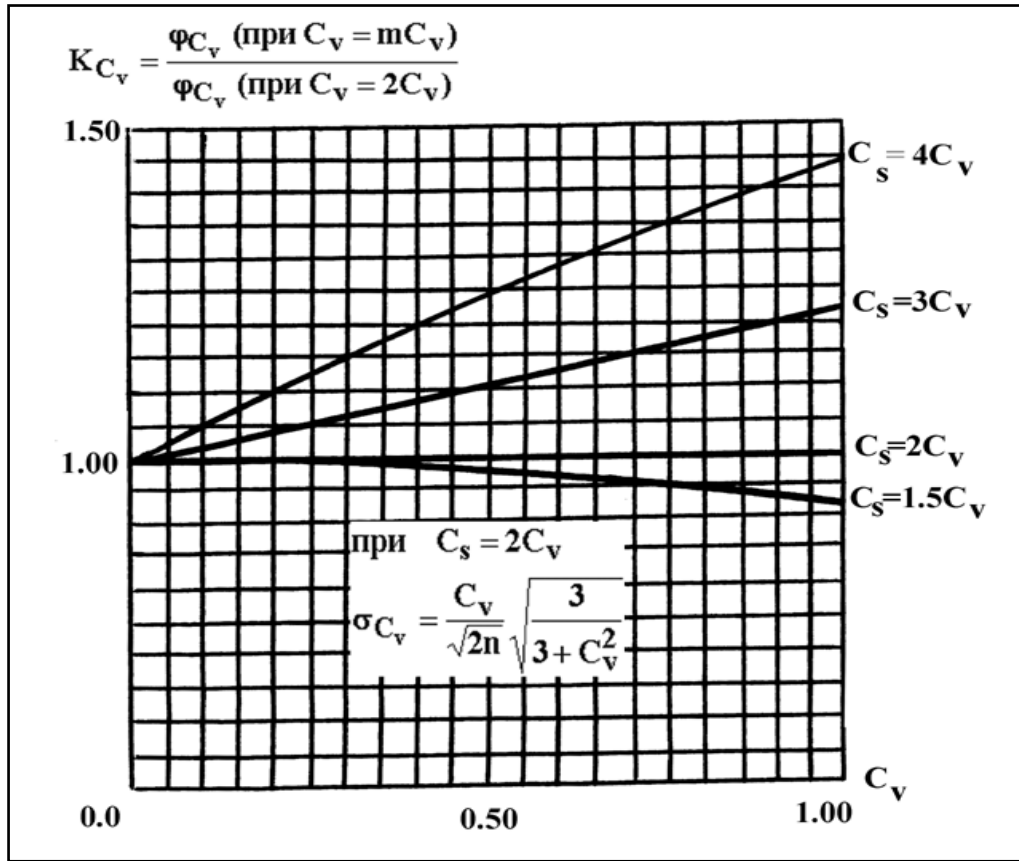


Рисунок 1.8 – Графік для визначення коефіцієнта  $K_{C_v}$

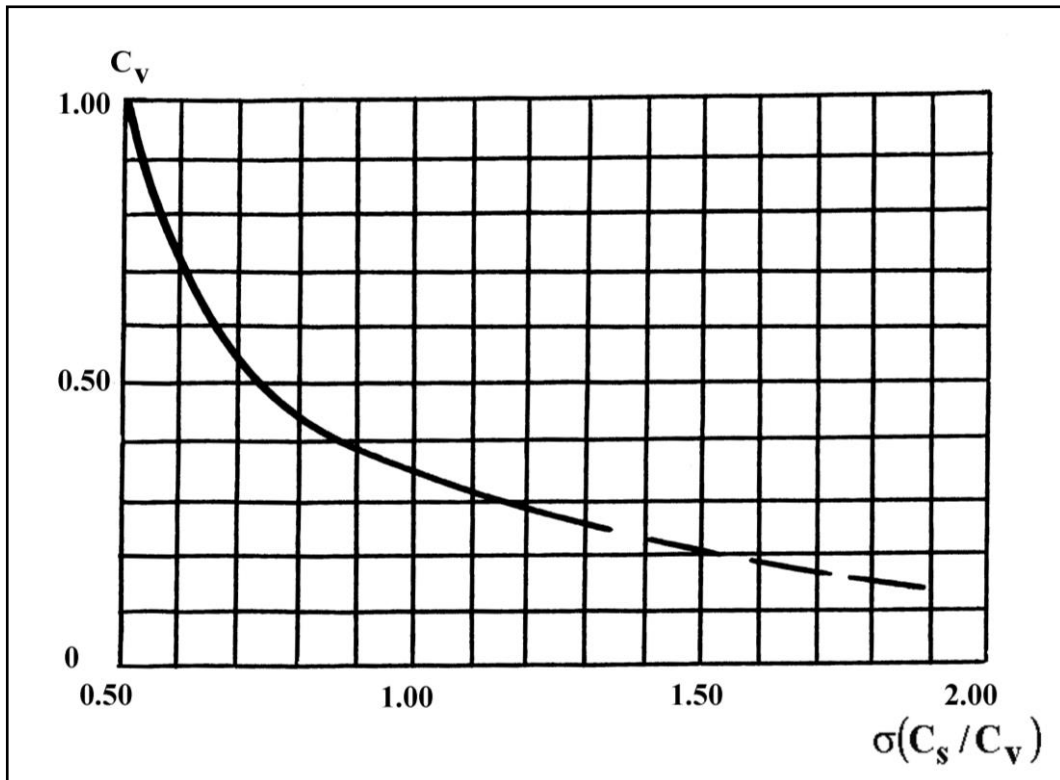


Рисунок 1.9 - Графік для визначення  $\sigma_{C_s/C_v}$

### 1.3.4 Графо-аналітичний метод Г.А.Алексєєва

Графо-аналітичний метод являє собою спрощений спосіб розрахунків статистичних параметрів. В ньому використовується емпірична крива забезпеченостей і аналітичний (теоретичний) закон розподілу. При його застосуванні приймається умова збіжності теоретичної кривої розподілу з емпіричною хоча б у трьох точках. Цей метод дозволяє оцінити статистичні параметри по рядах стоку безпосередньо для того теоретичного розподілу ймовірностей, який в більшій мірі відповідає емпіричному. Розглянемо випадок, коли за теоретичний розподіл стокової величини прийнятий закон розподілу Пірсона III.

У гідрологічних розрахунках закон розподілу випадкової величини задається у вигляді функції забезпеченості. Забезпеченість значення  $x$  випадкової величини  $X$  є ймовірність перевищення  $x$ , тобто

$$P(x) = p(X > x). \quad (1.89)$$

Оцінкою забезпеченості значення  $x_i$  з вибірки випадкових величин довжиною  $n$  є відносна частота події  $X \geq x_i$ . Для обчислення емпіричної забезпеченості  $\hat{p}$  використовується формула

$$\hat{p}(x_i) = \frac{m \cdot 100\%}{n + 1}, \quad (1.90)$$

де  $m$  - порядковий номер елемента  $x_i$  у ранжованому статистичному ряді (мається на увазі статистичний ряд, в якому всі члени розміщені в порядку зменшення). Фактично  $m$  являє собою абсолютну частоту події  $X \leq x_i$ .

З графіка емпіричної кривої забезпеченості, побудованої на відповідній клітчатці ймовірностей, знімаються величини стоку в характерних точках з забезпеченістю 5, 50 та 95 відсотків (%). Виходячи з припущення, що ці точки емпіричної кривої забезпеченості співпадають з теоретичною, звернемося до закону розподілу Пірсона III. Цей теоретичний закон розподілу випадкової величини надається в СНІП 2.01.14-83 таблицею нормованих відхилень  $\Phi_p(C_s)$ , які залежать від забезпеченості  $P$  і коефіцієнта асиметрії  $C_s$ :

$$\Phi_p(C_s) = \frac{x_p - \bar{x}}{\sigma_x}. \quad (1.91)$$



Для трьох характерних точок за формулою (1.91) визначаються випадкові величини  $x_p$  :

$$x_5 = \bar{x} + \hat{\sigma}_x \Phi_5 ; \quad (1.92)$$

$$x_{50} = \bar{x} + \hat{\sigma}_x \Phi_{50} ; \quad (1.93)$$

$$x_{95} = \bar{x} + \hat{\sigma}_x \Phi_{95} \quad (1.94)$$

з трьома невизначеними параметрами:  $\bar{x}$ ,  $\hat{\sigma}_x$  та  $\hat{C}_s$ . Параметр  $\hat{C}_s$  входить у рівняння (1.92 -1.94) в силу того, що  $\Phi_p$  є функцією  $C_s$ .

Для визначення коефіцієнта асиметрії використовується коефіцієнт скісності  $S$ , який функціонально пов'язаний з  $C_s$ .

Коефіцієнт скісності розраховується за такою формулою

$$\hat{S} = \frac{x_5 + x_{95} - 2x_{50}}{x_5 - x_{95}}. \quad (1.95)$$

За таблицею 1.1 відповідно  $\hat{S}$  встановлюється коефіцієнт  $\hat{C}_s$  та нормовані ординати  $\Phi_5$ ,  $\Phi_{50}$ ,  $\Phi_{95}$ .

Для отримання математичного виразу, який буде визначати середньоквадратичне відхилення, віднімаємо з лівої та правої частин рівняння (1.92) відповідні частини рівняння (1.94):

$$x_5 - x_{95} = \hat{\sigma}_x (\Phi_5 - \Phi_{95}), \quad (1.96)$$

звідки

$$\hat{\sigma}_x = \frac{x_5 - x_{95}}{\Phi_5 - \Phi_{95}}. \quad (1.97)$$

Середнє арифметичне значення знаходять з рівняння (1.93)

$$\bar{x} = x_{50} - \hat{\sigma}_x \Phi_{50}. \quad (1.98)$$

Коефіцієнт варіації розраховується як

$$\widehat{C}_v = \frac{\widehat{\sigma}_x}{\bar{x}}. \quad (1.99)$$

Якщо коефіцієнт скісності  $S$  від'ємний, то це свідчить про від'ємну асиметрію ( $C_S < 0$ ) розподілу.

У таких випадках наведені в таблиці ординат кривої забезпеченості Пірсона III величини беруться з протилежним знаком:

$$\frac{x_p - \bar{x}}{\sigma_x} = -\Phi_p, \quad (1.100)$$

для значень забезпеченості  $P^* = 100 - P$  і при додатному значенні коефіцієнта асиметрії  $C_S^* = -C_S$ .

Тоді

$$S^* = -S = \frac{2x_{50} - x_5 - x_{95}}{x_5 - x_{95}}; \quad (1.101)$$

$$\widehat{\sigma}_x = \frac{x_5 - x_{95}}{\Phi_5 - \Phi_{95}}; \quad (1.102)$$

$$\bar{x} = x_{50} - \widehat{\sigma}_x \Phi_{50}. \quad (1.103)$$

Зрозуміло, що хоч у графо-аналітичному методі розрахунків статистичних параметрів і використовується теоретичний закон розподілу, отримані статистичні характеристики є все ж таки тільки оцінками статистичних параметрів генеральної сукупності, бо вони спираються на емпіричну криву забезпеченості, побудовану за даними вибірок з генеральних сукупностей.

Випадкові похибки обчислення статистичних параметрів визначаються за тими ж формулами, що і в методі моментів.

Таблиця 1.1 – Значення коефіцієнта асиметрії  $C_s$  та скісності  $S$  для біноміальної кривої розподілу

$C_s$	$\frac{x_p - \bar{x}}{\sigma_x} = \frac{k_p - 1}{C_v} = \Phi(P, C_s)$						$\Phi_5 - \Phi_{95}$	$S$
	$\Phi_1$	$\Phi_2$	$\Phi_5$	$\Phi_{10}$	$\Phi_{50}$	$\Phi_{95}$		
0.0	2.33	2.02	1.64	1.28	0.00	-1.64	3.28	0.00
0.1	2.40	2.11	1.67	1.29	-0.02	-1.61	3.28	0.03
0.2	2.47	2.16	1.70	1.30	-0.03	-1.58	3.28	0.06
0.3	2.54	2.21	1.72	1.31	-0.05	-1.52	3.27	0.09
0.4	2.61	2.26	1.75	1.32	-0.07	-1.52	3.27	0.11
0.5	2.68	2.31	1.77	1.32	-0.08	-1.49	3.26	0.16
0.6	2.75	2.35	1.80	1.33	-0.10	-1.45	3.25	0.17
0.7	2.82	2.40	1.82	1.33	-0.12	-1.42	3.24	0.20
0.8	2.89	2.45	1.84	1.34	-0.13	-1.38	3.22	0.22
0.9	2.96	2.50	1.86	1.34	-0.15	-1.35	3.21	0.25
1.0	3.02	2.54	1.88	1.34	-0.16	-1.32	3.20	0.28
1.1	3.09	2.58	1.89	1.34	-0.18	-1.28	3.17	0.31
1.2	3.15	2.62	1.92	1.34	-0.19	-1.24	3.16	0.34
1.3	3.21	2.57	1.94	1.34	-0.21	-1.20	3.14	0.37
1.4	3.27	2.71	1.95	1.34	-0.22	-1.17	3.12	0.39
1.5	3.33	2.74	1.96	1.33	-0.24	-1.13	3.09	0.42
1.6	3.39	2.78	1.97	1.33	-0.25	-1.10	3.07	0.45
1.7	3.44	2.82	1.98	1.32	-0.27	-1.06	3.04	0.49
1.8	3.50	2.85	1.99	1.32	-0.28	-1.02	3.01	0.51
1.9	3.55	2.88	2.00	1.31	-0.29	-0.98	2.98	0.54
2.0	3.60	2.91	2.00	1.30	-0.31	-0.95	2.95	0.57
2.1	3.65	2.94	2.01	1.29	-0.32	-0.91	2.92	0.59
2.2	3.68	2.95	2.02	1.27	-0.33	-0.88	2.90	0.63
2.3	3.73	2.98	2.01	1.26	-0.34	-0.85	2.86	0.64
2.4	3.78	3.02	2.00	1.25	-0.35	-0.82	2.82	0.68
2.5	3.82	3.05	2.00	1.23	-0.36	-0.79	2.79	0.69
2.6	3.85	3.05	2.00	1.21	-0.37	-0.76	2.76	0.72
2.7	3.92	3.10	2.00	1.19	-0.38	-0.74	2.74	0.74
2.8	3.96	3.12	2.00	1.18	-0.39	-0.71	2.71	0.76
2.9	4.01	3.12	1.99	1.15	-0.39	-0.69	2.68	0.78
3.0	4.05	3.14	1.97	1.13	-0.40	-0.66	2.63	0.80
3.1	4.09	3.14	1.97	1.11	-0.40	-0.64	2.62	0.81
3.2	4.11	3.14	1.96	1.09	-0.41	-0.62	2.59	0.83

Продовження таблиці 1.1

$C_s$	$\frac{x_p - \bar{x}}{\sigma_x} = \frac{k_p - 1}{C_v} = \Phi(P, C_s)$						$\Phi_5 - \Phi_{95}$	$S$
	$\Phi_1$	$\Phi_2$	$\Phi_5$	$\Phi_{10}$	$\Phi_{50}$	$\Phi_{95}$		
3.3	4.15	3.14	1.95	1.08	-0.41	-0.60	2.56	0.85
3.4	4.18	3.15	1.94	1.08	-0.41	-0.59	2.53	0.86
3.5	4.21	3.16	1.93	1.04	-0.41	-0.57	2.50	0.87
3.6	4.24	3.17	1.93	1.03	-0.42	-0.56	2.48	0.89
3.7	4.26	3.18	1.91	1.01	-0.42	-0.54	2.45	0.90
3.8	4.29	3.18	1.90	1.00	-0.42	-0.53	2.43	0.91
3.9	4.32	3.20	1.90	0.98	-0.41	-0.51	2.41	0.92
4.0	4.34	3.20	1.90	0.96	-0.41	-0.50	2.40	0.92
4.1	4.36	3.22	1.89	0.95	-0.41	-0.49	2.38	0.93
4.2	4.39	3.21	1.88	0.93	-0.41	-0.48	2.36	0.94
4.6	4.46	3.27	1.84	0.87	-0.40	-0.44	2.28	0.97
4.7	4.49	3.28	1.83	0.85	-0.40	-0.43	2.26	0.97
4.8	4.50	3.29	1.81	0.82	-0.39	-0.42	2.23	0.98
4.9	4.51	3.30	1.80	0.80	-0.39	-0.41	2.21	0.98
5.0	4.54	3.32	1.78	0.78	-0.38	-0.40	2.18	0.98
5.1	4.57	3.32	1.76	0.76	-0.38	-0.39	2.15	0.98
5.2	4.59	3.33	1.74	0.73	-0.37	-0.38	2.15	0.98

### 1.3.5 Теоретичні криві розподілу, їх особливості та межі застосування

Теоретичні закони розподілу ймовірностей базуються або на визначених теоретичних схемах, або є узагальненням емпіричних розподілів.

Вимоги до теоретичних кривих розподілу стокових величин:

1. У рівнянні кривої повинно бути якнайменше параметрів, які чисельно визначаються за вибірковими даними.
2. Через те, що значення стоку завжди позитивні, крива розподілу не повинна знаходитися в області від'ємних значень.
3. Верхня межа кривої розподілу необмежена.
4. Теоретичні криві розподілу повинні бути одномодальними, що витікає з умови однорідності і незалежності стокових величин.

Розподіл імовірностей  $y = f(z)$  повинен задовольняти такі умови: на початку та на кінці графіка щільності ймовірностей  $y = 0$  та між початком і кінцем досягає максимального значення.

Отже, перша похідна  $\frac{dy}{dz}$  повинна дорівнювати нулю у трьох точках: на початку, в кінці та у точці, яка відповідає моді.

Величина  $z$  представляє собою центровану і нормовану вихідну величину  $x$ , тобто

$$z = \frac{x - m_x}{m_x} = k - 1, \quad (1.104)$$

де  $k$  - модульний коефіцієнт.

Сукупність теоретичних кривих розподілу випадкових величин Пірсона можна отримати в результаті вирішення диференціального рівняння

$$\frac{dy}{dz} = \frac{y(z+d)}{\varphi(z)}, \quad (1.105)$$

де  $d$  - відстань між модою та математичним сподіванням на графіку кривої розподілу (рис.1.8);

$y=f(z)$  - щільність імовірності;

$\varphi(z)$ - ряд Маклорена, який має вигляд

$$\varphi(z) = b_0 + b_1z + b_2z^2 + \dots + b_nz^n. \quad (1.106)$$

Рівняння (1.105) відповідає усім висунутим раніше вимогам. Дійсно, на кінцях розподілу при  $y = 0$ ,  $\frac{dy}{dz} = 0$ ; при  $z = -d$ , тобто в точці

максимуму ( $m_0$ ) значення  $\frac{dy}{dz}$  також дорівнює нулю.

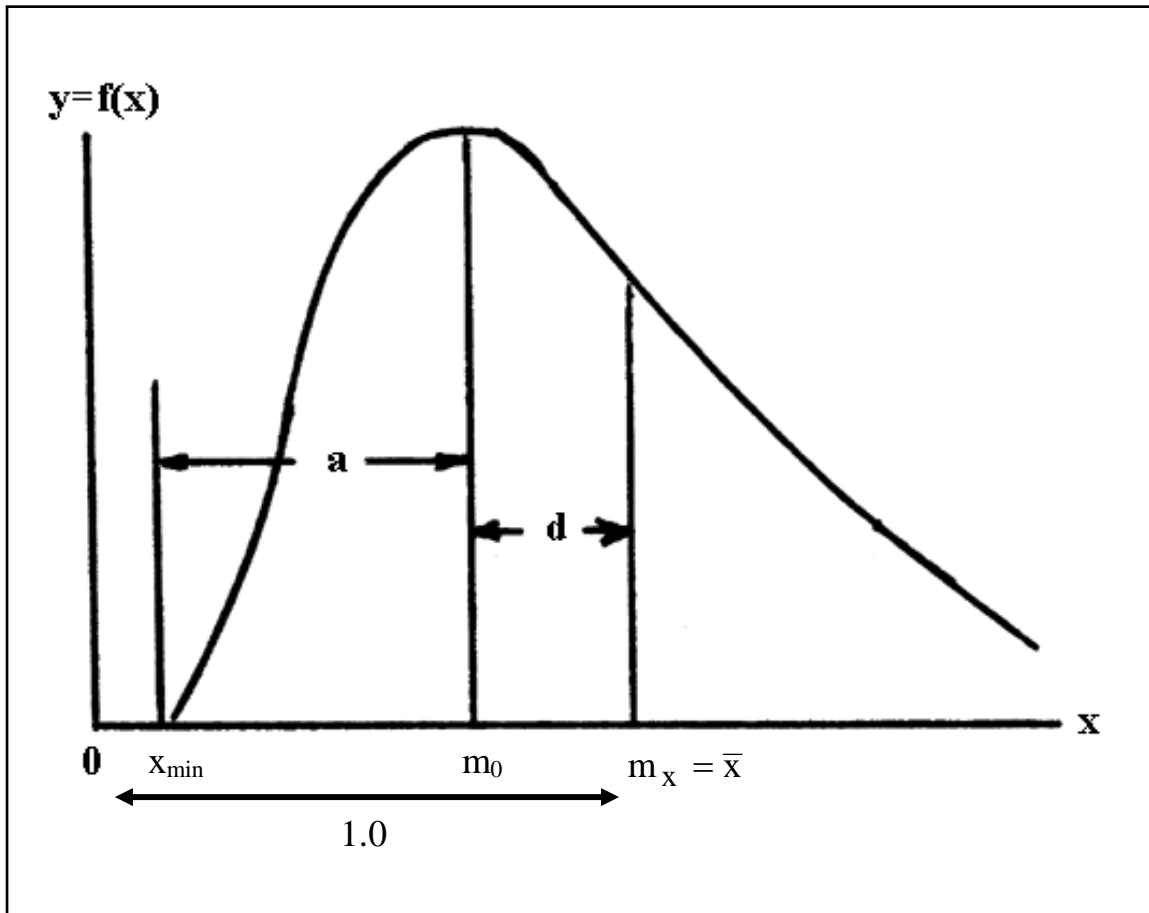


Рисунок 1.10 - Крива Пірсона III ( $C_S > 0$ )

Для побудови кривих розподілу Пірсона використовував тільки три перших члена ряду Маклорена, тобто

$$\varphi(z) = b_0 + b_1 z + b_2 z^2, \quad (1.107)$$

звідки

$$\frac{dy}{dz} = \frac{y(z+d)}{b_0 + b_1 z + b_2 z^2}. \quad (1.108)$$

Значення  $z$  та  $d$  вважаються відомими. Для розрахунків щільності ймовірності необхідно знайти значення  $b$ . Після інтегрування (1.108) приходимо до рівняння

$$-nb_0 \beta_{n-1} - b_1(n+1)\beta_n - b_2(n+2)\beta_{n+1} = \beta_{n+1} + d\beta_n, \quad (1.109)$$

де  $\beta_n$  - центральні моменти порядку  $n$ .

Розглянувши цей вираз при різних  $n$  ( $n = 0,1,2$ ) та враховуючи, що  $\beta_0 = 1$  і  $\beta_1 = 0$ , отримуємо систему рівнянь

$$\begin{cases} -b_1 = d \\ -b_0 - 3b_2\beta_2 = \beta_2 \\ -3b_1\beta_2 - 4b_2\beta_3 = \beta_3 + d\beta_2 \end{cases}, \quad (1.110)$$

яка дозволяє виразити  $b_0$ ,  $b_1$ ,  $b_2$  через центральні моменти другого та третього порядків.

З системи рівнянь (1.110) витікає, що для розрахунків перших трьох параметрів ряду Маклорена, необхідно використовувати третій центральний момент, який визначається за експериментальними даними з великими похибками. Отже подальше збільшення числа членів Маклорена немає сенсу.

В залежності від числових значень коефіцієнтів  $b_0, b_1, b_2$  з вихідного рівняння можна отримати сім різних типів кривих забезпеченостей. В гідрологічних розрахунках використовуються два з них: нормальний та біноміальний асиметричний (Пірсона III).

### ***1.3.5.1 Нормальний закон розподілу випадкових величин***

Нормальний закон розподілу є частковим випадком розподілу Пірсона. Для знаходження відповідної функції щільності ймовірностей Пірсон використав лише перший член ряду Маклорена  $b_0$ , приймаючи, що інші його члени дорівнюють нулю, тобто

$$\frac{dy}{dz} = \frac{y(z+d)}{b_0}. \quad (1.111)$$

Звернемося до загального розв'язку системи рівнянь (1.110):

1. Оскільки  $b_1=0$ , математичне сподівання і мода співпадають, тобто  $m_x=m_0$ .
2. Параметр  $-b_0$  дорівнює дисперсії випадкової величини.
3. Оскільки  $\beta_3=0$ , то  $C_s=0$ , тобто розподіл симетричний.

Розв'язання диференціального рівняння (1.111) зводиться до такого

$$\frac{dy}{dz} = -\frac{y(z)}{b_0}; \quad (1.112)$$

$$\int \frac{dy}{y} = \int -\frac{z}{\sigma_z^2} dz; \quad (1.113)$$

$$\ln|y| = -\frac{1}{\sigma_z^2} \frac{z^2}{2} + c \quad (1.114)$$

Але ж  $y = f(x) \geq 0$ , тоді  $|y| = y$  та  $\ln|y| = \ln y$ , тобто

$$\ln y = -\frac{1}{\sigma_z^2} \frac{z^2}{2} + c. \quad (1.115)$$

У точці  $z=0$  величина  $c = \ln y$ . Позначимо її як  $y_0$ . Це значення буде максимальним через те, що для всіх інших  $z$   $\ln y < \ln y_0$ . Крім того, якщо  $d=0$ , то значення  $k$ , відносно якого ведуться розрахунки, буде співпадати з модою, тобто мати найбільшу щільність імовірності. З урахуванням цього

$$y = y_0 e^{-\frac{z^2}{2\sigma_z^2}}. \quad (1.116)$$

З властивостей щільності ймовірності витікає, що

$$\int_{-\infty}^{\infty} f(z) dz = 1. \quad (1.117)$$

Отже,

$$y_0 \int_{-\infty}^{\infty} e^{-\frac{z^2}{2\sigma_z^2}} dz = 1. \quad (1.118)$$

Після деяких перетворень отримаємо

$$y_0 = \frac{1}{\sigma_z \sqrt{2\pi}}. \quad (1.119)$$



Загальний вигляд рівняння кривої нормального розподілу такий

$$f(z) = \frac{1}{\sigma_z \sqrt{2\pi}} e^{-\frac{z^2}{2\sigma_z^2}}. \quad (1.120)$$

Для вихідного ряду  $x$

$$f(x) = y = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}. \quad (1.121)$$

Якщо замість значень  $x$  використати нормовані значення, представлені у вигляді

$$t = \frac{x - m_x}{\sigma_x} = \frac{k - 1}{Cv}, \quad (1.122)$$

і, ураховуючи, що  $\sigma_t = 1$ , отримаємо

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}. \quad (1.123)$$

Вираз (1.123) задається у вигляді таблиці для практичного використання. Нормальний розподіл має декілька особливостей.

1. У зв'язку з тим, що функція (1.121) має дійсні значення при будь-яких значеннях незалежної змінної  $X$ , область її визначення така:  $-\infty < x < +\infty$ .
2. Функція  $f(x)$  - парна, тобто  $f(x) = f(-x)$ .
3. Нормальна крива не перетинає осі  $x$ .
4. Крива симетрична відносно моди.
5. Це двопараметричний розподіл.
6. Параметр  $\sigma_x$  є характеристикою форми кривої: чим більше  $\sigma_x$ , максимальна ордината менше, а крива розтягується.
7. Якщо змінювати  $m_x$ , то крива щільності ймовірностей буде переміщуватися уздовж осі  $x$ , зберігаючи свою форму.

Вперше нормальний закон розподілу випадкових величин був розроблений для аналізу похибок вимірювань. На цій основі він і отримав розповсюдження у багатьох галузях науки і техніки, в тому числі і в гідрології, де він широко використовується для оцінки точності розрахунків, визначення довірчих інтервалів і таке інше. У практичних

дослідженнях іноді допускається використовувати його до деяких рядів з асиметричним розподілом. Однак, похибка прийнятого припущення про нормальність розподілу може привести до неправильних рішень гідрологічних задач.

### 1.3.5.2. Розподіл Пірсона III

Диференціальне рівняння Пірсона III (1.105) може бути розписаним для перших двох членів ряду Маклорена. Для нормального закону розподілу була прийнята жорстка умова про те, що розподіл симетричний, тобто мода розподілу співпадає з математичним сподіванням і радіус асиметрії  $d=0$ .

Припустимо, що випадкова величина  $X$  позитивна, отже  $x_{min} \geq 0$ . Позначимо через  $a$  відстань від  $m_x$  до моди (рис. 1.8). Тоді сума  $a + d$  є відстанню від математичного сподівання до  $m_x$ .

Звідки

$$a + d = m_x - x_{min} = 1 - k_{min}, \quad (1.124)$$

де  $k_{min} = \frac{x_{min}}{m_x}$ .

Через те, що  $k_{min} \geq 0$ , то  $a + d \leq 1$ .

Рівняння біноміального розподілу Пірсона III типу може бути знайдено, коли розглянути два перших члени ряду Маклорена ( $b_0$  та  $b_1$ ). Тобто,

$$\frac{dy}{dz} = \frac{y(z + d)}{b_0 + b_1 z}. \quad (1.125)$$

Після інтегрування (1.125) при використанні системи (1.110) вираз біноміального розподілу Пірсона III типу набуває вигляду:

$$f(z) = y = y_0 e^{-z/d} (1 + z/a)^{a/d}, \quad (1.126)$$

де  $a$  - відстань від початку кривої до моди ;  
 $y_0$  - модальна ордината.

Загальний вигляд кривої Пірсона III при  $C_v > 0$  (позитивній асиметрії) показаний на рис.1.8.

Параметри рівняння (1.126) зв'язані з центральними моментами розподілу

$$a + d = 2\beta_2^2 / \beta_3. \quad (1.127)$$

Для практичного застосування в (1.127)  $\beta_2$  та  $\beta_3$  замінюють на статистичні параметри  $m_x$ ,  $C_V$ ,  $C_S$ . Знаючи, що  $\beta_2 = \sigma_x^2$ , а  $C_V = \sigma_x / \bar{x}$ , запишемо

$$\beta_2 = C_V^2 m_x^2. \quad (1.128)$$

Третій центральний момент, як було показано, зв'язаний з коефіцієнтом асиметрії  $C_S$  та середнім квадратичним відхиленням

$$\beta_3 = C_S \sigma_x^3 = C_S C_V^3 m_x^3. \quad (1.129)$$

Підставляючи (1.128) та (1.129) в (1.127), отримуємо

$$a + d = 2C_V^4 m_x^4 / (C_S C_V^3 m_x^3) = 2C_V m_x / C_S. \quad (1.130)$$

В загальному вигляді

$$m_x - x_{min} = 2C_V m_x / C_S. \quad (1.131)$$

Або

$$1 - k_{min} = 2C_V / C_S. \quad (1.132)$$

де  $k_{min} = x_{min} / m_x$  - модульний коефіцієнт мінімального значення ряду.

Розглянемо три можливих варіанти:

- 1) якщо  $k_{min} = 0$ ,  $C_S = 2C_V$
- 2) якщо  $k_{min} > 0$ ,  $C_S > 2C_V$
- 3) якщо  $k_{min} < 0$ ,  $C_S < 2C_V$

Третій випадок суперечить природі стокових рядів, які завжди позитивні, отже крива розподілу Пірсона III може використовуватись тільки тоді, коли  $C_S \geq 2C_V$ . З цього слід зазначити таке: крива обмежена нижньою ( $x=0$ ) і не обмежена верхньою границями. При  $x \rightarrow \infty$  крива

наближається до осі абсцис; опирається вона на три статистичні параметри:  $m_x$ ,  $C_V$ ,  $C_S$ , а при  $x_{min}=0$  є двопараметричним розподілом.

### 1.3.5.3 Логарифмічно-нормальний закон розподілу

Нормальний закон розподілу використовується для випадкових величин, котрі змінюються в межах від  $-\infty$  до  $+\infty$ . Для того, щоб використати цей закон для стокових величин, область визначення яких змінюється від 0 до  $+\infty$ , необхідно перетворити вихідну змінну.

У нашому розпорядженні є суттєво позитивна випадкова величина  $X$  з областю визначення  $0 \leq x < \infty$ , несиметрична. Перетворимо її у нову, яка підлягає нормальному закону розподілу. З цією метою введемо нову змінну  $u = \ln x$ . Якщо  $0 \leq x < +\infty$ , то  $-\infty < \ln x < +\infty$ . Тоді рівняння (1.121) набере вигляду

$$y = f(u) = \frac{1}{\sigma_u \sqrt{2\pi}} e^{-\frac{(u-\bar{u})^2}{2\sigma_u^2}}. \quad (1.133)$$

Параметри випадкової величини  $u$  можна представити наступним чином

$$\hat{m}_u = \frac{\sum_{i=1}^n \ln x_i}{n} = \hat{m}_{\ln x}; \quad (1.134)$$

$$\sigma_{\ln u} = \sqrt{\frac{\left( \sum_{i=1}^n (\ln x_i - \hat{m}_{\ln x})^2 \right)}{n-1}}. \quad (1.135)$$

Для побудови відповідної кривої розподілу спочатку у такому випадку використовують таблицю нормального закону розподілу.

З метою отримання закону розподілу, який можна застосувати до величин стоку, слід виконати трансформацію не вихідної змінної, а самого закону.

Нехай нам відомий закон розподілу деякої величини  $u$ . Необхідно отримати закон розподілу величини  $x$ , розподіл якої невідомий, але  $x$  функціонально зв'язаний з  $u$ , тобто

$$x = \varphi(u), \text{ а } u = \psi(x), \quad (1.136)$$

тоді новий закон розподілу представляється у вигляді

$$f(x) = f(u) \frac{du}{dx}. \quad (1.137)$$

Якщо  $u = \ln x$  підлягає нормальному закону розподілу, то

$$\frac{du}{dx} = \frac{1}{x}, \quad (1.138)$$

і

$$f(x) = f(u) \frac{1}{x}. \quad (1.139)$$

Отже, рівняння логарифмічно-нормального закону розподілу запишеться у вигляді

$$f(x) = \frac{1}{\sigma_{\ln x} \sqrt{2\pi}} e^{-\frac{(\ln x - m_{\ln x})^2}{2\sigma_{\ln x}^2}} \frac{1}{x}. \quad (1.140)$$

**Логарифмічно-нормальний закон розподілу використовується для рядів з високим коефіцієнтом асиметрії, а саме**

$$C_s = C_v(3 + C_v). \quad (1.141)$$

#### **1.3.5.4 Крива трипараметричного гама-розподілу С.М. Крицького і М.Ф. Менкеля**

Трипараметричний гама-розподіл базується на модифікації рівняння кривої Пірсона III при умові  $C_s = 2C_v$ . Для цього випадку рівняння Пірсона III представляється у вигляді гама-функції

$$f(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)} e^{-\alpha z} z^{\alpha-1}, \quad (1.142)$$

де  $z$  - значення випадкової величини в модульних коефіцієнтах, тобто

$$z = k = \frac{z_i}{\bar{z}};$$

$\Gamma(\alpha)$  - гама-функція;

$$\alpha = \frac{4}{C_s^2} = \frac{1}{C_v^2}.$$

Рівняння гама-функції є двопараметричним, через те що залежить тільки від двох аргументів  $z$  і  $C_v$ . Позитивною якістю гама-функції є те, що значення  $z$  змінюється від 0 до  $\infty$ , що відповідає фізичній природі стоку. Однак, відношення  $C_s = 2C_v$  є одним з часткових випадків можливих розподілів. Щоб зробити рівняння (1.142) більш загальним, вихідна величина  $z$  була трансформована. Замість  $z$  вводиться нова змінна  $u$ , причому

$$z = au^h \text{ та } u = \left(\frac{z}{a}\right)^{1/h}, \quad (1.143)$$

де  $a$  та  $h$  - параметри, які є функціями  $C_v$  та  $C_s$ .

Тоді

$$f(u) = \frac{\alpha^\alpha}{\Gamma(\alpha)} e^{-\alpha u} u^{\alpha-1}, \quad (1.144)$$

а

$$\frac{du}{dz} = \frac{1}{h} \left(\frac{1}{a}\right)^{1/h} z^{1/h-1}. \quad (1.145)$$

Згідно з (1.137), новий закон розподілу можна записати у вигляді

$$f(z) = \frac{\alpha^\alpha}{\Gamma(\alpha)} e^{-\alpha u} u^{\alpha-1} \frac{1}{h} \left(\frac{1}{a}\right)^{1/h} z^{1/h-1}. \quad (1.146)$$

Якщо замінити  $u$  на  $z$  і врахувати, що  $z = k$ , то

$$f(k) = \frac{\alpha^\alpha}{\Gamma(\alpha)} \left(\frac{1}{a}\right)^{\alpha/h} \frac{1}{h} e^{-\alpha\left(\frac{k}{a}\right)^{1/h}} k^{a/h-1}, \quad (1.147)$$

де  $a$  і  $h$  - параметри розподілу, які зв'язані із стандартними параметрами – коефіцієнтами варіації та асиметрії.

Розрахунковий алгоритм дуже складний і вирішення надається у вигляді статистик  $\lambda_1, \lambda_2, \lambda_3$ .

Головна перевага кривих розподілу Крицького-Менкеля полягає у тому, що при будь-яких співвідношеннях між  $C_s$  та  $C_v$  нижня границя ознаки розподілу завжди дорівнює нулю.

*При використанні трипараметричного гама-розподілу слід пам'ятати, що як статистичні параметри, можна використовувати тільки параметри, визначені за методом найбільшої правдоподібності.*

### 1.3.6 Клітчатка ймовірностей, їх призначення

Криві розподілу, побудовані в декартових шкалах координат, які використовуються в гідрології, мають вигнутий або опукло-вгнутий вигляд. Тому виникають проблеми екстраполяції емпіричних кривих забезпеченостей в області малих та великих забезпеченостей. Для усунення цих недоліків використовують клітчатка ймовірностей, які значною мірою вирівнюють залежності. Вирівнювання досягається за рахунок трансформації шкали забезпеченостей таким чином, щоб у системі координат розглядався інтегральний закон розподілу або крива забезпеченості у вигляді прямої лінії. Як правило, трансформація відбувається за допомогою графічних методів. Стосовно до закону нормального розподілу, використовується клітчатка з помірною асиметрією. При позитивній асиметрії крива має дещо вгнуту форму, при від'ємній - опуклу.

При використанні логарифмічно-нормального закону розподілу застосовують клітчатка значної асиметрії. Вони можуть бути побудовані на основі клітчатка нормального закону, якщо вісь ординат виразити у вигляді логарифмічної шкали.

Клітчатка ймовірності, вирівнюючі трипараметричний гама-розподіл, будують для рядів, асиметрія яких значно відрізняється від нуля.

### 1.3.7 Розрахунки характеристик стоку заданої забезпеченості за теоретичними законами розподілу

Емпіричні криві забезпеченості не дають можливості визначити величини стоку за межами вихідної інформації. Безпосередньо графічна екстраполяція емпіричної кривої забезпеченості в область невеликих ( $p < 5\%$ ) та великих ( $p > 95\%$ ) забезпеченостей має суб'єктивний характер і може призвести до значних похибок у розрахунках. Теоретичні закони розподілу у даному випадку "вирівнюють" емпіричний розподіл стокових величин та надають можливості екстраполювати його за межі спостережених даних.

Найчастіше у гідрологічних розрахунках застосовують криву Пірсона III та трипараметричний гама-розподіл С.Н.Крицького та М.Ф.Менкеля. В СНіП 2.01.14-83 рекомендується використовувати закон Пірсона III при  $C_v \leq 0.5$  та  $C_s \geq 2C_v$ , а при  $C_v > 0.5$  та  $C_s < 2C_v$  - трипараметричний гама-розподіл. Ці теоретичні розподіли представлені у вигляді таблиць. Наприклад, ординати теоретичної кривої забезпеченості Пірсона III подаються у вигляді нормованих відхилень від середньої величини стоку

$$\Phi(P, C_s) = \frac{x_P - \bar{x}}{\sigma_x} = \frac{k_P - 1}{C_v}, \quad (1.148)$$

а ординати трипараметричного гама-розподілу - у вигляді модульних коефіцієнтів

$$k(P, C_v) = \frac{x_P}{\bar{x}} \quad (1.149)$$

при заданому відношенні  $C_s/C_v$ .

Щоб скористатися таблицями теоретичних розподілів, спочатку необхідно визначити оцінки статистичних параметрів по матеріалах спостережень за стоком. Слід зазначити, що при використанні методу найбільшої правдоподібності, для розрахунків вибірових значень статистичних параметрів за теоретичний розподіл можна прийняти тільки трипараметричний гама-розподіл, тому що формули методу найбільшої правдоподібності виведені саме виходячи з цього закону. Що стосується закону розподілу Пірсона III, то при  $C_s < 2C_v$  та при  $P > 90\%$  ординати теоретичної кривої забезпеченості можуть мати від'ємні значення, а це суперечить фізичній суті стокових величин.



Розрахунки характеристик стоку за теоретичним розподілом Пірсона III виконуються за виразом

$$x_p = (\Phi_{P,C_s} C_v + 1) \bar{x}, \quad (1.150)$$

а при застосуванні трипараметричного гама-розподілу -

$$x_p = k_p \bar{x}. \quad (1.151)$$

#### 1.4 Приклади розрахунків

**Приклад 1. Розрахувати статистичні параметри річного стоку з водозбору р.Ока – м.Калуга за період з 1936 по 1970 рр., використовуючи метод моментів.**

Вихідні дані та результати обчислень повинні бути оформлені в таблиці такого вигляду:

Таблиця 1.2 – Розрахунок статистичних параметрів річного стоку за методом моментів (р.Ока – м.Калуга)

№ пп	Рік	Q <sub>i</sub> , м <sup>3</sup> /с	$k_i = \frac{Q_i}{\bar{Q}}$	$k_i - 1$	$(k_i - 1)^2$	$(k_i - 1)^3$
1	1936	281	1.65	0.65	0.4225	0.2746
2	1937	304	1.38	0.38	0.1444	0.0555
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....
.....	.....	.....	.....	.....	.....	.....
35	1970	450	0.65	-0.35	0.1225	0.04274
Сума		9531	35.08	-0.01	2.0056	0.4247

**Розрахунок здійснюється в такій послідовності:**

1. Обчислюється середня арифметична витрата води  $\bar{Q}$  (м<sup>3</sup>/с) за формулою (1.16)

$$\text{При } n=35 \quad \bar{Q} = \sum_{i=1}^n Q_i = \frac{9531}{35} = 272;$$

2. Коефіцієнт варіації  $C_v$  обчислюється за формулою (1.34)

$$\hat{C}_v = \sqrt{\frac{\sum_{i=1}^n (k_i - 1)^2}{n-1}} = \sqrt{\frac{2.0056}{34}} = 0.24;$$

3. Коефіцієнт асиметрії обчислюється за виразом (1.40)

$$\hat{C}_s = \frac{\sum_{i=1}^n (k_i - 1)^3}{C_v^3} \frac{n}{(n-1)(n-2)} = \frac{35 \times 0.4247}{34 \times 33 \times 0.0138} = \frac{14.9}{16.0} = 0.96;$$

4. Середнє квадратичне відхилення визначається у м<sup>3</sup>/с за (1.32) або (1.33).

Зокрема, при  $C_v = 0.24$  і  $\bar{Q} = 272$  м<sup>3</sup>/с

$$\sigma_Q = C_v \cdot \bar{Q} = 0.24 \times 272 = 65.3;$$

5. Випадкові похибки статистичних параметрів обчислені за формулами, наведеними в п. 1.3.2.3:

за формулою (1.54)

$$\sigma_{\bar{Q}} = \frac{\sigma_Q}{\sqrt{n}} = \frac{65.3}{\sqrt{35}} = \frac{65.3}{5.92} = 11.0;$$

за формулою (1.56)

$$\sigma_{C_v} = \frac{C_v}{n + 4C_v^2} \sqrt{\frac{n(1 + C_v^2)}{2}} = \frac{0.24}{35 + 4 \times 0.24^2} \sqrt{\frac{35(1 + 0.24^2)}{2}} = 0.029;$$

за формулою (1.57)

$$\sigma_{C_s} \sqrt{\frac{6}{n}(1 + 6C_v^2 + 5C_v^4)} = \sqrt{\frac{6}{35}(1 + 6 \times 0.24^2 + 5 \times 0.24^4)} = 0.48;$$

6. Відносні випадкові похибки визначення статистичних параметрів визначаються за формулами (1.59)-(1.61):

$$\varepsilon_{\bar{Q}} = \frac{C_v}{\sqrt{n}} 100\% = \frac{0.24}{5.92} 100\% = 4.05\%;$$

$$\varepsilon_{C_v} = \frac{\sigma_{C_v}}{C_v} 100\% = \frac{0.029}{0.24} 100\% = 12.1\%;$$

$$\varepsilon_{C_s} = \frac{\sigma_{C_s}}{C_s} 100\% = \frac{0.48}{0.93} 100\% = 51.6\%.$$

На підставі отриманих результатів можна зробити висновок, що точність розрахунків статистичних параметрів  $\bar{Q}$  та  $C_v$  задовільна, оскільки  $\varepsilon_{\bar{Q}} < 5\%$ , а  $\varepsilon_{C_v} < 15\%$ .

Відносне середньоквадратичне відхилення параметра  $C_s$  становить 51.6%, тобто цей параметр рекомендується розраховувати по середньому для розглядуваного регіону співвідношенню  $C_s / C_v$ .

**Приклад 2. Розрахувати статистичні параметри річного стоку з водозбору р. Сіверський Донець – м. Біла Калитва за період з 1934 по 1962 роки, використовуючи метод найбільшої правдоподібності.**

Вихідні дані та результати обчислень оформлюються у вигляді таблиці (табл. 1.3).

Довжина ряду становить 24 роки.

Для визначення статистик  $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3$  використані формули (1.81), (1.84), (1.85). Зокрема,

$$\bar{Q} = \lambda_1 = \frac{\sum_{i=1}^n Q_i}{n} = \frac{3526}{24} = 147;$$

$$\hat{\lambda}_2 = \frac{\sum_{i=1}^n \lg k_i}{n-1} = \frac{-1.103}{23} = -0.0480$$

$$\hat{\lambda}_3 = \frac{\sum_{i=1}^n k_i \lg k_i}{n-1} = \frac{1.072}{23} = 0.0466$$

Таблиця 1.3 – Обчислення статистичних параметрів ряду максимальних витрат води за методом найбільшої правдоподібності р. Сіверський Донець – м.Біла Калитва, 1934-1962 рр.

№ пп	Рік	$Q_i, \text{м}^3/\text{с}$	$k_i$	$\lg k_i$	$k_i \lg k_i$
1	1934	271	1.84	0.404	1.013
...	.....	.....	.....	.....	.....
...	.....	.....	.....	.....	.....
24	1962	54.9	0.37	-0.432	-0.160
Сума		3526	24	-1.103	1.072

Примітка: у цьому часовому ряді мають місце пропуски спостережень (1941-45 рр.)

За номограмами, наведеними в СНіП 2.01.14-83, знаходимо значення  $C_v$  та співвідношення  $C_s / C_v$ , які дорівнюють 0.48 і 1.44, відповідно до р. Сіверський Донець – м.Біла Калитва.

Таким чином,  $\bar{Q} = 147 \text{ м}^3 / \text{с}$ ,  $C_v = 0.48$ ,  $C_s = 3C_v = 1.44$ . При цьому:  
 $\sigma_Q = C_v \cdot \bar{Q} = 0.48 \times 147 = 70.6$ ;

$$\sigma_{\bar{Q}} = \frac{\sigma_Q}{\sqrt{n}} = \frac{70.6}{\sqrt{24}} = 14.4;$$

$$\varepsilon_{\bar{Q}} = \frac{C_v}{\sqrt{n}} 100\% = \frac{0.48}{\sqrt{24}} 100\% = 9.8\%.$$

Середньоквадратичне відхилення коефіцієнта варіації при  $C_s = 2C_v$  розраховується за формулою (1.87)

$$\sigma_{C_v} = \frac{C_v}{\sqrt{2n}} \sqrt{\frac{3}{3 + C_v^2}} = \frac{0.48}{\sqrt{2 \times 24}} \sqrt{\frac{3}{3 + 0.48^2}} = 0.067;$$

У нашому випадку  $C_s = 3C_v$ , отже для отриманого  $\sigma_{C_v}$  необхідно ввести поправковий коефіцієнт  $K$ , використовуючи рис.1.5, де  $K = 1.1$ . Тоді  
 $\sigma_{C_v} = 0.0668 \times 1.1 = 0.074$ ,

а

$$\varepsilon_{C_V} = \frac{\sigma_{C_V}}{C_V} 100\% = \frac{0.074}{0.48} 100\% = 15.4\%.$$

Середньоквадратичне відхилення співвідношення  $C_S / C_V$  знаходиться в залежності від  $C_V$  (рис. 1.6), відповідно до якого

$$\sigma_{C_S / C_V} = 0.73,$$

а

$$\varepsilon_{C_S} = \frac{\sigma_{C_S}}{C_S} 100\% = \frac{0.73}{1.44} 100\% = 50.7\%.$$

Отримані результати свідчать про необхідність приведення цього ряду до тривалого періоду, оскільки відносні середньоквадратичні відхилення параметрів  $Q$  та  $C_V$  близькі до припустимих ( $\varepsilon_{\bar{Q}} \leq 10\%$ ;  $\varepsilon_{C_V} \leq 15\%$ ) або дещо перевищують їх.

**Приклад 3. Визначити статистичні параметри максимального стоку весняного водопілля для водозбору р. Половина - с. Половина за період з 1951 по 1967 роки.**

Вихідні матеріали та матеріали розрахунків емпіричної забезпеченості розміщуються за формою, наведеною у скороченому вигляді в табл. 1.4.

Таблиця 1.4 – Розрахунки ординат емпіричної кривої забезпеченості максимального стоку весняного водопілля р.Половина-с.Половина, 1951-1967 рр.

№ пп	Рік	$Q_m, \text{м}^3/\text{с}$	$Q_m, \text{м}^3/\text{с}$	$\hat{p} = \frac{m}{n+1} \cdot 100$
1	2	3	4	5
1	1951	16.1	31.5	5.90
2	1952	14.3	30.1	11.7
...	.....	.....	.....	.....
....	.....	.....	.....	.....
17	1967	24.7	9.71	94.1

Хід виконання роботи:

1. На клітчатку ймовірностей з помірною асиметричністю (рис. 1.9) по гр.4 і 5 наносяться величини  $Q_m = f(\hat{p})$ .

2. По точках проводиться згладжена емпірична крива забезпеченості, для забезпеченостей  $\hat{p} = 5; 50$  і  $95\%$ .

3.3 цієї кривої знімаються три характерні ординати, причому:  $Q_5=35.0$ ;  $Q_{50}=20.0$ ;  $Q_{95}=8.50$  м<sup>3</sup>/с.

4. По значеннях  $Q_5, Q_{50}, Q_{95}$  за виразом (1.95) визначається коефіцієнт скісності  $S$

$$S = \frac{Q_5 + Q_{95} - 2Q_{50}}{Q_5 - Q_{95}} = \frac{35.0 + 8.50 - 2 \cdot 20.0}{35.0 - 8.50} = \frac{3.5}{26.5} = 0.13.$$

5. Використовуючи табл. 1.1, по  $S$  визначається коефіцієнт асиметрії

$$\hat{C}_s = \varphi(S) = 0.45.$$

6. Оціночні значення інших статистичних параметрів (середньоквадратичного  $\sigma_Q$ , коефіцієнта варіації  $C_v$ , середнього арифметичного  $\bar{Q}$ ) розраховуються за допомогою формул (1.97), (1.98) та (1.99), тобто

$$\sigma_Q = \frac{Q_5 - Q_{95}}{\Phi_5 - \Phi_{95}} = \frac{35.0 - 8.50}{3.26} = 8.13,$$

де  $\Phi_5 - \Phi_{95}$  вибираються з табл. 1.1 по  $S$ ;

$$\bar{Q}_m = Q_{50} - \sigma_Q \Phi_{50} = 20.0 - 8.13 \times (-0.08) = 20.6,$$

де  $\Phi_{50}$  - вибирається так само з табл. 1.1 по  $S$ ;

$$\hat{C}_v = \sigma_Q / \bar{Q}_m = 8.13 / 20.6 = 0.39.$$

**Приклад 4. Обчислити ординати кривої забезпеченості Пірсона III за статистичними параметрами  $\bar{Q} = 272\text{м}^3/\text{с}$ ,  $C_v = 0.24$ ,  $C_s = 0.86$  (приклад п.1.4.1)**

За визначеним по методу моментів значенням коефіцієнта асиметрії  $C_s=0.86$  та, використовуючи таблицю до кривої розподілу Пірсона III (її фрагмент наводиться на стор. 50-51 в табл. 1.1), знаходимо нормовані відхилення ординат від одиниці  $\Phi_p$  для різних забезпеченостей  $P$  (табл. 1.5).

Таблиця 1.5 – Ординати кривої забезпеченості річного стоку р.Ока – м.Муром

$C_s$	Величина	Забезпеченість, P%					
		0.01	0.10	1.0	3.0	...	99.9
0.86	$\Phi_p$	5.64	4.32	2.93	2.20	....	-1.95
	$\Phi_p \hat{C}_v$	1.35	1.04	0.70	0.53	....	-0.47
	$k_p = \Phi_p \hat{C}_v + 1$	2.35	2.04	1.70	1.53	....	0.53
	$Q_p = k_p \bar{Q}$	639	551	462	416	....	138

## РОЗДІЛ 2

### ОСНОВИ ТЕОРІЇ КОРЕЛЯЦІЇ

#### 2.1 Системи випадкових величин. Закони розподілу випадкових величин

В процесі узагальнення результатів гідрометеорологічних спостережень при вирішенні інженерно-гідрологічних задач часто доводиться мати справу з декількома (двома чи більш) випадковими величинами, що утворюють систему. Наприклад, зливові опади описуються не однією, а як мінімум двома випадковими величинами: тривалістю дощу та його інтенсивністю. Течія води характеризується швидкістю і напрямком; гідрологічний режим річкової системи може характеризуватися стоком декількох річок і ін.

При вивченні системи випадкових величин вже не можна обмежитися дослідженням властивостей окремих складових: необхідно враховувати залежності між цими складовими.

При аналізі систем випадкових величин корисною є їх геометрична інтерпретація. У найпростішому випадку (саме такий випадок і буде розглядатися надалі) систему двох випадкових величин  $X$  і  $Y$  можна зобразити у вигляді точки на площині з координатами  $x$  та  $y$ .

Як і для незалежних випадкових величин, можливий повний опис імовірнісних властивостей системи шляхом побудови їхнього закону розподілу і частковий, заснований на обчисленні окремих числових характеристик.

Функція розподілу системи двох випадкових величин  $X$  і  $Y$  є ймовірністю спільного виконання двох нерівностей:  $X < x, Y < y$ ; тобто

$$F(x, y) = p((X < x), (Y < y)). \quad (2.1)$$

Геометрична інтерпретація визначення (2.1) полягає в тому, що  $F(x, y)$  є ймовірністю попадання випадкової величини  $(X, Y)$  в нескінченний квадрант з вершиною в точці  $(x, y)$ , що лежить лівіше і нижче її (рис. 2.1).



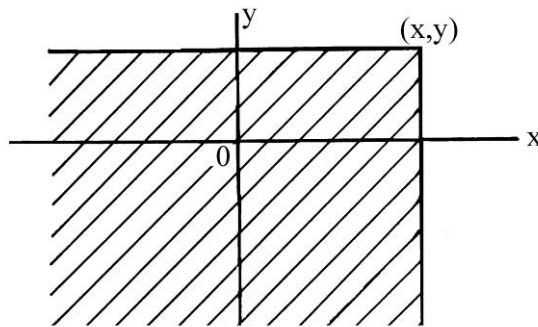


Рисунок 2.1 – Геометрична інтерпретація рівняння (2.1)

Основні властивості функції розподілу системи двох випадкових величин  $F(x, y)$  такі:

-  $F(x, y)$  є неспадною функцією обох аргументів, тобто

$$\text{при } x_2 < x_1 \quad F(x_2, y) \geq F(x_1, y); \quad (2.2)$$

$$\text{при } y_2 < y_1 \quad F(x, y_2) \geq F(x, y_1) \quad (2.3)$$

Дійсно, зміщуючи праву границю квадранта вправо, або зміщуючи верхню грань вгору, не можна зменшити ймовірність попадання у квадрант.

- при  $x \rightarrow -\infty$  або  $y \rightarrow -\infty$  функція розподілу дорівнює нулю

$$F(x, -\infty) = F(-\infty, y) = F(-\infty, -\infty) = 0. \quad (2.4)$$

- при одному з аргументів, який дорівнює  $+\infty$ , функція розподілу системи перетворюється у функцію розподілу випадкової величини, що відповідає іншому аргументу

$$F(x, +\infty) = F_1(x), \quad F(+\infty, y) = F_2(y), \quad (2.5)$$

де  $F_1(x)$  і  $F_2(y)$  - відповідно функції розподілу випадкових величин  $X$  і  $Y$ .

- якщо обидва аргументи відповідають  $+\infty$ , то функція розподілу системи дорівнює одиниці

$$F(+\infty, +\infty) = 1. \quad (2.6)$$

Для подальшого викладу необхідно розглянути наступну задачу: виразити через функцію розподілу системи ймовірність попадання випадкової величини в заданий прямокутник, обмежений абсцисами  $a$  і  $b$  та ординатами  $c$  і  $d$ .

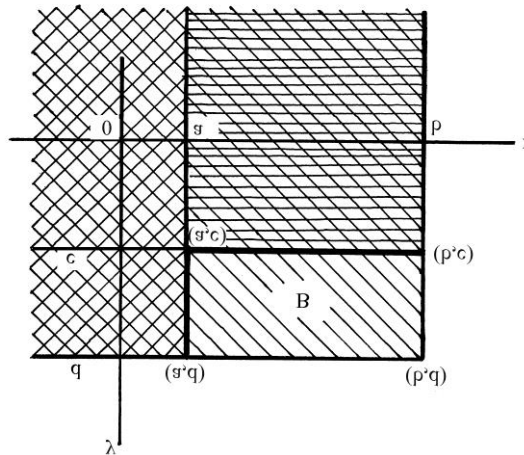


Рисунок 2.2 – Ймовірність попадання у заданий прямокутник

Як видно з рис. 2.2, ймовірність попадання в прямокутник  $B$  дорівнює ймовірності попадання в квадрант  $(b,d)$  мінус імовірність влучення в квадрант  $(a,d)$ , мінус ймовірність попадання в квадрант  $(b,c)$  і плюс ймовірність попадання в квадрант  $(a,c)$ , оскільки відбулося подвійне відняття ймовірності попадання в цей квадрант. Таким чином,

$$p[(X, Y) \in B] = F(b, d) - F(a, d) - F(b, c) + F(a, c). \quad (2.7)$$

Розподіл системи безперервних випадкових величин може бути охарактеризований не тільки функцією розподілу, але й *щільністю розподілу*. Щільність розподілу однієї випадкової величини визначалася як границя відношення ймовірності попадання на малу ділянку до довжини цієї ділянки при її зменшенні. Аналогічно визначимо і щільність розподілу системи двох випадкових величин (рис. 2.3).

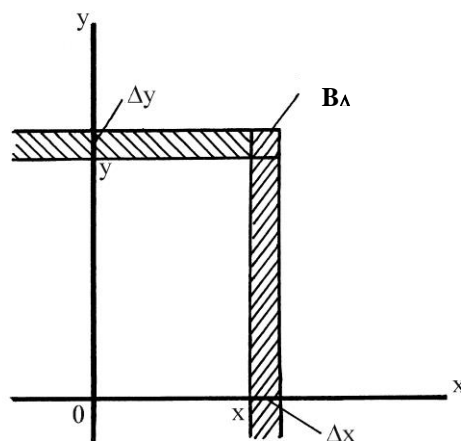


Рисунок 2.3 – Прямокутник  $B_{\Delta}$ .

З цією метою обчислимо границю відношення ймовірності попадання в прямокутник  $B_{\Delta}$  до його площі (рис. 2.3):

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{F(x + \Delta x, y + \Delta y) - F(x + \Delta x, y) - F(x, y + \Delta y) + F(x, y)}{\Delta x \Delta y} \quad (2.8)$$

Якщо  $F(x, y)$  не тільки безперервна, але й диференціюється, то вираз (2.8) являє собою другу змішану частинну похідну функції  $F(x, y)$  по  $x$  і по  $y$ . Позначимо цю похідну як  $f(x, y)$

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} \quad (2.9)$$

і назвемо щільністю розподілу системи.

Для  $f(x, y)$  можлива геометрична інтерпретація у вигляді деякої поверхні, яка називається *поверхнею розподілу* (рис. 2.4).

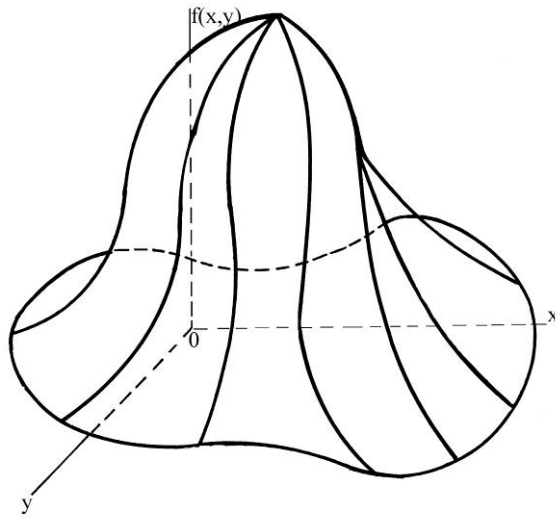


Рисунок 2.4 – Поверхня розподілу двох випадкових величин

За аналогією з однією випадковою величиною можна ввести поняття елемента ймовірності

$$f(x, y) dx dy, \quad (2.10)$$

який представляє собою ймовірність попадання в елементарний прямокутник зі сторонами  $dx, dy$ , котрі примикають до точки  $(x, y)$ .

Використовуючи поняття елемента ймовірності, одержимо вираз для ймовірності попадання випадкової величини в довільну область  $G$ . Шукана ймовірність може бути отримана інтегруванням елементів ймовірності по всій області  $G$

$$p[(X, Y) \in G] = \iint_G f(x, y) dx dy. \quad (2.11)$$

З формули (2.11) випливає формула для ймовірності попадання в прямокутник  $B$ , обмежений абсцисами  $a$  і  $b$  та ординатами  $c$  і  $d$

$$p[(X, Y) \in B] = \int_a^b \int_c^d f(x, y) dx dy \quad (2.12)$$

Використовуючи формулу (2.12), виразимо функцію розподілу системи  $F(x, y)$  через щільність розподілу  $f(x, y)$ . Оскільки інтегральна функція розподілу є ймовірністю попадання в необмежений квадрант, який можна

розглядати як прямокутник з абсцисами  $-\infty$  й  $x$  і ординатами  $-\infty$  і  $y$ , то відповідно до формули (2.12), отримаємо

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy. \quad (2.13)$$

Звернемо увагу на дві властивості щільності розподілу системи випадкових величин:

1. Щільність розподілу системи є функція невід'ємна

$$f(x, y) \geq 0,$$

оскільки щільність розподілу системи є границею відношення двох невід'ємних величин і від'ємною бути не може.

2. Подвійний інтеграл у нескінченних межах від щільності розподілу системи дорівнює одиниці

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1, \quad (2.14)$$

оскільки формула (2.14) виражає в даному випадку ймовірність достовірної події.

## 2.2 Умовні закони розподілу

Як відзначалося в попередньому розділі, знаючи закон розподілу системи двох випадкових величин, можна визначити розподіл окремих величин, що входять до системи:

$$F_1(x) = F(x, \infty); \quad F_2(y) = F(\infty, y)$$

чи інакше

$$F_1(x) = \int_{-\infty}^x \int_{-\infty}^{\infty} f(x, y) dx dy. \quad (2.15)$$

Після диференціювання (2.15) по  $x$ , одержимо вираз для щільності розподілу величини  $X$ :

$$f_1(x) = F_1'(x) = \int_{-\infty}^{\infty} f(x, y) dy. \quad (2.16)$$

Щільність розподілу величини  $Y$  запишемо аналогічно:

$$f_2(y) = F_2'(y) = \int_{-\infty}^{\infty} f(x, y) dx. \quad (2.17)$$

Формули (2.16) і (2.17) показують, що для того, щоб одержати розподіл окремої величини, яка входить до системи, необхідно проінтегрувати щільність розподілу системи по всій області визначення по аргументу, який відповідає іншій випадковій величині.

Обернену задачу в загальному випадку вирішити неможна: знаючи розподіл окремих величин, що входять до системи, не завжди можна знайти закон розподілу системи.

При цьому недостатньо знати розподіл кожної з величин, потрібно знати залежність між ними, яка може бути охарактеризована за допомогою умовних законів розподілу.

Умовним законом розподілу величини  $X$ , яка входить до системи  $(X, Y)$ , називається її закон розподілу, обчислений за умови, що інша випадкова величина  $Y$  набула визначеного значення  $y$ .

Умовний закон розподілу, який може бути заданий функцією розподілу або щільністю, позначаються в такий спосіб:

$$F(x/y) \text{ чи } f(x/y).$$

Маючи розподіл однієї з величин, що входять до системи, і умовний розподіл другої, можна одержати закон розподілу системи, для чого знову скористаємося поняттям елемента ймовірності. Розглянемо прилеглий до точки  $(x, y)$  елементарний прямокутник  $Bd$  зі сторонами  $dx$  і  $dy$  (рис. 2.5).

Ймовірність попадання в цей прямокутник (елемент ймовірності  $f(x, y) dx dy$ ) дорівнює ймовірності одночасного попадання в елементарні смуги  $I$  і  $II$ . Ймовірність добутку цих двох подій за теоремою множення ймовірностей дорівнює ймовірності попадання в елементарну смугу  $I$ , помножену на умовну ймовірність попадання в елементарну смугу  $II$ .

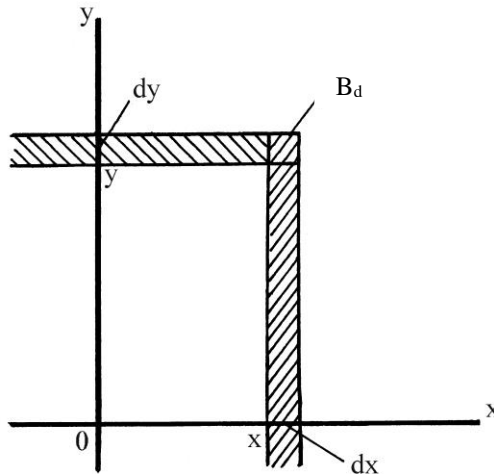


Рисунок 2.5 – Елементарний прямокутник  $B_d$ .

За умови, що перша подія мала місце,

$$f(x, y)dxdy = f_1(x)dx \cdot f(y/x)dy \quad (2.18)$$

або

$$f(x, y) = f_1(x) \cdot f(y/x). \quad (2.19)$$

Формулу (2.19) іноді називають теоремою множення законів розподілу, вона може бути переписана й в іншому вигляді, якщо фіксується значення випадкової величини  $Y$

$$f(x, y) = f_2(y) \cdot f(x/y) \quad (2.20)$$

Її також можна представити стосовно до умовного розподілу

$$f(y/x) = \frac{f(x, y)}{f_1(x)} \quad (2.21)$$

Зміст формул (2.19) і (2.20) може бути описаний таким чином: щільність розподілу системи двох випадкових величин дорівнює щільності розподілу однієї з них, помноженої на умовну щільність іншої, обчисленої за умови, що перша величина набула конкретного значення.

Для того, щоб уявлення було більш повним, звернемося до дискретних випадкових величин.

Нехай можливі результати випробувань становлять  $n$  випадків. Припустимо, що події  $A$  сприяють  $m$  випадків, а події  $B$  -  $k$  випадків.

Існують випадки, які сприяють події  $A$  і події  $B$  одночасово, кількість таких випадків становить  $\ell$  (рис. 2.6).

Тоді

$$p(A, B) = \frac{1}{n}; \quad p(A) = \frac{m}{n}.$$

Обчислимо  $p(B/A)$ , тобто умовну ймовірність події  $B$ , припускаючи, що подія  $A$  мала місце. Якщо відомо, що подія  $A$  відбулася, то з усіх раніше можливих  $n$  випадків залишаються можливими тільки ті  $m$ , які сприяли події  $A$ . З них  $\ell$  випадків сприяють події  $B$ . Отже,

$$p(B/A) = \frac{\ell}{m}.$$

Підставляючи вираз  $p(AB)$ ,  $p(A)$  і  $p(B/A)$  в формулу (2.19), отримуємо тотожність

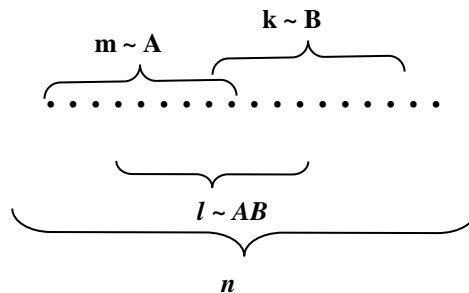


Рисунок 2.6 – Схема випадків для доведення теореми множення

### 2.3 Залежні і незалежні випадкові величини, кореляційний момент, коефіцієнт кореляції

Характер залежності випадкових величин, що утворюють систему, може бути різним. У деяких випадках залежність може бути настільки тісною, що, знаючи значення однієї випадкової величини, можна точно вказати значення іншої. Така надійна залежність близька до функціональної. Якщо ж, знаючи значення випадкової величини  $X$ , не можна точно вказати значення випадкової величини  $Y$ , а тільки закон її розподілу в залежності від значення, яке прийняла величина  $X$ , то така



залежність називається ймовірнісною чи стохастичною. Іншими словами, у стохастичній залежності зміна випадкової величини  $X$  зумовлює зміну розподілу випадкової величини  $Y$ . Цей розподіл може бути описаний рівнянням умовного закону розподілу вигляду (2.21). Функціональна залежність розглядається при цьому як крайній, граничний випадок ймовірнісної залежності. В іншому крайньому випадку залежність між випадковими величинами може бути слабкою настільки, що ці величини можна вважати незалежними.

Сформулюємо найважливіше для подальшого викладу поняття про незалежні випадкові величини. Випадкова величина  $Y$  називається незалежною від випадкової величини  $X$ , якщо закон розподілу  $Y$  не залежить від того, яке значення прийняла величина  $X$

$$f(y/x) = f_2(y). \quad (2.22)$$

Якщо  $Y$  залежить від  $X$ , то

$$f(y/x) \neq f_2(y). \quad (2.23)$$

Для незалежних випадкових величин теорема множення законів розподілу набуває вигляду:

$$f(x, y) = f_1(x) \cdot f_2(y) \quad (2.24)$$

Ймовірнісна залежність між випадковими величинами може бути більш-менш тісною. Числовою характеристикою тісноти зв'язку в системі двох випадкових величин може бути кореляційний момент, який визначається співвідношенням

$$K_{xy} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - m_x)(y - m_y) f(x, y) dx dy. \quad (2.25)$$

Оскільки для незалежних випадкових величин щільність ймовірності системи випадкових величин дорівнює добутку безумовних ймовірностей кожної з цих величин (2.24), інтеграл (2.25) можна представити як добуток двох інтегралів

$$K_{xy} = \int_{-\infty}^{\infty} (x - m_x) f_1(x) dx \int_{-\infty}^{\infty} (y - m_y) f_2(y) dy, \quad (2.26)$$

де інтеграл

$$\int_{-\infty}^{\infty} (x - m_x) f_1(x) dx \quad (2.27)$$

являє собою не що інше, як перший центральний момент, що дорівнює 0. З тієї же причини дорівнює нулю і другий співмножник, звідки можна зробити висновок, що для незалежних випадкових величин  $K_{xy} = 0$ . Таким чином, якщо кореляційний момент двох випадкових величин відмінний від нуля, то це є ознакою наявності залежності між ними. На практиці зазвичай використовується не сам кореляційний момент, для якого вживають термін *коваріація*  $cov(x, y)$ , а його нормоване (безрозмірне) значення, яке називають коефіцієнтом кореляції

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y}, \quad (2.28)$$

причому

$$-1 \leq r_{xy} \leq 1, \quad (2.29)$$

де  $\sigma_x, \sigma_y$  - середні квадратичні відхилення величин  $X, Y$ .

При  $K_{xy} = 0$   $r_{xy}$  також дорівнює 0, тобто для незалежних випадкових величин коефіцієнт кореляції дорівнює нулю. У той же час, для функціонально зв'язаних величин  $|r_{xy}| = 1.0$ . Випадкові величини, для яких коваріація і кореляційний момент дорівнюють нулю, називаються некорельованими. Однак, поняття некорельованості і незалежності не еквівалентні. З незалежних випадкових величин впливає їхня некорельованість, але з некорельованості не впливає їхня незалежність. Умова незалежності випадкових величин більш жорстка, ніж умова некорельованості, оскільки коефіцієнт кореляції цілком характеризує не всяку залежність, а тільки так звану лінійну. Лінійна стохастична залежність випадкових величин полягає в тім, що при зростанні однієї випадкової величини інша теж має тенденцію зростання чи убуття за лінійним законом

$$M[Y / X = x] = f(x), \quad (2.30)$$

де  $f(x)$  - лінійна функція  $x$ .

Позитивна кореляція між випадковими величинами означає, що при зростанні однієї з них інша має тенденцію в середньому зростати ( $0 < r_{xy} < 1$ ); негативна кореляція означає, що при зростанні однієї з випадкових величин інша має тенденцію в середньому убавати ( $-1 < r_{xy} < 0$ ).

## 2.4 Нормальний закон розподілу для системи двох випадкових величин

Найбільше поширення в практичному використанні має так званий класичний нормальний закон розподілу системи двох нормально-розподілених випадкових величин. У загальному випадку щільність нормального розподілу двох випадкових величин виражається формулою:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r_n^2}} e^{-\frac{1}{2(1-r_n^2)} \left[ \frac{(x-m_x)^2}{\sigma_x^2} - \frac{2r_n(x-m_x)(y-m_y)}{\sigma_x\sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2} \right]}, \quad (2.31)$$

де  $r_n$  - коефіцієнт кореляції між двома випадковими, нормально розподіленими величинами.

Використовуючи вирази (2.16) і (2.17), можна показати, що величини  $Y$  і  $X$  підлягають нормальному закону розподілу з центрами розсіювання  $m_y$  та  $m_x$  і середніми квадратичними відхиленнями  $\sigma_y$  та  $\sigma_x$ . А при підстановці у формулу (2.25) виразу (2.31), після деяких перетворень, отримаємо

$$K_{xy} = r_n\sigma_x\sigma_y \quad (2.32)$$

звідки

$$r_n = \frac{K_{xy}}{\sigma_x\sigma_y}. \quad (2.33)$$

У такий спосіб параметр  $r_n$  у формулі (2.32) є ніщо інше, як коефіцієнт кореляції величин  $X$  і  $Y$ , тобто  $r_n = r_{xy}$ . Припустимо, що випадкові величини  $X$  і  $Y$  підлягають нормальному закону і некорельовані, тобто  $r_n = 0$ . Тоді вираз (2.31) набере вигляду:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-m_x)^2}{2\sigma_x^2} - \frac{(y-m_y)^2}{2\sigma_y^2}}. \quad (2.34)$$

Легко переконатися в тім, що випадкові величини  $X$  і  $Y$ , підпорядковані закону розподілу з щільністю (2.34), не тільки некорельовані, але й незалежні. Дійсно, (2.34) можна представити як добуток щільностей розподілу окремих величин, що входять до системи

$$f(x, y) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} \cdot \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(y-m_y)^2}{2\sigma_y^2}}, \quad (2.35)$$

яка відповідає теоремі множення законів розподілу (2.24), що розглядається як необхідна і достатня умова незалежності випадкових величин.

Таким чином, для системи двох випадкових величин, підпорядкованих нормальному закону, з некорельованості величин випливає їхня незалежність.

Визначаючи умовний закон розподілу величини  $Y$  при  $r_{xy} \neq 0$  відповідно до виразу (2.21), одержимо:

$$f(y/x) = \frac{1}{\sigma_y\sqrt{1-r_{xy}^2}\sqrt{2\pi}} e^{-\frac{1}{2(1-r_{xy}^2)}\left[\frac{y-m_y}{\sigma_y} - r_{xy}\frac{x-m_x}{\sigma_x}\right]^2}. \quad (2.36)$$

Приведемо вираз (2.36) до вигляду:

$$f(y/x) = \frac{1}{\sigma_y\sqrt{1-r_{xy}^2}\sqrt{2\pi}} e^{-\frac{1}{2\sigma_y^2(1-r_{xy}^2)}\left[y-m_y - r_{xy}\frac{\sigma_y}{\sigma_x}(x-m_x)\right]^2}. \quad (2.37)$$

Очевидно, що це є щільність розподілу нормального закону з центром розсіювання

$$m_{y/x} = m_y + r_{xy} \frac{\sigma_y}{\sigma_x} (x - m_x) \quad (2.38)$$

і середнім квадратичним відхиленням

$$\sigma_{y/x} = \sigma_y \sqrt{1 - r_{xy}^2}. \quad (2.39)$$

З формул (2.38) і (2.39) випливає, що умовний закон розподілу величини  $Y$  характеризується тільки зміною математичного сподівання при зміні  $x$ , а умовна дисперсія від  $x$  не залежить. Величина  $m_{y/x}$  називається умовним математичним сподіванням величини  $Y$  при даному  $x$ . Залежність (2.38) можна зобразити на площині  $xOy$ , відкладаючи умовне математичне сподівання по осі ординат. У відповідності з (2.38), вийде пряма лінія, яка називається лінією регресії  $Y$  по  $X$  (рис. 2.7). Рівняння регресії  $X$  по  $Y$  можна представити у вигляді

$$m_{x/y} = m_x + r_{xy} \frac{\sigma_x}{\sigma_y} (y - m_y). \quad (2.40)$$

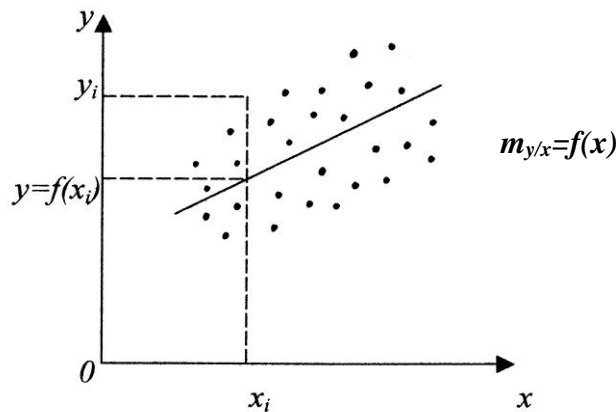


Рисунок 2.7 – Графічний вигляд рівняння лінійної регресії  $m_{y/x} = f(x)$

Рівняння (2.38) і (2.40) можуть бути приведені до вигляду рівнянь прямих ліній

$$m_{y/x} = \alpha_1 x + \beta_1, \quad (2.41)$$

$$m_{x/y} = \alpha_2 y + \beta_2, \quad (2.42)$$

де  $\alpha_1, \beta_1, \alpha_2, \beta_2$  - коефіцієнти рівнянь прямих, причому

$$\alpha_1 = r_{xy} \frac{\sigma_y}{\sigma_x}; \quad \beta_1 = m_y - \alpha_1 m_x; \quad (2.43)$$

$$\alpha_2 = r_{xy} \frac{\sigma_x}{\sigma_y}; \quad \beta_2 = m_x - \alpha_2 m_y. \quad (2.44)$$

Величини  $m_x, m_y$  - це безумовні математичні сподівання випадкових величин  $X$  і  $Y$ , тобто центри їх розподілу без урахування залежності між  $X$  і  $Y$ .

## **2.5 Рівняння парної лінійної регресії, оцінка їх параметрів за вибірковими даними**

Отримані рівняння парної лінійної регресії вигляду (2.43), (2.44) представлені у вигляді рівнянь умовного математичного сподівання для системи випадкових величин, широко застосовуються у гідрологічних розрахунках як математична модель.

Пошук рівнянь парної лінійної регресії може відбуватися для величин стоку з різних водозборів (метод приведення коротких рядів стоку до довгого періоду), для величин стоку і стокоформуєчих факторів (при розробці регіональних, стохастичних залежностей при відсутності даних спостережень), а також при пошуку зв'язків між перерізами процесу стоку.

У останньому випадку як випадкова величина  $x$ , виступає переріз випадкового процесу в момент  $t_j$ , а як величина  $y$  - переріз у момент часу  $t_j + \tau$ , де  $\tau$  - зсув у часі.

Прийнята до розрахунку математична модель (лінійна парна регресія) описує зв'язок генеральних сукупностей залежних випадкових величин  $X$  і  $Y$ . Задача користувача полягає в тому, щоб за обмеженими даними спостережень (вибірками), зробити висновки про характер зв'язку та оцінити статистичні параметри рівняння регресії.

Для вибіркових даних рівняння умовного математичного сподівання представляється у вигляді

$$\tilde{y}_i = \tilde{y}(x_i) = \hat{m}_{y/x} = ax_i + b, \quad (2.45)$$

де  $x_i$  - дискретні значення випадкової величини  $X$ ;  $y$  - дискретні значення випадкової величини  $Y$ ;  $\tilde{y}_i$  - значення випадкової величини  $Y$ , розраховані за рівнянням регресії;  $a, b$  - шукані параметри рівняння.

### 2.5.1 Оцінка параметрів рівняння

Для отримання розрахункових формул визначення параметрів  $a$  і  $b$  використовується метод найменших квадратів. Оскільки йдеться про рівняння умовного математичного сподівання, то можна сказати, що функція регресії  $\tilde{y}(x) = m_{y/x}$  є функцією, яка мінімізує середню квадратичну похибку визначення  $\tilde{y}(x)$  при обчисленні  $\tilde{y}_i$  по величині  $x_i$ . Це означає, що для довільної функції  $U(x)$  справедлива нерівність

$$M[Y - U(X)]^2 \geq M[Y - \tilde{y}(x)]^2 \quad (2.46)$$

Для того, щоб виконувалася нерівність (2.46), оцінки параметрів, які входять до рівняння регресії, обчислюються за методом найменших квадратів. Це метод обробки емпіричного числового матеріалу, вимога якого полягає в тому, щоб сума квадратів відхилень даних спостережень  $y_i$  від лінії регресії була найменшою, тобто

$$\Delta = \sum_{i=1}^n [y_i - \tilde{y}(x_i)]^2 = \min, \quad (2.47)$$

де  $y_i$  - спостережені значення випадкової величини  $Y$ ;  $\tilde{y}(x_i)$  - значення випадкової величини  $Y$ , розраховані по рівняннях регресії для заданих  $x_i$ ;  $n$  - число спільно спостережених значень  $y_i$  і  $x_i$ .

Допустимо, що за математичну модель, яка описує стохастичний кореляційний зв'язок між випадковими величинами  $Y$  і  $X$ , обране рівняння лінійної регресії вигляду (2.41). На основі вихідних даних за спільний період спостережень необхідно оцінити коефіцієнти  $a$  й  $b$  і одержати тим самим можливість розрахунку («передбачення») значень випадкової величини  $Y$  при заданих  $x_i$ . Запис рівняння набирає вигляду

$$\tilde{y}(x_i) = ax_i + b, \quad (2.48)$$

де  $a = \hat{\alpha}_1; b = \hat{\beta}_1$ ; тобто  $a$  і  $b$  є оцінками параметрів моделі (2.41).

Відповідно до методу найменших квадратів  $a$  і  $b$  повинні бути такими, щоб сума  $\Delta$  досягала свого мінімуму. Вимога екстремума означає, що частинні похідні  $\Delta$ , узяті по  $a$  і  $b$ , дорівнюють нулю

$$\frac{\partial \Delta(a,b)}{\partial a} = \frac{\partial \left[ \left( \sum_{i=1}^n y_i - ax_i - b \right)^2 \right]}{\partial a} = 0 ; \quad (2.49)$$

$$\frac{\partial \Delta(a,b)}{\partial b} = \frac{\partial \left[ \left( \sum_{i=1}^n y_i - ax_i - b \right)^2 \right]}{\partial b} = 0 \quad (2.50)$$

Вирішуючи рівняння (2.49) і (2.50) відносно  $a$  і  $b$ , одержимо

$$a = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2}; \quad (2.51)$$

$$b = \bar{y} - a\bar{x} . \quad (2.52)$$

Чисельник дробу, що знаходиться в правій частині рівності (2.51), являє собою оцінку кореляційного моменту  $\hat{K}_{xy}$  (коваріації), розраховану за дискретною вибіркою довжиною  $n$

$$\hat{K}_{xy} = \text{cov}(x, y) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \quad (2.53)$$

а знаменник – оцінку дисперсії випадкової величини  $X$

$$\hat{\sigma}_x^2 = S_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2 . \quad (2.54)$$

Оцінка коефіцієнта кореляції, який відображає тісноту лінійного зв'язку відповідно до виразу (2.53), може бути представленою у вигляді



$$\hat{r}_{xy} = r = \frac{cov(x, y)}{S_x S_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}}, \quad (2.55)$$

де  $S_y$  - оцінка середнього квадратичного відхилення випадкової величини  $Y$ .

Оцінка дисперсії  $S_y^2$  випадкової величини  $Y$  виконується за формулою

$$\hat{\sigma}_y^2 = S_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \overline{y^2} - \bar{y}^2. \quad (2.56)$$

Оцінка параметра рівняння лінійної регресії  $a$  зазвичай виражається через коефіцієнт кореляції  $r$

$$a = \hat{r}_{x,y} \frac{S_y}{S_x}. \quad (2.57)$$

### 2.5.2 Оцінка вірогідності лінійної регресійної моделі

Оцінка вірогідності моделі передбачає установлення ступеня відповідності між реальним стохастичним зв'язком випадкових величин  $X$  і  $Y$  та математичною моделлю, яка описує цей зв'язок. Основою такої оцінки може служити величина  $\Delta$  (2.47), яка являє собою суму квадратів нев'язок  $\varepsilon_i$  між спостереженими і розрахованими по рівнянню лінійної регресії значеннями випадкової величини

$$\Delta = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [y_i - (ax_i + b)]^2. \quad (2.58)$$

Критерієм якості розрахунку є середнє квадратичне відхилення спостережених  $y_i$  і розрахованих  $\tilde{y}_i = ax_i + b$  значень.

$$S = \sqrt{\frac{\sum_{i=1}^n \varepsilon_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^n [y_i - (ax_i + b)]^2}{n}}. \quad (2.59)$$

Величина  $S$  носить також назву похибки апроксимації.

Рівняння (2.59) легко перетвориться до вигляду

$$S^2 = \overline{y^2} - \bar{y}^2 - 2a \cdot \text{cov}(x, y) + a^2 [\overline{x^2} - \bar{x}^2]. \quad (2.60)$$

З урахуванням (2.55)-(2.57) похибка апроксимації  $S$  представляється у вигляді

$$S^2 = S_y^2 (1 - r^2) \quad (2.61)$$

чи

$$S = S_y \sqrt{1 - r^2}. \quad (2.62)$$

Однак середня квадратична похибка апроксимації  $S$  не є досить інформативною оцінкою, оскільки нев'язка розрахунку визначається в першу чергу вірогідністю визначення коефіцієнтів регресії. Щоб переконатися у вірогідності побудованої моделі, необхідно оцінити вірогідність коефіцієнтів  $a$  і  $b$ , для чого будуються довірчі інтервали.

### 2.5.3 Визначення довірчих інтервалів для коефіцієнтів рівняння лінійної регресії

Обчислюючи на підставі наявних вибірових даних оцінку  $\hat{\theta}$  деякого параметра  $\theta$ , ми усвідомлюємо, що насправді величина  $\hat{\theta}$  є лише наближеним значенням невідомого параметра  $\theta$ , навіть у тому випадку, коли ця оцінка обґрунтована, незміщена й ефективна. Виникає питання, як сильно може відхилитися це наближене вибірове значення (оцінка) від відповідного значення генеральної сукупності? Рішення цього питання досягається шляхом установаження довірчих інтервалів.

*Довірчим* називається інтервал числової осі, відносно якого з імовірністю, близькою до 1, можна стверджувати, що оцінюваний параметр  $\hat{\theta}$  знаходиться у середині нього

$$p[\theta_1 < \hat{\theta} < \theta_2] = 1 - q, \quad (2.63)$$

де  $q$  - доповнення до одиниці довірчої ймовірності, який звичайно приймається рівним 0.05 чи 0.01 і називається рівнем значущості. Чим вужчий інтервал, тим вище вірогідність оцінки, тим точніше модель описує реальність. Якщо розподіл оцінки параметра близький до нормального, довірчий інтервал має вигляд

$$[\hat{\theta} - t_{\nu, \alpha} \sigma_{\hat{\theta}}; \hat{\theta} + t_{\nu, q} \sigma_{\hat{\theta}}], \quad (2.64)$$

де  $t_{\nu, \alpha}$  - статистика Стьюдента, яка залежить від рівня значущості  $\alpha$  й обсягу вибірки  $\nu = n - 2$ ;  $\sigma_{\hat{\theta}}$  - середнє квадратичне відхилення вибіркової оцінки  $\hat{\theta}$ .

Так, наприклад, довірчий інтервал для математичного сподівання визначається співвідношенням

$$\bar{x} - t_{\nu, \alpha} \sigma_{\bar{x}} < m_x < \bar{x} + t_{\nu, \alpha} \sigma_{\bar{x}}, \quad (2.65)$$

де

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}. \quad (2.66)$$

Якщо величина  $Y$  розподілена нормально, то не тільки оцінки середнього арифметичного цієї величини, але й оцінки коефіцієнтів регресії розподілені нормально. Щоб показати це, досить представити коефіцієнт  $a$  (2.51) у вигляді лінійної комбінації спостережених значень  $y_i$

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \sum_{i=1}^n (x_i)^2} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{n S_x^2} \quad (2.67)$$

чи

$$a = \sum_{i=1}^n K_i y_i, \quad (2.68)$$

де

$$K_i = \frac{x_i - \bar{x}}{nS_x^2}. \quad (2.69)$$

Оскільки комбінація нормально розподілених величин також може бути описана нормальним законом розподілу, коефіцієнти рівняння регресії  $a$  і  $b$  варто визнати нормально розподіленими. Отже, довірчі інтервали коефіцієнтів регресії можуть бути побудовані в такий спосіб

$$a - t_{v,q}\sigma_a < m_a < a + t_{v,q}\sigma_a; \quad (2.70)$$

$$b - t_{v,q}\sigma_b < m_b < b + t_{v,q}\sigma_b. \quad (2.71)$$

У цьому випадку проблема зводиться до пошуку стандартів  $\sigma_a$  і  $\sigma_b$ .

Розглянемо, чому дорівнює математичне сподівання коефіцієнта регресії  $a$ , розраховане за вибіркою, довжиною  $n$ , відповідно до рекомендацій методу найменших квадратів та за умови, що вираз (2.67) представлений як лінійна комбінація спостережених значень  $y_i$  (2.68):

$$m_a = M[a] = M\left[\sum_{i=1}^n K_i y_i\right] = \sum_{i=1}^n M[K_i y_i] = \sum_{i=1}^n K_i M[y_i], \quad (2.72)$$

де  $M[y_i]$  - математичне сподівання значень випадкової величини  $Y$  при заданому  $x_i$ , тобто умовне математичне сподівання  $m_{y/x_i}$ .

Тоді

$$m_a = \sum_{i=1}^n K_i m_{y/x_i} = \sum_{i=1}^n \frac{x_i - \bar{x}}{nS_x^2} (\alpha_1 x_i + \beta_1) \quad (2.73)$$

чи

$$m_a = \frac{\alpha_1 \left[ \overline{x^2} - \bar{x}^2 \right]}{S_x^2} = \alpha_1. \quad (2.74)$$

Іншими словами,  $a_1$  є незміщеною функцією параметра  $\alpha$  в рівнянні лінійної регресії для системи нормально розподілених випадкових величин  $X$  і  $Y$ .

Визначимо дисперсію коефіцієнта регресії  $a$  :

$$\sigma_a^2 = D[a] = D\left[\sum_{i=1}^n K_i y_i\right] = \sum_{i=1}^n D[K_i y_i] = \sum_{i=1}^n K_i D[y_i], \quad (2.75)$$

де  $D[y_i]$  - умовна дисперсія випадкової величини  $Y$ , яка входить до системи нормально розподілених і лінійно залежних величин  $X$  і  $Y$

$$\sigma_{y/x_i}^2 = D[y_i] = M\{[y_i - \alpha_1 x_i - \beta_1]^2\}. \quad (2.76)$$

Не важко помітити, що величина  $S$ , яка визначається за виразом (2.59), є по суті оцінкою умовного середнього квадратичного відхилення  $\sigma_{y/x_i}^2$ . Оскільки для системи нормально розподілених величин умовний стандарт  $\sigma_{y/x_i}$  не залежить від значення  $x_i$ , можна прийняти, що  $\sigma_{y/x_i}^2 \cong S^2$  при усіх  $x_i$ .

Тоді

$$\sigma_a^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n^2 S_x^4} \cdot \sigma_{y/x_i}^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n^2 S_x^4} \cdot S^2. \quad (2.77)$$

Підставляючи в (2.77) вираз (2.62), одержимо

$$\sigma_a^2 = \frac{S_y^2}{S_x^2 n} (1 - \bar{r}_{xy}^2) \quad (2.78)$$

і

$$\sigma_a = \frac{S_y}{S_x} \sqrt{\frac{1 - \bar{r}_{xy}^2}{n}}. \quad (2.79)$$

Зробивши виклади, аналогічні тим, що зроблені для коефіцієнта  $a$ , стосовно до  $b$ , одержимо

$$m_b = \beta_1;$$

$$\sigma_b = \sqrt{S_x^2 + (\bar{x})^2} \cdot \frac{S_y}{S_x} \sqrt{\frac{1 - \hat{r}_{x,y}^2}{n}}. \quad (2.80)$$

#### 2.5.4 Визначення довірчого інтервалу для коефіцієнта кореляції в рівнянні лінійної регресії

Емпіричний (вибірковий) коефіцієнт кореляції  $\hat{r}$  являє собою обґрунтовану оцінку. При досить великому обсязі вибірки ( $n > 30$ ) і невисокому коефіцієнті кореляції ( $r \leq 0.3$ ) розподіл його оцінок може вважатись нормальним. Середня квадратична похибка визначення  $r_{xy}$  по спостережених даних у цьому випадку визначається за формулою

$$\hat{\sigma}_r = S_r = \frac{1 - \hat{r}_{x,y}^2}{\sqrt{n-1}}. \quad (2.81)$$

Довірчий інтервал коефіцієнта кореляції дорівнює

$$\hat{r}_{x,y} - t_{v,q} \frac{1 - r^2}{\sqrt{n-1}} < \hat{r}_{xy} < \hat{r}_{x,y} + t_{v,q} \frac{1 - \hat{r}_{xy}^2}{\sqrt{n-1}}. \quad (2.82)$$

Якщо ж довжина ряду невелика ( $n < 30 \dots 40$ ), а коефіцієнт кореляції значний ( $r \geq 0.3$ ), то розподіл вибіркових значень  $r$  істотно відрізняється від нормального. При значеннях  $r$ , що наближаються до 1, крива розподілу вибіркових оцінок коефіцієнта кореляції стає усе більш асиметричною. Тому, для оцінки точності розрахунку значення коефіцієнта кореляції використовують  $z$ -перетворення Фішера

$$z = \frac{1}{2} \ln \left[ \frac{1 + \hat{r}_{xy}}{1 - \hat{r}_{xy}} \right]. \quad (2.83)$$

Закон розподілу оцінок  $z$  близький до нормального із середнім значенням

$$\bar{z} = \ln \left[ \frac{1 + \hat{r}_{x,y}}{2(1-r)} + \frac{\hat{r}_{x,y}}{2(n-1)} \right] \quad (2.84)$$

і дисперсією

$$\sigma_z^2 = 1/(n-3). \quad (2.85)$$

Розрахунки довірчого інтервалу коефіцієнта кореляції виконуються наступним чином:

1. Надається оцінка статистики  $z$

$$\hat{z} = \frac{1}{2} \ln \frac{1 + \hat{r}_{x,y}}{1 - \hat{r}_{x,y}} \quad (2.86)$$

2. Будується довірчий інтервал

$$\hat{z} - t_{v,q} \sigma_z < z < \hat{z} + t_{v,q} \sigma_z, \quad (2.87)$$

тобто

$$z_1 < z < z_2$$

3. Відбувається зворотний перехід від  $z$  до  $r_{xy}$

$$2\hat{z} = \ln \frac{1 + \hat{r}_{x,y}}{1 - \hat{r}_{x,y}}; \quad \frac{1 + \hat{r}_{x,y}}{1 - \hat{r}_{x,y}} = \exp\{2\hat{z}\} \quad (2.88)$$

$$1 + \hat{r}_{xy} = \exp 2\hat{z} - \hat{r}_{xy} \cdot \exp 2\hat{z} \quad (2.89)$$

або

$$\hat{r}_{x,y} (1 + \exp 2\hat{z}) = \exp 2\hat{z} - 1, \quad (2.90)$$

звідки

$$\hat{r}_{x,y} = \frac{\exp 2\hat{z} - 1}{\exp 2\hat{z} + 1}. \quad (2.91)$$

4. Використовуючи границі довірчих інтервалів для статистики  $z$ , за допомогою рівняння (2.87), одержуємо границі довірчого інтервалу коефіцієнта кореляції

$$r_1 = \frac{\exp 2\hat{z}_1 - 1}{\exp 2\hat{z}_1 + 1} \quad (2.92)$$

i

$$r_2 = \frac{\exp 2\widehat{z}_2 - 1}{\exp 2\widehat{z}_2 + 1}; \quad (2.93)$$

тоді

$$r_1 < \widehat{r}_{x,y} < r_2. \quad (2.94)$$

### 2.5.5 Перевірка гіпотези про статистичну значущість коефіцієнта кореляції і коефіцієнтів рівняння регресії

Часто в гідрологічних дослідженнях постає питання про реальність установлених на основі спостережених даних зв'язків, тому що можливо, що встановлені значення  $\widehat{r}_{x,y}$ ,  $a$  й  $b$  обумовлені випадковістю вибірок. Вирішення такого питання називають оцінкою статистичної значущості параметрів і воно зв'язане з перевіркою статистичних гіпотез.

Висунемо нульову гіпотезу щодо тісноти розглянутого зв'язку

$$H_0 : r_{xy} = 0. \quad (2.95)$$

Якщо розподіл вибірових оцінок  $r_{xy}$  задовольняє нормальний закон розподілу (що вірно при великих  $n$ ), то для перевірки нульової гіпотези як критерій можна використовувати статистику

$$t = \frac{r_{xy}}{\sqrt{\sigma_{r_{xy}}^2}}, \quad (2.96)$$

розподіл якої підлягає розподілу Стюдента.

Значення  $t$  визначається за вибіровими оцінками коефіцієнта кореляції  $\widehat{r}_{x,y}$  і його стандарту  $S_r$

$$\widehat{t} = \frac{|\widehat{r}_{xy}|}{S_r} \quad (2.97)$$

і порівнюється з критичним  $t_{кр}$ , котре залежить від числа степенів вільності  $\nu = n - 1$  і рівня значущості  $q$ .



При  $t < t_{кр}$  нульова гіпотеза приймається, а при  $t > t_{кр}$  - відхиляється, тобто значення коефіцієнта кореляції визнається статистично значущим.

При невеликих  $n$  і високих значеннях  $\hat{r}_{x,y}$  оцінка статистичної значущості  $\hat{r}_{x,y}$  виконується за допомогою  $z$ -перетворення Фішера, тобто оцінюється не величина  $\hat{r}_{x,y}$  безпосередньо, а статистика  $\hat{z}$ . Якщо  $z$  значуще, то і коефіцієнт кореляції є значимою величиною.

Оцінка значущості коефіцієнтів рівняння регресії  $a$  і  $b$  виконується аналогічним чином.

Нульова гіпотеза полягає в тому, що коефіцієнти регресії визнаються незначущими, якщо

$$H_0 : \alpha = 0 \quad (2.98)$$

$$H_0 : \beta = 0. \quad (2.99)$$

Як вже відзначалося в наведених вище розділах, розподіл вибірових оцінок  $a$  і  $b$  допускається нормальним. Тоді у відповідності до теорем математичної статистики для перевірки гіпотези про  $a$  і  $b$  можна використовувати статистику  $t$ , яка підлягає розподілу Стьюдента

$$\hat{t}_a = \frac{|a|}{S_a} \quad \text{и} \quad \hat{t}_b = \frac{|b|}{S_b} \quad (2.100)$$

Якщо  $\hat{t} > t_{кр}(v, q)$ , то коефіцієнти регресії вважаються значущими.

## РОЗДІЛ 3

### ЗАГАЛЬНІ УЯВЛЕННЯ ПРО ТЕОРІЮ ВИПАДКОВИХ ПРОЦЕСІВ

#### 3.1 Поняття про випадкову функцію

На практиці доводиться оперувати величинами, які змінюються у процесі випробувань. Такі величини отримали назву випадкових функцій. Вивченням подібних випадкових явищ, в яких випадковість набуває форми процесу, займається спеціалізований розділ теорії ймовірностей – теорія випадкових функцій або стохастичних процесів.

**Випадкова функція** – це функція, значення якої встановлюються за допомогою випробувань і можуть бути різними в залежності від ходу випробувань. **Випадковою функцією називається функція, яка в разі випробувань може набрати того чи іншого конкретного вигляду, наперед невідомо, якого саме. Конкретний вигляд, якого набуває випадкова функція в результаті випробувань, називається реалізацією випадкової функції.**

Зустрічаються випадкові функції, які залежать не від одного аргументу, а від декількох. У гідрологічних розрахунках найчастіше розглядаються функції тільки одного аргументу. Якщо аргументом випадкової функції є час, то для її позначення використовується термін «випадковий процес».

Випадкову функцію вивчають аналітичним способом, заснованим на визначенні багатовимірного закону її розподілу, і статистичним, заснованим на визначенні тільки числових характеристик такої функції. Перший спосіб застосовують у теоретичних дослідженнях. Другий спосіб широко використовують для вирішення різних прикладних задач у теорії випадкових функцій.

Розглянемо випадкову функцію  $X(t)$ , над якою проведено  $n$  незалежних випробувань і отримано  $n$  реалізацій. Кожна реалізація є звичайною, тобто не випадковою функцією.

Сімейство реалізацій функції, що відповідає ймовірнісному процесу з безперервною зміною функції в залежності від аргументу, наведене на рис.3.1.

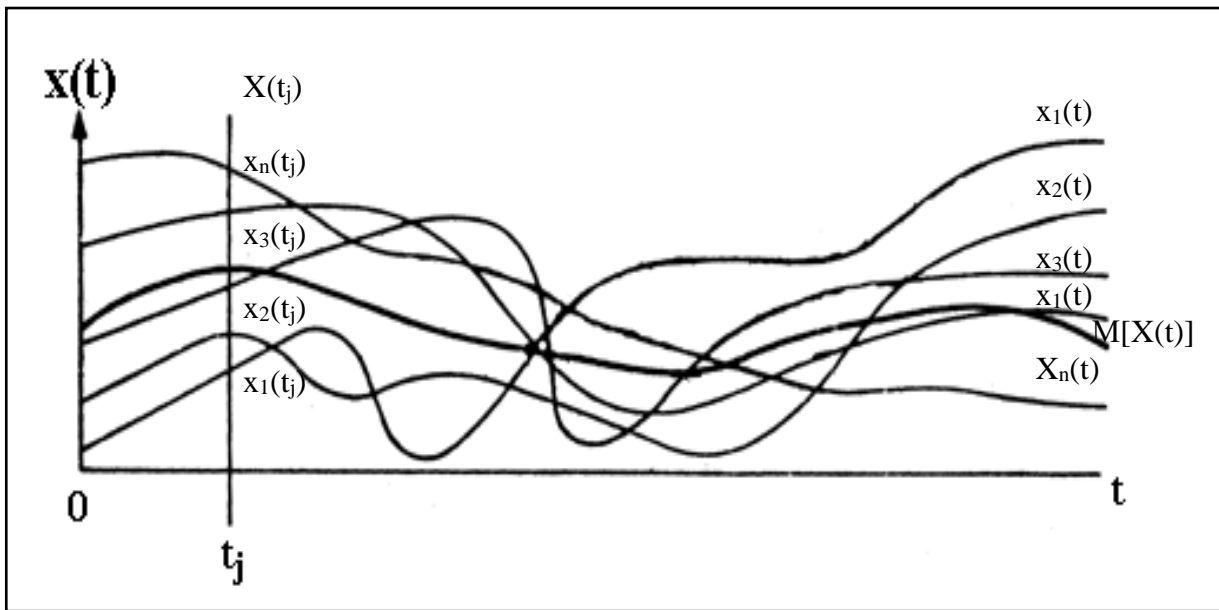


Рисунок 3.1 - Сімейство реалізацій випадкової функції

Реалізації випадкової функції можна отримати за допомогою самописного приладу, осцилограми запису змін процесу в часі, кінозйомки, періодичної фотозйомки та інш.

### 3.2 Закон розподілу випадкової функції

Закон розподілу однієї випадкової величини є функція одного аргументу, закон розподілу системи двох величин – функція двох аргументів і т.п. Однак, для практичного використання функції розподілу багатьох аргументів застосовувати незручно. Зазвичай відмовляються від використання законів розподілу і розглядають тільки числові характеристики випадкових функцій.

Закон розподілу випадкової функції може представляти собою функцію безлічі аргументів, яку чисто формально можна записати в символічній формі.

*Якщо зафіксувати деяке значення аргументу  $t_j$ , то при заданому значенні аргументу  $t_j$  (рис.3.1) випадкова функція перетворюється у випадкову величину  $X(t_j)$ , яка називається перетином випадкової функції, що відповідає моменту  $t_j$ . Значеннями цієї випадкової величини будуть значення  $x_1(t_j), x_2(t_j), \dots, x_n(t_j)$ , де  $n$  - кількість реалізацій.*

Таким чином, випадкова функція суміщує у собі риси випадкової величини і функції. Якщо зафіксувати значення аргументу, то отримуємо випадкову величину. В результаті кожного випробування випадкова функція, у свою чергу, перетворюється у звичайну не випадкову функцію.

Розглянемо випадкову величину  $X(t)$  - перетин випадкової функції в момент часу  $t$ . Ця випадкова величина має закон розподілу, що у загальному випадку залежить від  $t$  і позначається як  $f(x, t)$ . Функція  $f(x, t)$  називається одновимірним законом розподілу випадкової функції  $X(t)$ . Функція  $f(x, t)$  не є повною вичерпною характеристикою випадкової функції  $X(t)$ . Дійсно, ця функція характеризує тільки закон розподілу  $X(t)$  для даного, хоча й довільного  $t$ ; вона не відповідає на запитання про взаємозв'язок випадкових величин  $X(t)$  при різних  $t$ . З цього погляду більш повною характеристикою випадкової функції  $X(t)$  є так званий двовимірний закон розподілу

$$f(x_1, x_2; t_1, t_2). \quad (3.1)$$

Це закон розподілу системи двох випадкових величин  $X(t_1), X(t_2)$ , тобто закон розподілу для двох довільно узятих перетинів випадкової функції  $X(t)$ . Однак і ця характеристика не є вичерпною, ще більш повним описом випадкової функції був би тривимірний закон:

$$f(x_1, x_2, x_3; t_1, t_2, t_3). \quad (3.2)$$

Теоретично можна необмежено збільшувати число аргументів, але оперувати ними вкрай незручно. У зв'язку з цим на практиці розглядають не закони розподілу, а найпростіші характеристики випадкових функцій, аналогічні числовим характеристикам випадкових величин.

### 3.3 Характеристики випадкових функцій та їх визначення

Основними числовими характеристиками випадкового процесу є: математичне сподівання, що визначає таку функцію, навколо якої групуються реалізації випадкової функції; дисперсія, що показує ступінь розсіювання цих реалізацій відносно математичного сподівання, кореляційна функція, що визначає внутрішню структуру випадкового процесу і характеризує силу взаємного зв'язку між реалізаціями випадкової функції.

На відміну від числових характеристик випадкових величин, які представляють собою числа, характеристики випадкових функцій представляють собою у загальному випадку не числа, а функції.

Математичне сподівання випадкової величини визначається в такий спосіб. Розглянемо перетин випадкової функції  $X(t)$  при фіксованому  $t$ . У цьому перетині ми маємо звичайну випадкову величину; визначимо її математичне сподівання. Очевидно, у загальному випадку воно залежить від  $t$ , тобто являє собою деяку функцію  $t$ :

$$m_x(t) = M[X(t)]. \quad (3.3)$$

Таким чином, математичним сподіванням випадкової функції  $X(t)$  називається не випадкова функція  $m_x(t)$ , яка при кожному значенні аргументу  $t$  дорівнює математичному сподіванню відповідного перетину випадкової функції. На рис.3.1 більш тонкими лініями показані реалізації випадкової функції, жирною лінією – її математичне сподівання.

Дисперсією випадкової функції  $X(t)$  називається не випадкова функція  $D_x(t)$ , значення якої для кожного  $t$  дорівнюють дисперсії відповідного перетину випадкової функції

$$D_x(t) = D[X(t)]. \quad (3.4)$$

Дисперсія  $D_x(t)$  є невід'ємною функцією. Добуваючи з неї квадратний корінь, одержимо функцію  $\sigma_x(t)$  - середнє квадратичне відхилення випадкової функції

$$\sigma_x(t) = \sigma[X(t)] = \sqrt{D[X(t)]}. \quad (3.5)$$

Досить часто користуються центрованою характеристикою випадкової функції

$$\delta[X(t)] = X(t) - M[X(t)] \quad (3.6)$$

Очевидно, що

$$M\delta[X(t)] = 0. \quad (3.7)$$

Нормована випадкова функція представляється у вигляді

$$X_0(t) = \delta[X(t)] / \sigma[X(t)], \quad (3.8)$$

для якої справедливим рівнянням

$$M[X_0(t)] = 0; \quad D[X_0(t)] = 1. \quad (3.9)$$

Дисперсія характеризує розсіювання реалізацій випадкової функції відносно функції математичного сподівання. Не виключена можливість однакої дисперсії для всіх значень аргументу  $t$ , тобто  $D[X(t)] = D(X)$ . Якщо така дисперсія дорівнює нулю, то можна вважати, що випадкова функція  $X(t) = M[X(t)]$  має ймовірність появи, яка дорівнює одиниці.

Бувають випадки, коли математичні сподівання випадкових функцій однакові, а їхні дисперсії різні. Наприклад, на рис. 3.2 показані реалізації випадкових функцій  $X(t), Y(t), Z(t)$ , що мають однакові математичні сподівання, але різні дисперсії.

Математичне сподівання і дисперсія визначають тільки смугу можливих реалізацій випадкової функції, але не поведінку реалізацій усередині такої смуги.

□

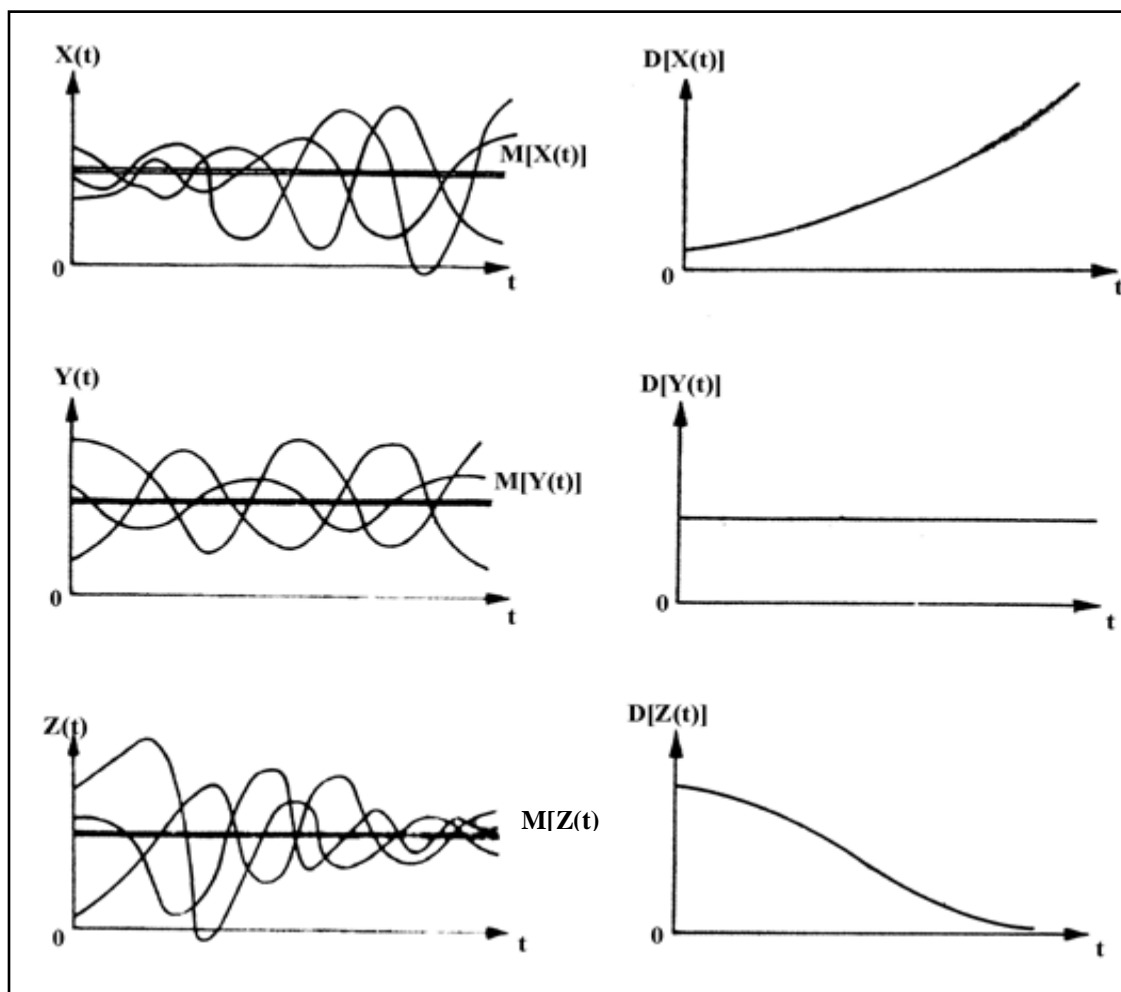


Рисунок 3.2 - Зміна дисперсії у часі для різних випадкових процесів

На рис. 3.3 наведені сімейства реалізацій (а) випадкових функцій, скедастичні криві (б), що показують зміну дисперсій, і кореляційні еліпси (в), які характеризують взаємний зв'язок випадкових ординат по перетинах  $t_1$  і  $t_2$ . Зв'язок між випадковими величинами  $X(t_i)$  і  $X(t_j)$  може бути виражений за допомогою кореляційного моменту чи коефіцієнта кореляції. Ступінь залежності між перетинами випадкової функції у різних аргументах  $t$ , характеризується коваріаційною або кореляційною функціями.

Коваріаційним моментом випадкової функції називається не випадкова функція двох аргументів  $K_x$ , яка для кожної пари перетинів випадкової функції відповідних аргументів  $t_i$  і  $t_j$  дорівнює кореляційному моменту

$$K_x = K[X(t_i), X(t_j)] = M\{\delta[X(t_i)] \cdot \delta[X(t_j)]\}. \quad (3.10)$$

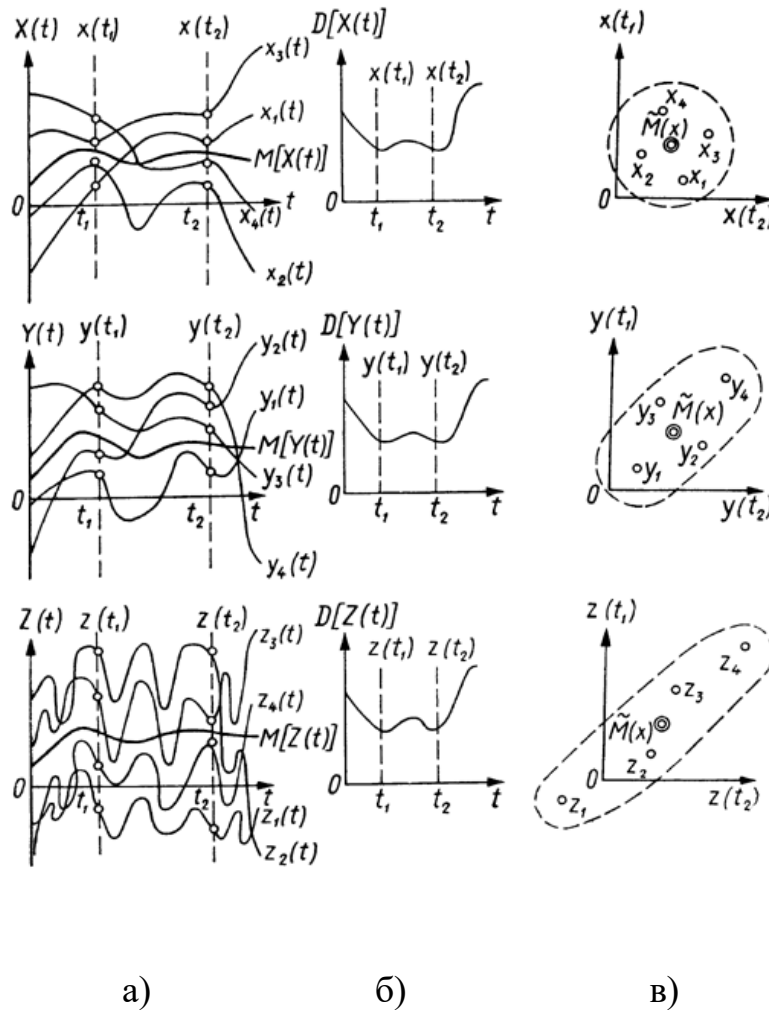


Рисунок 3.3 - Випадкові функції, їх дисперсії та еліпси розсіювання

Розглянемо два випадкових процеси  $X_1(t)$  і  $X_2(t)$  з однаковими математичними сподіваннями та дисперсіями (рис. 3.4, 3.5).

Коваріаційна функція випадкової функції  $X_1(t)$  повільно убуває в міру збільшення проміжку  $t = t'$ . Навпаки, коваріаційна функція випадкової функції  $X_2(t)$  убуває швидко. Ці особливості випадкових процесів характеризуються саме коваріаційними моментами.

На практиці також використовують кореляційну функцію, яка являє собою нормований кореляційний момент

$$r_x = r[X(t_i), X(t_j)] = \frac{K[X(t_i), X(t_j)]}{\sigma(t_i) \cdot \sigma(t_j)}. \quad (3.11)$$

Кореляційну функцію випадкового процесу часто називають **автокореляційною**, що означає кореляцію процесу із самим собою.

При  $t_i = t_j = t$ , одержимо

$$K[X(t), X(t)] = M[\delta^2(t)] = D_x(t), \quad (3.12)$$

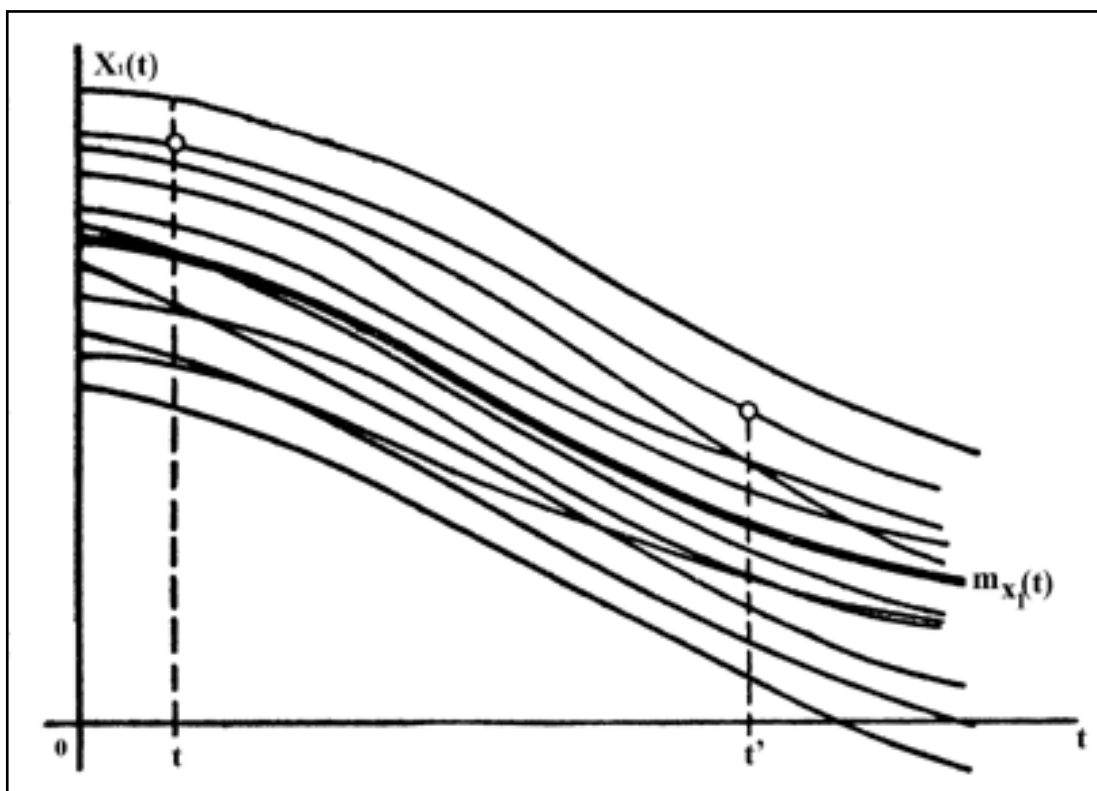


Рисунок 3.4 - Реалізації випадкової функції  $X_1(t)$



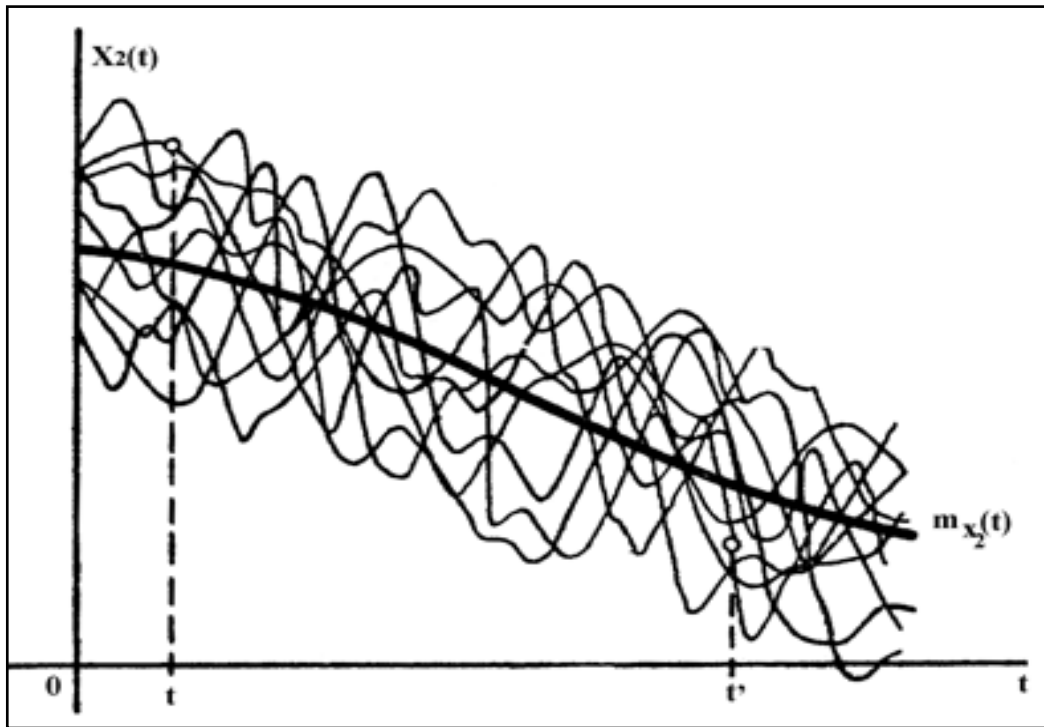


Рисунок 3.5 - Реалізації випадкової функції  $X_2(t)$

тобто при  $t_i = t_j$  коваріаційна функція стає дисперсією випадкової функції, а кореляційна функція дорівнює 1, тобто

$$r_x = 1.0 \quad (3.13)$$

Таким чином, при  $t_i = t_j$  необхідність в дисперсії, як в окремій характеристиці випадкової функції, відпадає: як основною характеристикою випадкової функції достатньо розглядати її математичне сподівання і кореляційну функцію.

Коваріаційний момент двох випадкових величин  $X(t)$  і  $X(t')$  не залежить від послідовності, при якій ці величини розглядаються, отже, коваріаційна функція симетрична відносно своїх аргументів, тобто не змінюється при переміні аргументів місцями

$$K_x(t, t') = K_x(t', t). \quad (3.14)$$

Інтервал  $(t_i - t_j)$  називають «запізнюванням» чи «зсувом».

Очевидно, що значення коваріаційної і автокореляційної функцій залежать від інтервалу між  $t_i$  і  $t_j$ . Вони зменшуються із збільшенням цього інтервалу. При одних і тих же математичних сподіваннях і

дисперсіях дві випадкові функції можуть мати різні автокореляційні функції. Кореляційна функція може бути залежною не від окремих значень  $t_i$  і  $t_j$ , а від різностей  $t_i - t_j$ . Якщо при одному із значень  $t_i$  чи  $t_j$  випадкова функція  $X(t)$  стає не випадковою величиною, то  $K[X(t_i), X(t_j)] = 0$  при будь-якому значенні іншого аргументу.

Коваріаційна (кореляційна) функція симетрична щодо своїх аргументів. Вона зображується в тривимірній системі прямокутних координат  $t_i$ ,  $t_j$  і  $K[X(t_i), X(t_j)]$  у вигляді поверхні, симетричної щодо вертикальної площини  $Q$ , яка проходить через бісектрису кута  $t_i O t_j$  (рис.3.6). Ця властивість коваріаційної функції відповідає властивості коваріаційної матриці системи випадкових величин (табл.3.1), які приблизно виражають випадкову функцію. По головній діагоналі цієї матриці розташовуються дисперсії. Симетрично щодо неї розміщуються кореляційні моменти (коваріації), які задовольняють умову  $K[X(t_i), X(t_j)] = K[X(t_j), X(t_i)]$ . У прикладному варіанті теорії випадкових функцій оперують числовими характеристиками, визначеними за спостереженими даними (з випробувань).

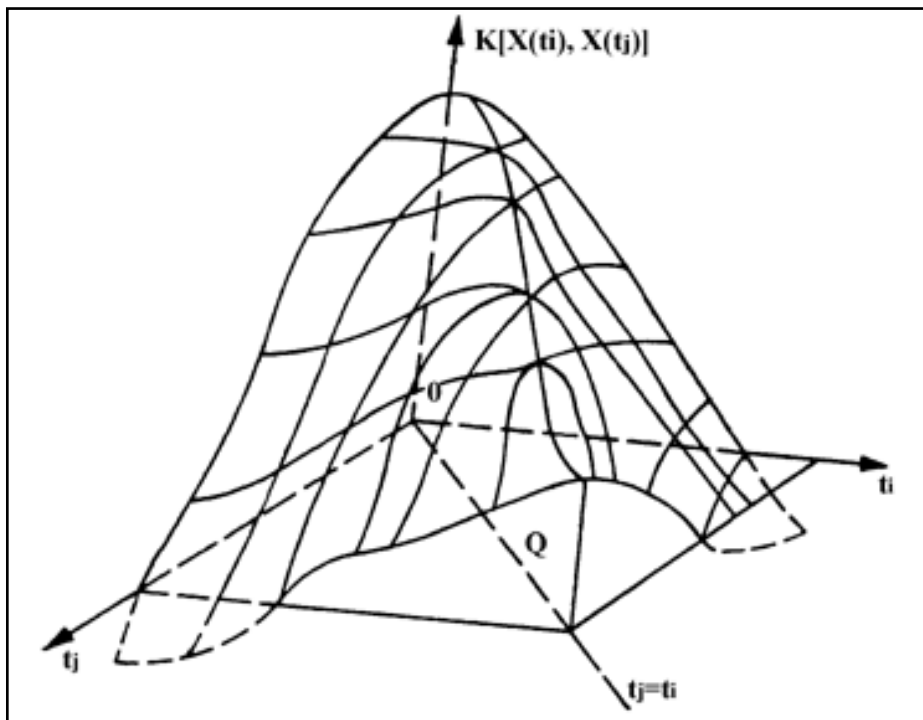


Рис. 3.6 - Коваріаційна функція у вигляді поверхні

Таблиця 3.1 – Коваріаційна матриця системи випадкових величин

$t_i$	$t_j$					
	$t_1$	$t_2$	...	$t_j$	...	$t_n$
$t_1$	$D_1$	$K_{12}$	...	$K_{1j}$	...	$K_{1n}$
$t_2$	$K_{21}$	$D_2$	...	$K_{2j}$	...	$K_{2n}$
...	...	...	...	...	...	...
$t_i$	$K_{i1}$	$K_{i2}$	...	$D_j$	...	$K_{in}$
...	...	...	...	...	...	...
$t_n$	$K_{n1}$	$K_{n2}$	...	$K_{nj}$	...	$D_n$

Нехай над випадковою функцією  $X(t)$  зроблено  $n$  незалежних випробувань (спостережень), у результаті отримано  $n$  реалізацій випадкової функції. Якщо число реалізацій дорівнює  $m$ , то для кожного перетину  $t_i$  можна обчислити емпіричні (статистичні) оцінки характеристик випадкової функції

Якщо число реалізацій дорівнює  $m$ , то для кожного перетину  $t_i$  можна обчислити емпіричні (статистичні) оцінки характеристик випадкової функції  $X(t)$ . Так, оцінкою математичного сподівання в перетині  $t_i$  є просте середнє арифметичне

$$\bar{x}(t_i) = \frac{\sum_{j=1}^m x(t_i)}{m}, \quad (3.15)$$

оцінкою дисперсії -

$$\hat{\sigma}_x(t_i) = S_X^2(t_i) = \frac{1}{m} \sum_{j=1}^m \{ \delta [X(t_i)] \} = \frac{\sum_{j=1}^m [x(t_i) - \bar{x}(t_i)]^2}{m-1} \quad (3.16)$$

Для кореляційних і коваріаційних функцій розрахунок виконується по наступних формулах

$$\hat{r}_x(t_i, t_k) = \frac{\sum_{j=1}^m [x(t_i) - \bar{x}(t_i)] \cdot [x(t_k) - \bar{x}(t_k)]}{(m-1)S_x(t_i)S_x(t_k)} ; \quad (3.17)$$

$$\hat{K}_x(t_i, t_k) = \frac{\sum_{j=1}^m [x(t_j) - \bar{x}(t_j)] \cdot [x(t_k) - \bar{x}(t_k)]}{(m-1)}. \quad (3.18)$$

Після того, як характеристики обчислені, можна побудувати залежності середньої арифметичних  $\bar{x}$  і дисперсії  $S_x^2$  від часу. Функції двох аргументів  $\hat{r}_x(t_i, t_k)$  і  $\hat{K}_x(t_i, t_k)$  також відтворюються в прямокутній сітці точок. За необхідністю всі ці функції апроксимуються аналітичними виразами.

### 3.4 Поняття про стаціонарність випадкового процесу. Ергодичність стаціонарних випадкових процесів

*Випадкові процеси, що протікають у часі приблизно однорідно, називають стаціонарними.* Вони мають вигляд безперервних коливань відносно деякого середнього значення. Реалізації таких процесів знаходяться ніби-то у стані статистичної рівноваги. З часом середня амплітуда і характер коливань істотно не змінюються.

У строгому розумінні слова випадковий процес називають стаціонарним, якщо всі його закони розподілу не змінюються при додаванні до всіх значень аргументу одного й того ж числа, тобто якщо усі вони залежать тільки від взаємного розташування значень аргументу одного й того ж числа або якщо усі вони залежать тільки від взаємного розташування значень аргументу, але не від самих цих значень.

Таким чином, випадковий процес  $X(t)$  є стаціонарним, якщо при будь-якому  $n$  і будь-якому  $t_0$  будуть виконуватися рівності

$$f_n(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_n(x_1, x_2, \dots, x_n; t_1 + t_0, t_2 + t_0, \dots, t_n + t_0) \quad (3.19)$$

Отже, щільності розподілу інваріантні стосовно зсуву початку відліку аргументу  $t$ .

Зокрема, для одновимірної щільності розподілу  $f_1(x, t)$  стаціонарного випадкового процесу, беручи  $t_0 = -t$ , отримуємо

$$f_1(x, t) = f_1(x; t - t) = f_1(x; 0) = f_1(x). \quad (3.20)$$

Таким чином, одновимірна щільність розподілу стаціонарного процесу не залежить від  $t$ , вона є однією і тією ж для усіх перетинів випадкового процесу. Двовимірна щільність розподілу при  $t_0 = -t_1$  представляється у вигляді

$$f_2(x_1, x_2; t_1, t_2) = f_2(x_1, x_2; 0; t_2 - t_1) = f_2(x_1, x_2; t_2 - t_1) = f_2(x_1, x_2; \tau), \quad (3.21)$$

тобто двовимірна щільність розподілу залежить не від двох аргументів  $t_1, t_2$ , а тільки від одного аргументу – їхньої різниці  $\tau = t_2 - t_1$ . Звідси для стаціонарного випадкового процесу, згідно з (3.20), одержуємо

$$m_x(t) = \int_{-\infty}^{\infty} x f_1(x) dx = m_x = const, \quad (3.22)$$

тобто математичне сподівання стаціонарного випадкового процесу не залежить від аргументу  $t$  і є постійною величиною.

Згідно з (3.21) вираз для кореляційного моменту буде наступний

$$K_x(t_1, t_2) = \iint_{-\infty}^{\infty} (x_1 - m_x)(x_2 - m_x) \cdot f_2(x_1, x_2; \tau) dx_1 dx_2 = K_x(\tau). \quad (3.23)$$

Таким чином, коваріаційна функція стаціонарного випадкового процесу є функцією тільки одного аргументу  $\tau = t_2 - t_1$ . Умови (3.22) і (3.23) виконуються для будь-якого стаціонарного процесу. Однак вони не є достатніми для визнання стаціонарності, тобто їхнє виконання не гарантує виконання умов (3.19). Таке розуміння стаціонарності випадкових функцій називають стаціонарністю у вузькому розумінні.

У прикладній теорії стаціонарних випадкових функцій користуються поняттям про стаціонарність у широкому розумінні. За звичай припускають, що стаціонарний випадковий процес повинний задовольняти три умови:

$$M[X(t_s)] = const, D[X(t_s)] = const; K_X[X(t_s), (t_k)] = K_\tau. \quad (3.24)$$

Таким чином, *випадкова функція називається стаціонарною, якщо усі її статистичні характеристики не залежать від  $t$ , точніше, не змінюються при будь-якому зсуві аргументів, від яких вони залежать.*

Однак не всі ці умови рівноцінні. Не можна вважати сталість математичного сподівання істотною вимогою до стаціонарної випадкової функції. При переході до центрованої випадкової функції відразу ж

виконується рівність  $M\{\delta[X(t)]\}=0$ , тобто перша умова (3.24). Отже, у випадку змінного математичного сподівання  $M[X(t)]$  досить перейти до центрування для того, щоб виконати одну з вимог до стаціонарного випадкового процесу.

Виконання другої умови (3.24) є частинним випадком третього. Для стаціонарного випадкового процесу характерна незалежність значень коваріаційної функції від положення відрізка  $\tau$  на осі  $t$ . Коваріаційна функція повинна залежати тільки від довжини проміжку  $\tau$  (3.23). Покладемо  $t_s = t; t_k = t + \tau$ . Якщо  $\tau = 0$ , то згідно (3.12)

$$D_x(t) = K_X[X(t), X(t+0)] = K_X(\tau=0) = const. \quad (3.25)$$

Іншими словами, у випадку стаціонарного випадкового процесу його дисперсія є не що інше, як значення коваріаційної функції при  $\tau = 0$ , тобто не є функцією аргументу. Слід зазначити, що з властивості симетричності коваріаційної і кореляційної функцій

$$K_x(\tau) = K_x(-\tau); \quad (3.26)$$

$$r_x(\tau) = r_x(-\tau), \quad (3.27)$$

витікає незалежність цих функцій від знака величини  $\tau$ , тобто можна вважати  $\tau = |t_2 - t_1|$ . Тому коваріаційну і кореляційну функції знаходять тільки для позитивного аргументу.

Найчастіше зустрічаються стаціонарні випадкові процеси, коваріаційні функції яких апроксимуються функціями наступних типів (рис. 3.7)

$$K_X(\tau) = \sigma^2 e^{-a|\tau|}, \quad a > 0 \quad (3.28)$$

$$K_X(\tau) = \sigma^2 e^{-a\tau^2}, \quad a > 0 \quad (3.29)$$

$$K_X(\tau) = \sigma^2 e^{-a|\tau|} \cdot \cos \beta\tau, \quad a > 0 \quad (3.30)$$

$$K_X(\tau) = \sigma^2 e^{-a\tau^2} \cdot \cos \beta\tau, \quad a > 0 \quad (3.31)$$

$$K_X(\tau) = \sigma^2 e^{-a|\tau|} \left( \cos \beta\tau + \frac{\alpha}{\beta} \sin \beta|\tau| \right), \quad a > 0, \beta > 0 \quad (3.32)$$

$$K_X(\tau) = \begin{cases} \sigma^2 \left(1 - \frac{\tau}{\tau_0}\right) & \text{при } \tau \leq \tau_0 \\ 0 & \text{при } \tau > \tau_0 \end{cases} \quad (3.33)$$

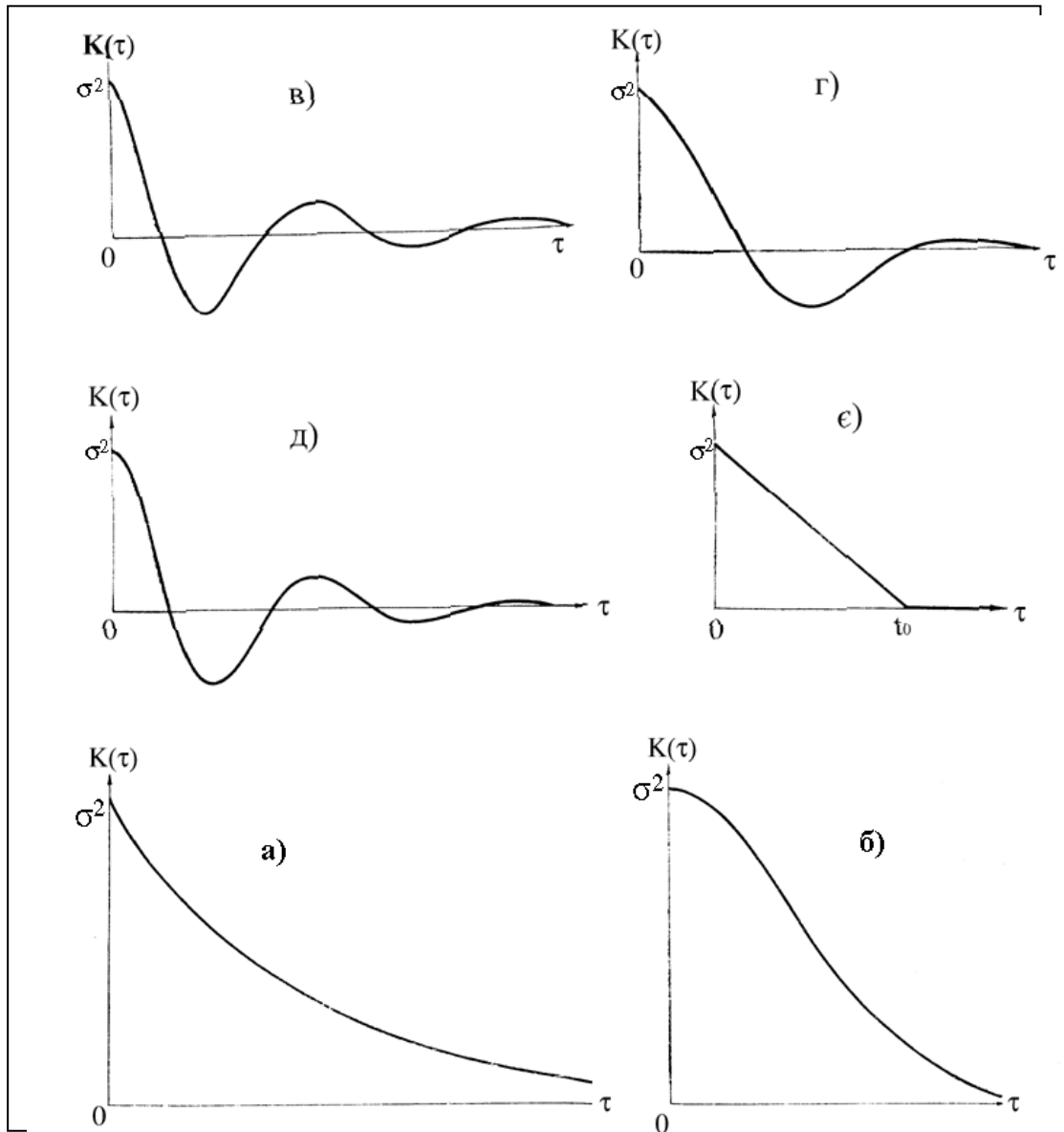


Рис. 3.7. Приклади коваріаційних функцій стандартних процесів

На рис. 3.7 наведені графіки коваріаційних функцій тільки для  $\tau > 0$ , у силу парності цих функцій при  $\tau < 0$  їм будуть відповідати криві, симетричні щодо осі ординат. З рис. 3.7 видно, що значення кореляційних функцій убивають з ростом  $\tau$ , тобто кореляційний зв'язок між різними

перетинами випадкових функцій убуває зі збільшенням інтервалу часу між ними. Криві на рис.3.7.а, 3.7.б мають вигляд гармонійних коливань з амплітудою, яка зменшується у часі. Вигляд цих кривих указує на наявність періодичності в структурі випадкової функції.

Одержання від'ємних значень  $K_X(\tau)$  означає прояв негативних зв'язків між перетинами випадкової функції. Для всіх наведених випадків коваріаційна, а, отже, і кореляційна функція прагне до нуля при прагненні  $\tau$  до нескінченності. Ця властивість виконується практично для всіх гідрометеорологічних рядів. Коли в структурі випадкової функції, як постійний доданок, є деяка випадкова величина, при  $\tau \rightarrow \infty$  коваріаційна функція буде прагнути не до 0, а до дисперсії  $D$  цієї величини (рис.3.8). Як правило, вигляд автокореляційної функції, наведений на рис.3.8, характерний для випадкових процесів, які знаходяться під впливом монотонно зростаючих антропогенних перетворень.

Переважній більшості стаціонарних випадкових функцій властива ергодичність. Це відноситься до випадкових функцій, що характеризуються стаціонарністю у вузькому і широкому розумінні.

*Властивість ергодичності полягає у тому, що кожна окрема реалізація стаціонарної випадкової функції є повноправним представником усієї сукупності можливих реалізацій. Одна реалізація достатньої довжини може замінити при статистичній обробці всі реалізації в цілому. Статистичні характеристики кожної реалізації є одними й тими ж для усіх реалізацій.*

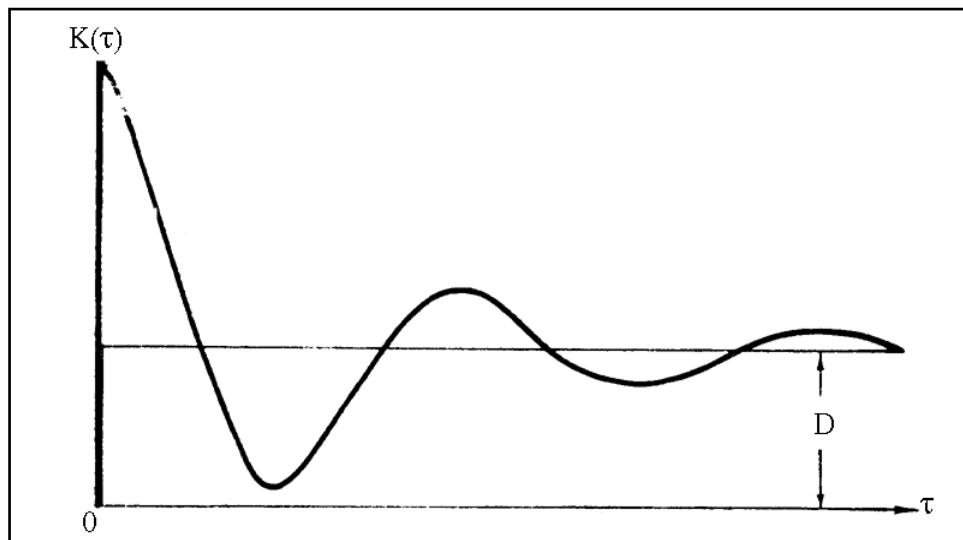


Рисунок 3.8 - Кореляційна функція при наявності постійного впливу на випадковий процес

Розглянемо дві стаціонарні випадкові функції  $X_1(t)$  і  $X_2(t)$ , наведені на рис.3.9. Кожна з реалізацій випадкової функції  $X_1(t)$  має одні й тіж



самі характерні ознаки: середнє значення, навколо якого коливаються реалізації; розмах коливань. Функція  $X_2$  таких властивостей не має. Якщо вибрати довільно одну з таких

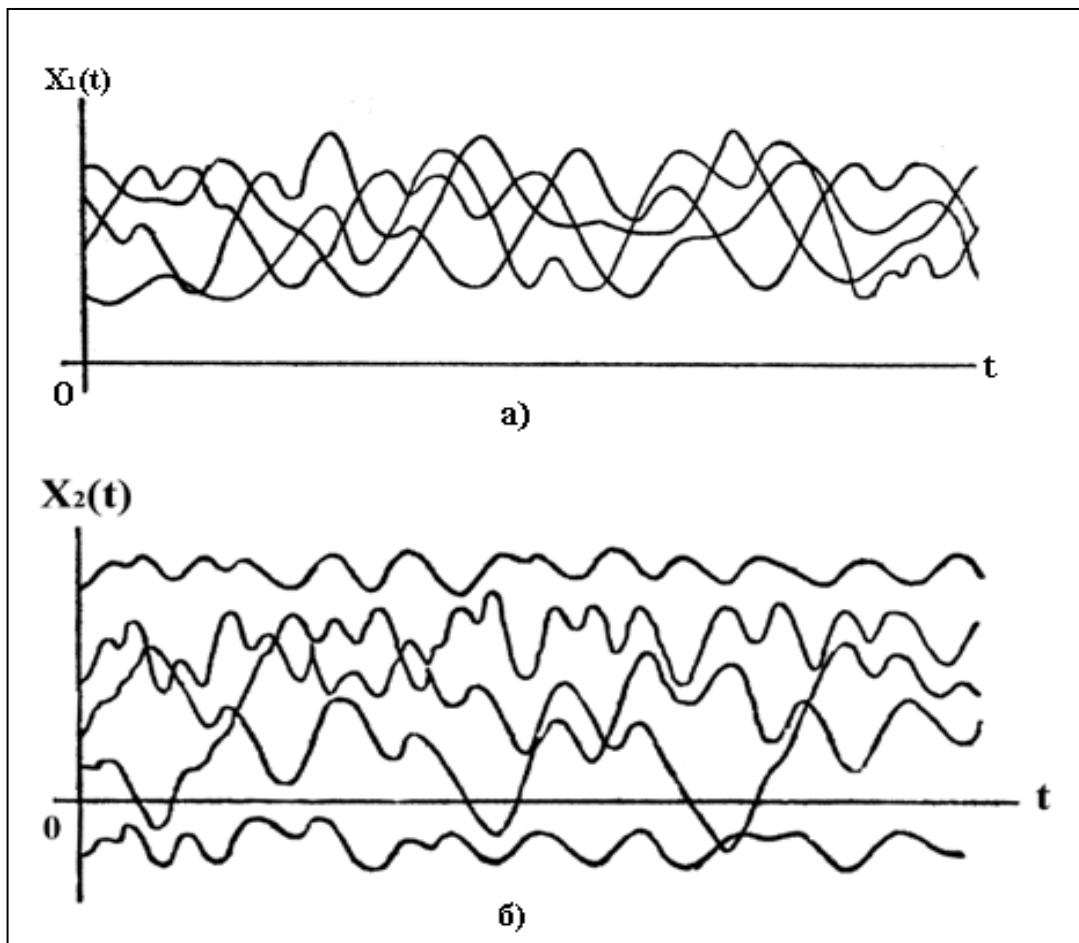


Рисунок 3.9 - Приклади стаціонарних випадкових процесів

реалізацій, то при досить довгій її тривалості  $T$  по ній можна одержати повне уявлення про випадкову функцію в цілому.

Осереднюючи значення цієї реалізації за часом, одержують оцінку математичного сподівання:

$$\hat{m}_x \approx \frac{1}{T} \int_0^T x(t) dt, \quad (3.34)$$

автокореляційної функції:

$$\hat{K}_x(\tau) \approx \frac{1}{T-\tau} \int_0^{T-\tau} [x(t) - \hat{m}_x] \cdot [x(t+\tau) - \hat{m}_x] dt. \quad (3.35)$$

Якщо розбити інтервал запису випадкової функції на  $n$  рівних частин довжиною  $\Delta t$  і позначити середини отриманих ділянок  $t_1, t_2, \dots, t_n$  (рис.3.10), то інтеграли (3.34) і (3.35) можна представити у вигляді:

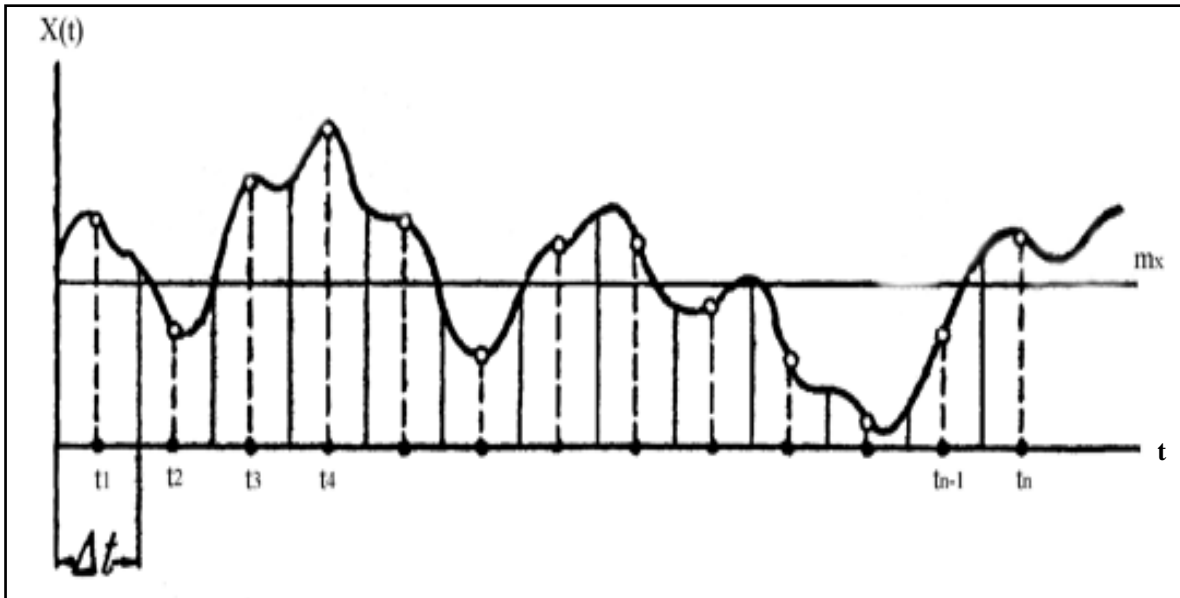


Рисунок 3.10 - Запис випадкової функції, яка розбивається на  $n$  різних частин

$$\hat{m}_x = \frac{1}{n} \sum_{i=1}^n x(t_i) = \bar{x}; \quad (3.36)$$

$$\hat{K}_x(\tau) = \frac{1}{n - \tau} \sum_{i=1}^{n-\tau} [x(t_i) - \bar{x}] \cdot [x(t_i + \tau) - \bar{x}], \quad (3.37)$$

де  $\tau = m\Delta t$ , причому  $m = 0, 1, 2, \dots$

У подальшому викладенні для простоти запису знак  $\hat{\phantom{x}}$ , який означає, що мова йде про оцінку характеристики за вибірковими даними, прибирається, а  $t_i$  представляється як  $i$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad (3.38)$$

$$K_x(\tau) = \frac{1}{n - \tau} \sum_{i=1}^{n-\tau} (x_i - \bar{x})(x_{i+\tau} - \bar{x}). \quad (3.39)$$

Загальний вигляд функції  $K_x(\tau)$  відтворюється по окремих точках. Для того, щоб математичне сподівання і кореляційна функція були визначені з задовільною точністю, потрібно, щоб число  $n$  було досить велике (порядку сотні, а в деяких випадках навіть і більше). Вибір довжини елементарної ділянки  $\Delta t$  визначається характером зміни випадкової функції. Якщо випадкова величина змінюється порівняно плавно, ділянки  $\Delta t$  можна вибирати більшими, ніж коли функція робить різкі і часті коливання. Чим більш високочастотний склад мають коливання, що утворюють випадкову функцію, тим частіше потрібно розміщувати опорні точки при обробці. Орієнтовно можна вибирати елементарні ділянки  $\Delta t$  так, щоб на повний період самої високочастотної гармоніки в складі випадкової функції припадало порядку 5-10 опорних точок.

Значення величини  $m$  задаються послідовно, аж до таких  $m$ , при яких кореляційна функція практично дорівнює нулю чи починає робити невеликі нерегулярні коливання біля нуля.

### 3.5 Автокореляційна функція, коефіцієнт автокореляції

У гідрологічних розрахунках річковий стік розглядається як ергодичний процес з річним періодом. Така гіпотеза зручна й необхідна, оскільки по кожному гідрологічному створу є лише одна реалізація процесу стоку, яка представлена гідрологічними спостереженнями. Проте точного математичного доказу справедливості цієї гіпотези немає. Очевидно, що прийнята гіпотеза буде повністю знехтувана, якщо вимірювати час геологічними масштабами. В межах проміжків, що охоплюються періодом гідрологічних спостережень у минулому і перспективного планування в майбутньому за умови незначних антропогенних навантажень, порушення ергодичності стоку маловірогідні.

Кореляційний аналіз є одним з традиційних методів в дослідженнях зв'язків між послідовними членами рядів річного стоку. Проте значне випадкове розсіювання вибірових оцінок коефіцієнтів кореляції при невеликій тривалості вихідних рядів і невеликих числових значеннях  $r_x(\tau)$  викликає природне утруднення в інтерпретації результатів розрахунку.

Ординати автокореляційної функції за вибіровими даними обчислюються з урахуванням зміщеності за наступною формулою:

$$r_x(\tau) = \frac{\sum_{i=1}^{n-\tau} (x_i - \bar{x})(x_{i+\tau} - \bar{x})}{\hat{\sigma}_x^2(n-\tau-1)}, \quad (3.40)$$

де  $n$  - число членів вихідного ряду спостережень;

$\tau$  - зсув ряду по відношенню до самого себе при підрахунку коефіцієнта кореляції (запізнювання);

$x_i$  - член ряду від  $x_i$  до  $x_{n-\tau}$ ;

$x_{i+\tau}$  - члени ряду від  $x_{i+\tau}$  до  $x_n$ ;

$\bar{x}, \hat{\sigma}_x$  - відповідно середнє і стандарт вихідного ряду.

В [3] рекомендується враховувати зміну середнього і стандарту вибірки з ростом запізнювання автокореляційної функції. При цьому (3.40) набуває вигляду:

$$r_x(\tau) = \frac{\sum_{i=1}^{n-\tau} (x_i - \bar{x})(x_{i+\tau} - \bar{x}_{i+\tau})}{\sigma_i \sigma_{i+\tau} (n - \tau - 1)}, \quad (3.41)$$

де  $x_i, \bar{x}_{i+1}, \sigma_i$  і  $\sigma_{i+\tau}$  - середні і стандарт відповідних відрізків вихідного ряду.

На думку більшості дослідників, вираз перетворює вибіркове значення  $r_x(\tau)$ . Суть перетворення полягає у тому, що лінія регресії проходить при цьому через центр тяжіння не всієї сукупності точок, а тільки точок, які безпосередньо використовуються для підрахунку коефіцієнта кореляції: перетворення посилюється із зростанням запізнювання, причому чим більше  $T$ , тим сильніше відрізняються середнє і стандарт вихідного ряду від відповідних значень тієї його частини, яка при зсуві використовується для підрахунку коефіцієнта кореляції.

Характеру автокореляційних функцій рядів річного стоку присвячена велика кількість досліджень як у колишньому СРСР, так і за межами країни. Найбільш суттєві роботи А.Ш.Резніковського, О.І.Чеботарьова і А.В.Рождественського, Д.Я.Ратковича, М.В. Болгова. На цей же час встановлена наявність позитивного зв'язку між річним стоком суміжних років. Загальноприйнятого погляду на значущість внутрішньорядних зв'язків в послідовностях річного стоку поки немає. Навіть при відносно тривалих рядах стоку (50-60 років) недостовірні як окремі ординати автокореляційної функції, так і її контури. Тому при розрахунках річного стоку його ряди розглядаються як простий ланцюг Маркова, тобто враховуються кореляційні зв'язки тільки між стоком суміжних років. *Послідовність випадкових випробувань, які мають таку властивість, що ймовірність результатів наступного випробування залежить тільки від результатів безпосередньо попередніх випробувань, називаються простим ланцюгом Маркова.* Таким чином, з усіх ординат автокореляційної функції береться до уваги тільки перша, яка відповідає  $\tau = 1$ .

Коефіцієнт кореляції між стоком суміжних років обчислюють згідно з виразом (3.41), виходячи з умови, що часовий зсув становить одиницю ( $\tau = 1$ ):

$$r(\tau = 1) = r(1) = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (3.42)$$

Середня квадратична похибка емпіричного коефіцієнта автокореляції визначається з рівняння:

$$\sigma_{r(1)} = \frac{1 - r(1)^2}{\sqrt{n-1}}. \quad (3.43)$$

Розрахункове значення коефіцієнта кореляції вважається значущим, якщо виконується умова

$$r(1) \geq 2\sigma_{r(1)}. \quad (3.44)$$

Внутрішньорядні зв'язки у рядах річного стоку зменшують обсяги незалежної інформації. В той же час методи розрахунків статистичних параметрів широко припускають незалежність членів вихідної вибірки. Для усунення впливу внутрішньорядних зв'язків між суміжними членами рядів річного стоку на точність розрахунків статистичних параметрів до коефіцієнтів варіації та асиметрії вносяться поправки

$$C_v = (a_1 + a_2/n) + (a_3 + a_4/n)\tilde{C}_v + (a_5 + a_6/n)\tilde{C}_v^2; \quad (3.45)$$

$$C_s = (b_1 + b_2/n) + (b_3 + b_4/n)\tilde{C}_s + (b_5 + b_6/n)\tilde{C}_s^2, \quad (3.46)$$

де  $a_1, \dots, a_6; b_1, \dots, b_6$  - коефіцієнти, які визначаються по табл.3.2 і 3.3;  
 $\tilde{C}_v, \tilde{C}_s$  - оцінки коефіцієнтів варіації і асиметрії, установлені без урахування внутрішньорядних зв'язків

$$\tilde{C}_v = \sqrt{\frac{\sum_{i=1}^n (k_i - 1)^2}{n-1}}; \quad (3.47)$$

$$\tilde{C}_s = \frac{n}{(n-1)(n-2)} \frac{\sum_{i=1}^n (k_i - 1)^3}{\tilde{C}_v^3}; \quad (3.48)$$

де  $k_i = x_i / \bar{x}$ .

Таблиця 3.2 – Коефіцієнти  $a$  у формулі (3.45)

$C_s / C_v$	$r(1)$	$a_1$	$a_2$	$a_3$	$a_5$	$a_6$	$a_7$
1	2	3	4	5	6	7	8
2.00	0.00	0.00	0.19	0.99	-0.88	0.01	1.54
	0.30	0.00	0.22	0.99	-0.41	0.01	1.51
	0.50	0.00	0.18	0.98	0.41	0.02	1.47
3.00	0.00	0.00	0.69	0.98	-4.34	0.01	6.78
	0.30	0.00	1.15	1.02	-7.53	-0.04	12.38
	0.50	0.00	1.75	1.00	-11.79	-0.05	21.13
4.00	0.00	0.00	1.36	1.02	-9.68	-0.05	15.55
	0.30	-0.02	2.61	1.13	-19.85	-0.22	34.15
	0.50	-0.02	3.47	1.18	-29.71	-0.41	58.08

Таблиця 3.3 – Коефіцієнти  $b$  у формулі (3.46)

$r(1)$	$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$
0.00	0.03	2.00	0.92	-5.09	0.03	8.10
0.30	0.03	1.77	0.93	-3.45	0.03	8.03
0.50	0.03	1.63	0.92	-0.97	0.03	7.94

Середні квадратичні похибки вибірових параметрів обчислюються таким чином:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \sqrt{\frac{1+r(1)}{1-r(1)}}; \quad (3.49)$$

$$\sigma_{c_v} = \frac{C_v}{n+4C_v^2} \sqrt{\frac{n(1+C_v^2)}{2} \left( 1 + \frac{3C_v r(1)^2}{1-r(1)} \right)}. \quad (3.50)$$

### 3.6 Приклад розрахунків

*Розрахувати коефіцієнт автокореляції та уточнити статистичні параметри річного стоку з водозбору р. Дерекойка – м. Ялта за 39-ти річний період спостережень*

Вихідні дані та результати обчислень оформлюють у вигляді табл.3.4.

- середнє багаторічне значення річного стоку  $\bar{q}=10.5$  л/с км<sup>2</sup>;
- середньоквадратичне відхилення  $\sigma_q = 4.2$  л/с км<sup>2</sup> ;
- коефіцієнт варіації річного стоку  $\hat{C}_v = 0.40$  ;
- коефіцієнт асиметрії річного стоку  $C_s = 1.16$  ;
- співвідношення  $\hat{C}_s / \hat{C}_v = 2.89 \approx 3.00$
- середня квадратична похибка розрахунку параметра  $\bar{q}$ :  $\sigma_{\bar{q}} = 0.67$  л/с· км<sup>2</sup>;

Коефіцієнт автокореляції  $r(1)$  при зсуві  $\tau = 1$ , розрахований по (3.42), дорівнює 0.38, а середньоквадратична похибка його  $\sigma_{r(1)}$  становить 0.152. Умова (3.45) виконується, тобто  $0.380 > 2 \times 0.152$ , через що значення параметра  $r(1)$  слід визнати значущим. Використовуючи формули (3.45) та (3.46), а також табл. 3.2 і 3.3, визначаємо параметри  $\hat{C}_v$  і  $\hat{C}_s$  з урахуванням внутрішньорядних зв'язків. Уточнені значення параметрів такі:  $\hat{C}_v = 0.40$ ,  $\hat{C}_s = 1.39$ ,  $\hat{C}_s / \hat{C}_v = 3.43$ ,  $\sigma_q = 0.67$ ,  $\sigma_{Cv} = 0.049$ . Як показують результати розрахунків, вплив внутрішньорядних зв'язків особливо виражений при розрахунках коефіцієнта асиметрії та співвідношення  $\hat{C}_s / \hat{C}_v$ .

Таблиця 3.4 – Обчислення коефіцієнту автокореляції річного стоку з водозбору р.Дерекойка - м.Ялта

№ ПП	$q_i$ л/скм <sup>2</sup>	$q_{i+1}$ л/скм <sup>2</sup>	$q_i - \bar{q}$	$(q_i - \bar{q})^2$	$q_{i+1} - \bar{q}$	$(q_i - \bar{q}) \times$ $(q_{i+1} - \bar{q})$
1	2	3	4	5	6	7
1	11.3		0.80	0.64		
2	10.3	11.3	-0.20	0.04	0.80	-0.16
3	9.46	10.3	-1.04	1.08	-0.20	0.21
4	8.45	9.46	-2.05	4.20	-1.04	2.13
5	7.44	8.45	-3.06	9.36	-2.05	6.27
6	13.9	7.44	3.40	11.6	-3.06	-10.4
7	14.5	13.9	4.00	16.0	3.40	13.6
8	7.85	14.5	-2.65	7.02	4.00	-10.6
9	9.66	7.85	-0.84	0.71	-2.65	2.23
10	13.3	9.66	2.80	7.84	-0.84	-2.35
11	4.43	13.3	-6.07	36.8	2.80	-17.0
12	7.85	4.43	-2.65	7.02	-6.07	16.1
13	9.05	7.85	-1.45	2.10	-2.65	3.84
14	8.25	9.05	-2.25	5.06	1.45	3.26
15	9.26	8.25	-1.24	1.54	-2.25	2.79
16	11.5	9.26	1.00	1.00	-1.24	-1.24
17	8.05	11.5	-2.45	6.00	1.00	-2.45
18	8.25	8.05	-2.25	5.06	-2.45	5.51
19	13.5	8.25	3.00	9.00	-2.25	-6.75
20	10.5	13.5	0.00	0.00	3.00	0.00
21	17.5	10.5	7.00	49.0	0.00	0.00
22	20.1	17.5	9.60	92.2	7.00	67.2
23	10.1	20.1	-0.40	0.16	9.60	-3.84
24	10.9	10.1	0.40	0.16	-0.40	-0.16
25	3.82	10.9	-6.68	44.6	0.40	-2.67
26	4.00	3.82	-6.50	42.2	-6.60	43.4
27	10.5	4.00	0.00	0.00	-6.50	0.00
28	7.24	10.5	-3.26	10.6	0.00	0.00



Продовження таблиці 3.4

№ пп	$q_i$ л/скм <sup>2</sup>	$q_{i+1}$ л/скм <sup>2</sup>	$q_i - \bar{q}$	$(q_i - \bar{q})^2$	$q_{i+1} - \bar{q}$	$(q_i - \bar{q}) \times$ $(q_{i+1} - \bar{q})$
1	2	3	4	5	6	7
29	9.26	7.24	-1.24	1.54	-3.26	4.04
30	8.85	9.26	-1.65	2.72	-1.24	1.55
31	6.84	8.85	-3.66	13.4	-3.66	6.04
32	5.63	6.84	-4.87	23.7	-4.87	17.8
33	11.5	5.63	1.00	1.00	1.00	-4.87
34	12.9	11.5	2.40	5.76	2.40	2.40
35	24.7	12.9	14.2	202	14.2	3.40
36	15.9	24.7	5.40	29.2	5.40	76.7
37	14.7	15.9	4.20	17.6	4.20	22.7
38	8.45	14.7	-2.05	4.20	-2.05	-8.61
39	9.86	8.45	-0.64	0.41	-0.64	1.31
Сума				672		262

## РОЗДІЛ 4

### СТОХАСТИЧНІ МОДЕЛІ В ГІДРОЛОГІЧНИХ РОЗРАХУНКАХ

В основі більшості інженерно-гідрологічних методів розрахунку лежить уявлення про ймовірнісний характер коливань водності річок і інших складових водного балансу. Існуюча методологія вирішення гідрологічних задач ґрунтується на уявленні про стаціонарність багаторічних коливань стоку, а основні математичні підходи розвиваються на базі теорії стаціонарних випадкових процесів. Зокрема, стохастична теорія багаторічних коливань стоку широко використовує як математичні моделі процеси (ланцюги) Маркова[23, 24].

Випадковий процес  $X(t)$  характеризується марковськими властивостями, якщо  $P(x(t)/x(t_i), i=0,1,...)$  - умовний розподіл імовірностей значень  $x(t)$ , за умови, що задані значення  $x(t_i)$  в попередні моменти часу  $t < t_i, i=0,1,2,...$  залежать тільки від значення  $x(t_0)$  в наступний момент

$$P\{x(t)/x(t_i), i=0,1,...\} = P\{x(t)/x(t_0)\}, \quad (4.1)$$

тобто при фіксованому значенні майбутнє марковського процесу не залежить від його попереднього. Це означає, що вся інформація про минуле для такого процесу полягає в його значенні в останній момент часу. Однак, звідси не випливає, що майбутнє не корелює з більш далеким минулим. Відповідна кореляційна функція має вид експоненти  $exp\{\alpha, t\}$ . Для дискретного марковського процесу випадкову послідовність (ланцюг Маркова), що відповідає даному вище визначенню, іноді ще називають простим ланцюгом Маркова, для того, щоб відрізнити її від більш загальної (складної), для якої залежність від минулого описується не одним попереднім значенням  $x(t_0)$ , а двома, трьома і ін. Методи розв'язання рівняння Маркова для випадкових величин з одновимірними законами розподілу із сімейства кривих К. Пірсона були розроблені в 50-і роки О.В.Сармановим [31-33].

Ряди річного стоку не можуть розглядатися як послідовності незалежних випадкових величин: у коливаннях річного стоку мають місце серії маловодних і багатоводних років, що свідчить про наявність

внутрішньорядних зв'язків. Аналіз автокореляційних функцій рядів річного стоку показав, що вони у першому наближенні задовільно описується простим ланцюгом Маркова [2,4,8,10,11,19,21,22,35].

Різні модифікації опису простого ланцюга Маркова приводять до різних результатів водогосподарського розрахунку, навіть при однаковому значенні їхнього числового параметра – коефіцієнта автокореляції. Тому для опису багаторічних коливань стоку варто рекомендувати таку з цих модифікацій, яка краще відповідає закономірностям коливань, що спостерігаються в природі стоку, і не призводить до істотних помилок у результатах водогосподарських розрахунків. Для зіставлення матеріалів натурних спостережень зі стохастичними моделями використовуються двовимірні щільності розподілу ймовірностей. З метою одержання шуканих еталонних співвідношень, які відповідають розглянутим стохастичним моделям, широко використовується метод Монте-Карло.

*У простого ланцюга Маркова розподіл ймовірностей наступної випадкової величини залежить лише від значення безпосередньо попередньої величини; ступінь залежності (при заданому характері зв'язаності) визначається величиною коефіцієнта кореляції між суміжними членами ряду.*

Використання простого ланцюга Маркова як стохастичної моделі стоку припускає загасання коефіцієнта кореляції між стоком послідовних років

$$r^\tau = r_1^\tau, \quad (4.2)$$

де  $r_\tau$  - коефіцієнт кореляції між величинами стоку в роки  $i$  та  $i + \tau$ ;  
 $\tau$  - число років.

Величина коефіцієнта кореляції між стоком безпосередньо суміжних років  $r_1 = r(1)$  визначає характер всієї автокореляційної функції, чим однак не досягається вичерпний опис характеру зв'язку між послідовними членами ряду. У такій послідовності розподіл ймовірностей якого-небудь члена ряду (умовний розподіл) залежить від величини попереднього члена. Умовний розподіл ймовірностей може відрізнитися як параметрами, так і типом від безумовного розподілу, справедливого для сукупності всіх членів ряду.

Сукупність умовних розподілів утворить двовимірний розподіл ймовірностей, що і є вичерпною характеристикою зв'язку між членами корельованої послідовності, яка представляє собою простий ланцюг Маркова. Геометричною інтерпретацією двовимірного розподілу є поверхня над координатною площиною, по одній з осей якої відкладається величина попереднього члена, а по іншій – наступного; ординати поверхні є щільності ймовірності. При тому самому вигляді автокореляційної

функції можна сконструювати необмежену кількість варіантів двовимірного розподілу – модифікацій описів простого ланцюга Маркова; нижче розглядається декілька характерних модифікацій, які застосовуються для опису коливань стоку.

Модель 1 припускає відсутність кореляції між членами ряду – послідовність незалежних випадкових величин. Якщо ряд стоку розглядається як вибірка незалежних випадкових величин, то відповідні значення забезпеченості стоку є незалежними випадковими величинами, рівномірно розподіленими в інтервалі від 0 до 1. У цьому випадку всі умовні розподіли ймовірностей збігаються з безумовним. Генерування штучних рядів стоку без урахування кореляції між стоком суміжних років розглянуто в розділі 1.

Модель 2 розглядає кореляцію між нормально-розподіленими випадковими величинами (з наступним переходом до гама – розподілених величин). У цій модифікації послідовність забезпеченостей нормально-розподілених корельованих величин приймається за послідовність забезпеченостей випадкових величин, які характеризуються гама-розподілом.

При відомому значенні попереднього  $x$  члена ланцюга Маркова умовний розподіл наступного члена  $y$  є, так само як і безумовний розподіл величин  $x$ , нормальним

$$f(y/x) = \frac{1}{\sigma_{y/x} \sqrt{2\pi}} \exp \left[ -\frac{(y - m_{y/x})^2}{2\sigma_{y/x}^2} \right] \quad (4.3)$$

з умовним математичним очікуванням, що змінюється за лінійним законом в залежності від величини попереднього члена

$$m_{y/x} = \bar{x}_{i+1/i} = \bar{x} + r_H(x_i - \bar{x}), \quad (4.4)$$

і умовним стандартом

$$\sigma_{y/x} = \sigma_{i+1/i} = \sigma \sqrt{1 - r_H^2}, \quad (4.5)$$

де  $f(x_{i+1/i})$  чи  $f_{y/x}$  умовна щільність імовірностей випадкової величини

$y = x_{i+1}$ ;

$m_x = \bar{x}$  - безумовне математичне очікування розподілу випадкової величини  $x$ ;

$\sigma$  - безумовний стандарт розподілу випадкової величини  $x$ ;

$r_n$  - коефіцієнт кореляції між суміжними членами ряду нормально-розподілених величин.

Вирази (4.4) і (4.5) визначають лінійність кореляції і незалежність умовного стандарту від величини попереднього члена (гомоскедастичність). Однак це є, апріорі, справедливим тільки для нормально-розподілених випадкових величин, а для гама - розподілених порушується як лінійність кореляції, так і гомоскедастичність умовного розподілу; ступінь нелінійності кореляції зростає зі збільшенням коефіцієнта варіації безумовного розподілу.

Генерування рядів стоку за моделлю 2 здійснюється в наступній послідовності.

1. Генерація випадкових чисел, що у подальшому розглядаються як забезпеченості, тобто ( $\delta_i = P_i$ ), де  $0 < \delta_i < 1$ .

2. Генерація значень стоку заданої забезпеченості за умови, що кореляція між стоком суміжних років відсутня ( $r_n = 0$ ).

У цьому випадку кожне значення модельованого ряду позначається як  $x_i^o$ .

3. Як розподіл, береться нормальний закон, у якому використовуються центровані і нормовані вихідні величини, представлені у вигляді:

а) для безумовного розподілу

$$\Phi_{i+1}^o = \frac{x_{i+1}^o - \bar{x}}{\sigma}; \quad (4.6)$$

б) для умовного розподілу

$$\Phi_{i+1/i} = \frac{x_{i+1/i} - \bar{x}_{i+1/i}}{\sigma_{i+1/i}} \quad (4.7)$$

Для визначення  $\Phi_{i+1}^o$  і  $\Phi_{i+1/i}$  застосовують таблиці інтегральної функції нормального закону розподілу. Оскільки між інтегральною функцією і функцією забезпеченості існує зв'язок, представлений у вигляді  $F(x) = 1 - P(x) = 1 - \delta$ , то можна зробити висновок, що кожному випадковому числу  $\delta$  відповідають однакові значення  $\Phi_{i+1}^o$  і  $\Phi_{i+1/i}$ , тобто

$$\frac{x_{i+1/i} - \bar{x}}{\sigma_{i+1/i}} = \frac{x_{i+1}^o - \bar{x}}{\sigma_x} \quad (4.8)$$

З урахуванням позначень (4.4) і (4.5) одержимо

$$x_{i+1/i} = \bar{x} + r_H(x_i - \bar{x}) + (x_{i+1}^o - \bar{x})\sqrt{1 - r_H^2}, \quad (4.9)$$

де  $x_{i+1/i}$  - кожне наступне значення модельованого ряду з урахуванням його залежності від попереднього значення;

$x_{i+1}^o$  - кожне наступне значення модельованого ряду при  $r_H = 0$ .

4.3 модельований за (4.9) ряд поєднує в собі властивості нормально розподілених величин. Для переходу до гама-розподілених величин вираз (4.9) представляють у вигляді

$$\Phi_{i+1} = r_H \Phi_i + \Phi_{i+1}^o \sqrt{1 - r_H^2}, \quad (4.10)$$

де

$$\Phi_{i+1} = \frac{x_{i+1/i} - \bar{x}}{\sigma}; \Phi_i = \frac{x_i - \bar{x}}{\sigma}; \Phi_{i+1}^o = \frac{x_{i+1}^o - \bar{x}}{\sigma}. \quad (4.11)$$

За виразом (4.10) визначається значення  $\Phi_{i+1}$ . Шукане (уточнене для гама-розподілу) значення забезпеченості  $P_{i+1}$  знаходять з виразу

$$P_{i+1} = 1 - F(\Phi_{i+1}), \quad (4.12)$$

де  $F(\Phi_{i+1})$  - інтегральна функція нормального закону розподілу.

Модель 3 розглядає кореляцію між випадковими величинами, які характеризуються гама – розподілом, із застосуванням апарату гама – кореляції (за Є.Г.Блохиновим-О.В.Сармановим [2]). У гідрологічних додатках безумовна щільність гама-розподілених величин  $x$  записується в таких позначеннях

$$p(x) = \frac{\gamma^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-\gamma x}, \quad (4.13)$$

де параметр  $\gamma$ , зв'язаний з коефіцієнтом варіації  $C_v$  співвідношенням

$$\gamma = \frac{1}{C_v^2} \quad (4.14)$$

Для щільності (4.6) безумовна двовимірна й умовна щільності записуються у такий спосіб:

$$f(x, y) = p(x)p(y) \left[ 1 + \sum_{k=1}^{\infty} r^k L_k^{\gamma-1}(\gamma x) L_k^{\gamma-1}(\gamma y) \right]; \quad (4.15)$$

$$f(y/x) = p(y) \left[ 1 + \sum_{k=1}^{\infty} R^k L_k^{\gamma-1}(\gamma x) L_k^{\gamma-1}(\gamma y) \right], \quad (4.16)$$

при  $\lambda_k = k\lambda, k = 1, 2, \dots, \lambda > 0,$

де  $y = x_{i+1}$ ;

$L_k^{\gamma}(x)$  - поліноми Лагера;

$R = \exp(-\lambda)$  - коефіцієнт кореляції.

Як і в розглянутому вище нормальному розподілі величин, що варіюють, кореляція є лінійною й умовне математичне очікування визначається за формулою (4.4). Стандарт умовного розподілу при використанні апарату гама-кореляції відповідно дорівнює

$$\sigma_{y/x} = \sigma_{i+1/i} = \sigma \sqrt{(1-r_r)^2 + 2x_i r_r (1-r_r)}, \quad (4.17)$$

де  $r_r$  - коефіцієнт кореляції гама – розподілених величин.

Умовний коефіцієнт асиметрії набирає значення

$$C_{S_{y/x}} = \frac{2C_v \left[ (1-r_r)^3 + 3r_r(1-r_r)^2 x_i \right]}{\left[ (1-r_r)^2 + 2r_r(1-r_r)x_i \right]^{3/2}}, \quad (4.18)$$

умовний ексцес -

$$E_{y/x} = \frac{6C_v^2 \left[ (1-r_r)^4 + 4r_r(1-r_r)^3 x_i \right]}{\left[ (1-r_r)^2 + 2r_r(1-r_r)x_i \right]^2} \quad (4.19)$$

Модель 4 використовує кореляцію між випадковими величинами, які характеризуються гама - розподілом, із застосуванням наближеного прийому кореляції (за С.Н.Крицьким і М.Ф.Менкелем [15,16]).

Наближений прийом кореляції гама - розподілених величин виходить з припущень, що умовний розподіл є, так само як і безумовний, гама - розподілом, причому умовний стандарт зв'язаний з безумовним тим же співвідношенням (4.5), що і при нормальному розподілі випадкових величин.

Перераховані модифікації описів простої моделі Маркова виходять з кореляції між суміжними значеннями випадкових величин (у даному випадку середньорічних витрат води). При цьому вплив попереднього члена на розподіл імовірностей наступного (умовний розподіл) виявляється тим більшим, чим сильніше відхиляється попередній член від середнього значення.

Однак можна представити й інший принцип конструювання ланцюга Маркова для опису стокових рядів, які спираються на наявність кореляції між забезпеченостями суміжних членів ряду; при цьому математичне очікування умовного розподілу ймовірностей наступного члена буде змінюватися в тій мірі, у якій значення забезпеченості попередніх членів відхиляються від медіани; кореляція буде зв'язувати випадкові величини, розподілені рівномірно в інтервалі від нуля до одиниці. Особливістю такого роду кореляції є те, що структура послідовності забезпеченостей випадкових величин не залежить ні від типу їхнього розподілу, ні від числових значень параметрів цього розподілу. Таким чином, модель 5 застосовує кореляцію між забезпеченостями (рівномірно розподіленими випадковими величинами).

Нехай  $x$  і  $y$  - випадкові величини, які розподілені рівномірно у проміжку  $[-1, 1]$  і мають відповідні щільності

$$f_1(x) = \frac{1}{2}, \quad f_2(x) = \frac{1}{2} \quad (4.20)$$

Потрібно побудувати двовимірну щільність  $f(x, y)$  у квадранті  $[-1 \leq x; y \leq 1]$  з урахуванням заданого коефіцієнта кореляції  $r_0$ . Шукана щільність представлена у вигляді розкладання по поліномах Лежандра. Поліноми Лежандра визначаються формулою Родріга

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[ (x^2 - 1)^n \right]; \quad n = 0, 1, 2, \dots \quad (4.21)$$

Зокрема,

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1) \quad (4.22)$$



Якщо обмежитися в білінійному розкладанні за системою ортонормованих поліномів першими трьома членами, то двовимірну щільність запишеться у вигляді:

$$f(x, y) = \frac{1}{4} \left[ 1 + 3r_0 xy + \frac{5}{4} r_0^2 (3x^2 - 1)(3y^2 - 1) \right], \quad (4.23)$$

де  $r_0$  - коефіцієнт кореляції рівномірно розподілених у  $[-1, 1]$  випадкових величин.

У [33] показано, що розкладання (4.16) зберігає ймовірнісний зміст при виконанні наступної умови

$$-0,55 \leq \lambda \leq 0,55 \quad (4.24)$$

Для побудови двовимірного закону розподілу рівномірно-розподілених випадкових величин у квадранті  $0 \leq u, v \leq 1$  (забезпеченостей) необхідно здійснити лінійну заміну змінних  $x = 2u - 1$ ,  $y = 2v - 1$ . Щільність  $f(u, v)$  запишеться у вигляді

$$f(u, v) = f(P_{i+1}, P_i) = 1 + 3\lambda(2u - 1)(2v - 1) + \frac{5}{4} \lambda^2 [3(2u - 1)^2 - 1] \cdot [3(2v - 1)^2 - 1] \quad (4.25)$$

Умовна щільність розподілу величини  $v$  буде представлена виразом

$$f(v/u) = \frac{f(u, v)}{f_1(u)}, \quad (4.26)$$

а умовна функція розподілу -

$$F(v/u) = \int_0^v f(u, v) dv = v + 3\lambda(2u - 1)(v^2 - v) + 5\lambda^2(6u^2 - 6u + 1)(2v^3 - 3v^2 + v) \quad (4.27)$$

Умовна функція розподілу для величини  $u$  записується у вигляді

$$F(u/v) = u + 3r_0(2v-1)(u^2 - u) + \frac{5}{2}r_0^2[3(2v-1)^2 - 1](2u^3 - 3u^2 + u) \quad (4.28)$$

Умовне математичне очікування дорівнює

$$m(u/v) = \bar{P}_{i+1/i} = \frac{1}{2} + \frac{r_0}{2}(2P_i + 1), \quad (4.29)$$

а умовна дисперсія

$$\sigma_{u/v}^2 = \sigma_{i+1/i}^2 = \frac{1-r_0^2}{12} \quad (4.30)$$

Умовна дисперсія не залежить від попереднього значення забезпеченості  $P_i$  і, отже, кореляція гомоскедастична.

Рівняння (4.27) є лінійним спільним законом розподілу рівномірно розподілених величин (наприклад, забезпеченостей). Цього досить для того, щоб моделювати ряд забезпеченостей необхідної довжини із заданим значенням коефіцієнта кореляції між ними  $|r_0| \leq 0,55$ . Для моделювання випадкових значень використовується такий спосіб: задаючи значення  $u = u_1$  і підставляючи в (4.27), одержують умовну функцію розподілу  $v$  для значень  $u = u_1$ ; далі з таблиці випадкових чисел  $[0, 1]$  беруться значення функції розподілу  $F(v/u)$  і для знаходження випадкового значення  $v_1$  вирішуємо кубічне рівняння (4.27); для генерування випадкового часового ряду із заданими властивостями отримане значення  $v_1$  ставиться на місце  $u_1$  й описана процедура генерації повторюється ще необхідне число раз. Незважаючи на простоту й універсальність методу, мають місце дві обставини, які необхідно взяти до уваги. Лінійна кореляція між забезпеченостями приводить (у даній теоретичній схемі) до нелінійної кореляції між самими випадковими величинами. Якщо ж спочатку будувати лінійну кореляцію між вихідними випадковими величинами, то кореляція між їхніми забезпеченостями буде нелінійною. У цьому випадку необхідно використовувати відповідну умовну функцію розподілу і, вирішуючи зворотну задачу, знаходити відповідне умовне значення по заданих  $F(x/y)$  і  $r_0$ .

Для моделювання випадкового процесу із заданим маргінальним законом розподілу необхідно вказати співвідношення між коефіцієнтом кореляції для забезпеченостей  $r_0$  і коефіцієнтом кореляції для вихідних випадкових величин  $r_r$ . Зокрема, для двовимірного закону, одержуваного

шляхом обмеження в розкладанні першими трьома членами і з нормальними апіорними щільностями, справедлива формула [24]

$$r_r = \frac{3}{\pi} r_0 \quad (4.31)$$

При груповому аналізі матеріалів спостережень за стоком, коли точні числові параметри розподілу ймовірностей для кожного окремого пункту невідомі, а іноді сам тип розподілу недостатньо достовірний, зручніше розглядати ряди забезпеченостей. У роботі [17] генерування штучних рядів річного стоку виконувалося стосовно до значень коефіцієнтів кореляції, які змінюються від 0 до 0,55 [22-24]. Генерування рядів стоку здійснювалося як моделювання послідовностей випадкових величин, зв'язаних кореляцією марковськ нелінійного типу з маргінальним трипараметричним гама-розподілом С.М.Крицького і М.Ф.Менкеля (модель 5) [15,16]. При цьому попередньо була послідовність з 4000 випадкових рівномірно розподілених чисел в інтервалі (0,1), зв'язаних лінійною лежандровською кореляцією.

## РОЗДІЛ 5

### МОДЕЛЮВАННЯ РІЧНОГО СТОКУ В УМОВАХ ВОДОГОСПОДАРСЬКИХ ПЕРЕТВОРЕНЬ

Методологія гідрологічних розрахунків базується на часових рядах спостережень за стоком. При цьому приймається умова про збереження ергодичності процесу і репрезентативність матеріалів спостережень. Ряди побутового (перетвореного діяльністю людини) річного стоку не можуть розглядатися як реалізації ергодичного випадкового процесу, а, отже, до них неможливо застосовувати стандартні статистичні методи. Для вивчення водного режиму річок в умовах антропогенної діяльності необхідна побудова математичної моделі стоку, яка дозволяє імітувати стан водних ресурсів в залежності від виду водогосподарських перетворень. У останні десятиріччя до проблеми оцінки стоку в умовах водогосподарських перетворень, таких як створення штучних водойм, забір води з місцевих водотоків і перекидання стоку, осушувальні і зрошувальні меліорації додалася проблема оцінки впливу на водні ресурси змін глобального клімату.

Вирішення поставлених проблем потребує реалізації таких етапів науково-дослідної роботи:

- оцінки характеристик стоку в природних (непорушених господарською діяльністю) умовах його формування;
- оцінки характеристик стоку в умовах проведення водогосподарських заходів (усередині окремих водозборів);
- оцінка характеристик стоку в результаті можливих змін клімату;
- узагальнення результатів досліджень з метою розробки моделі розрахунку характеристик стоку в умовах антропогенного впливу, включаючи як водогосподарські перетворення на водозборах (внутрішньобасейнові антропогенні чинники), так і зміни клімату.

Основою подібного роду досліджень є детальне вивчення закономірностей просторово-часового розподілу характеристик стоку. Традиційні методи розрахунку характеристик стоку “при недостатності і відсутності даних спостережень” [20] виявилися неприйнятними, у зв'язку з чим авторами розроблена методика оцінки характеристик річного стоку на основі методу водно-теплогового балансу з використанням метеорологічних даних.

### 5.1. Огляд методів оцінки річного стоку в порушених господарською діяльністю умовах

Головна особливість сучасних стохастичних моделей полягає в прийнятті гіпотези стаціонарності [8,11,12,24,26]. Але якщо розглядати геофізичні процеси за достатньо тривалий історичний період (сотні і тисячі років), то їх слід вважати нестаціонарними. Порушення однорідності і стаціонарності рядів гідрологічних характеристик відбувається за рахунок зростаючого впливу господарської діяльності на водозборах і руслах річок, а також за рахунок антропогенної зміни сучасного клімату. Ця обставина ставить перед сучасними ученими проблему оцінки характеристик стоку в умовах антропогенного впливу, вирішення якої може бути досягнуте за рахунок математичного моделювання [14,16,24,34 і ін.] Моделі часових рядів, як правило, бувають трьох видів: стохастичні, детерміновані і детерміновано-стохастичні. Для пояснення цього питання використовуємо системний підхід.

Природна гідрологічна система може бути представленою у вигляді одного векторного рівняння [13,14]

$$F_1(Y, \varepsilon, \Lambda, x, t), \quad (5.1)$$

де  $Y$  - вектор параметрів, що характеризують стан річкового басейну;

$\varepsilon$  - вектор зовнішніх впливів;

$\Lambda$  - вектор заданих функцій і коефіцієнтів, що характеризують фізичні, геометричні, морфологічні і т.п. параметри системи;

$x$  - вектор просторових координат;

$t$  - час.

За звичай, (5.1) представляється в диференціальному вигляді

$$\frac{\partial Y}{\partial t} + L(\Lambda, Y) = \varepsilon, \quad (5.2)$$

де  $L$  - оператор, який є складовою частиною математичної моделі, що описує перетворення стоку.

Інтегруючи (5.2) за часом, отримаємо

$$Y(t) = Y(t_0) - \int L(\Lambda, Y)dt + \int \varepsilon dt, \quad (5.3)$$

де  $Y(t_0)$  - початковий стан системи.

Таким чином, знаючи початковий стан системи  $Y(t_0)$  і прогнозуючи зовнішній вплив  $\varepsilon$ , можна імітувати поведінку гідрологічної системи, вирішуючи тим або іншим способом (5.3). Рівняння (5.3) можна розглядати як динамічну систему, в якій під зовнішнім впливом слід розуміти стокоформуванняльні кліматичні чинники, а під  $L$  - оператори моделей, які враховують динаміку гідрологічних процесів в ланцюгу “схиловий стік - русловий стік”.

Якщо йдеться про гідрологічну систему, яка знаходиться під “пресом” водогосподарських перетворень, то розглядається природно-технічна система. Складова  $Y(t_0)$  у такому разі характеризує початковий стан гідрологічної системи в непорушених господарською діяльністю умовах. Зміни параметрів водозборів (площ осушених боліт, площ водної поверхні штучних ставків і водосховищ, площ зрошуваних масивів) включаються у вектор  $\Lambda$ . Тоді  $L$  - набір операторів, що враховують зміни стоку в результаті водогосподарських перетворень на водозборах.

Таким чином, задача оцінки впливу господарської діяльності на стік зводиться до пошуку оператора перетворення початкового стану гідрологічної системи (в даному випадку це природний стік) в змінений водогосподарськими перетвореннями. Математичне представлення цього оператора може бути різним: функція відгуку, інтеграл згортки, функція пам'яті і т. ін.

У загальному випадку рівняння (5.3) носить стохастичний характер. Так, величину  $Y(t_0)$  можна задати тільки в ймовірнісній формі. Зовнішні впливи (опади, випаровування) також є випадковими процесами. Більш того, і антропогенні впливи включають випадкову складову, оскільки їх інтенсивність багато в чому визначається коливанням кліматичних факторів. Так, вплив сукупності ставків і водосховищ, як штучних об'єктів ландшафту, визначається не тільки сумарною площею водної поверхні, але і додатковим випаровуванням з цієї поверхні, яке, у свою чергу, залежить від водності року.

Класичний підхід до вирішення проблеми оцінки характеристик стоку в порушених водогосподарськими перетвореннями умовах формування стоку полягає в аналізі і розрахунках стоку в природних умовах і оцінці впливу окремих видів господарської діяльності [36,37].

Серед методів урахування впливу господарської діяльності на річний стік річок можна виділити наступні.

1. Метод порівняння характеристик стоку в природному і порушеному господарською діяльністю станах. Недоліки цього підходу полягають у відсутності даних по стоку річок в природних умовах його формування, що особливо актуально для південних областей України, де початок

спостережень за стоком практично співпадає з інтенсивним господарським освоєнням території. Саме у зв'язку з нестачею початкових даних нами і була розроблена методика оцінки характеристик природного річного стоку на основі методу водно-теплогового балансу, яка передбачає використання метеорологічних даних. Крім того, метод порівняння дозволяє одержати дані про зміну стоку тільки в інтегральному вигляді, без урахування динаміки водогосподарських факторів і їхнього індивідуального внеску в зміну стоку.

2. Метод водогосподарського балансу базується на урахуванні безпосередніх скидів і заборів води на водозборах

$$W_B = W_E - \Delta W_{II} + \Delta W_{CB}, \quad (5.4)$$

де  $W_B, W_E$  - об'єми побутового і природного стоку;

$\Delta W_{II}$  - безповоротні вилучення стоку з поверхневих водотоків (втрати);

$\Delta W_{CB}$  - скиди води в поверхневі водотоки.

Переваги цього підходу полягають в можливості оцінити вплив кожного водогосподарського чинника окремо. Але вирішення проблеми знову ж таки пов'язане із задачами приведення стоку до природного виду. Крім того, детальні експериментальні дані і надійні матеріали обліку водоспоживання і водовідведення відсутні, що ставить під сумнів якість отриманих оцінок.

Найперспективнішими в оцінці впливу водогосподарських перетворень на стік є методи імітаційного фізичного і математичного моделювання.

3. Методи фізичного моделювання найчастіше застосовуються для дослідження ділянок русел, при цьому виявляються можливі зміни режиму річки. Фізичне моделювання цілих водозборів використовується дуже рідко і, по суті, є унікальним.

4. В основі застосування методів математичного моделювання найчастіше лежить детерміністичний підхід. На початку відбувається вивчення процесу і складання детерміністичних рівнянь, які описують в часі і просторі зміну стоку під впливом господарської діяльності, а потім виконується числовий експеримент: за розробленою схемою та за різних початкових умов і при різних параметрах господарської діяльності програвуються різні ситуації на водозборі. Методи математичного моделювання дозволяють кількісно оцінити вплив найрізноманітнішого поєднання природних і антропогенних чинників і тим самим надають можливість розрахувати і спрогнозувати водний режим у майбутньому.

Оператором перетворення  $L$  початкового стану гідрологічної системи в рівнянні (5.2), яке описує її зміну при водогосподарських діях,

може служити нейромережева модель. Як науковий предмет штучні нейронні мережі вперше заявили про себе в 40-і роки минулого сторіччя. Прагнучи відтворити функції людського мозку, дослідники створили прості апаратні, а пізніше - програмні, моделі біологічного нейрона і системи його з'єднань. Перше систематичне вивчення штучних нейронних мереж було зроблене Маккаллоком і Піттсом в 1943 році. Пізніше вони досліджували мережеві парадигми для розпізнавання зображень, що піддаються зсуву і поворотам. Проста нейронна модель використовувалася в більшій частині їх робіт. Елемент  $\Sigma$  множить кожен вхід  $x$  на вагу  $w$  і підсумовує зважені входи. Якщо ця сума більше заданого порогового значення, вихід дорівнює одиниці, інакше – нулю. Ці системи (і багато їм подібних) одержали назву персептронів. Розенблатт довів теорему про навчання персептронів (Розенблатт Ф., 1965), але свій повноцінний розвиток нейромережі одержали після того, як був запропонований ефективний алгоритм навчання багат шарових персептронів, заснований на обчисленні градієнта функції помилки. Застосований для навчання нейромережі метод зворотного розповсюдження помилки фактично є узагальненням методу найменших квадратів стосовно багат шарових нейромереж. Здатність до навчання дала можливість застосування даного алгоритму для вирішення таких задач, як розпізнавання образів, класифікація, прогнозування. Зокрема, використовуючи дані про водоспоживання, можна виконати навчання нейромережі і моделювати побутовий стік залежно від заданого рівня господарської діяльності. Проте, для успішного вирішення цієї задачі необхідно створити еталони образів розпізнавання, що є надзвичайно трудомістким процесом, який вимагає досить повної інформації про величини водоспоживання на водозборі. Основна мета функціонування нейросистеми як системи штучного інтелекту полягає, по-перше, в можливості її навчання для відтворення динаміки того або іншого процесу. По-друге, і це головне, нейромережевий алгоритм повинен вирішувати задачі прогнозу і відтворення подальшої часової еволюції процесу або системи на основі виконаного навчання. В зв'язку з цим перспективи застосування нейромережевого підходу в задачах вивчення географічних і гідрологічних систем представляються багатообіцяючими.

Але, нейромережева модель не відповідає повністю поставленим задачам моделювання побутового стоку, оскільки у водогосподарському балансі присутні як невідповідні величини, обумовлені діяльністю людини, так і випадкові, пов'язані з коливаннями клімату. Вирішення проблеми оцінки характеристик річного стоку в умовах водогосподарських перетворень на водозборах може бути досягнуте шляхом стохастичного моделювання, при якому навчання мережі замінюється імітаційним



експериментом, що дозволяє моделювати різні ситуації, пов'язані як із змінами клімату, так і з господарською діяльністю.

## 5.2. Математичне моделювання побутового річного стоку на водозборах річок зони недостатнього зволоження

Розроблена в ОДЕКУ математична модель річного стоку в умовах водогосподарських перетворень на водозборах поєднує в собі стохастичний і детерміністичний підходи до розрахунків річного стоку, тобто відноситься до числа детерміновано-стохастичних моделей [9,10,17,18]. У її основі лежить рівняння водогосподарського балансу водозбору (5.4), яке для більшості водозборів зони недостатнього зволоження в межах України можна детальніше представити таким чином:

$$W_{\text{ПОБ}} = W_{\text{ПР}} - \Delta W_{\text{ЗР}} - \Delta W_{\text{ВИП}} + \Delta W_{\text{СК}}, \quad (5.5)$$

де  $W_{\text{ПОБ}}$ ,  $W_{\text{ПР}}$  - об'єми побутового і природного стоку;

$\Delta W_{\text{ЗР}}$  - об'єм безповоротних втрат стоку на зрошування;

$\Delta W_{\text{ВИП}}$  - об'єм втрат стоку на додаткове випаровування з водної поверхні ставків і водосховищ;

$\Delta W_{\text{СК}}$  - об'єм додаткового скидання води в поверхневі водотоки.

На першому етапі створення моделей розглядалося питання про оцінку окремих складових ( $\Delta W_{\text{ОР}}$ ,  $\Delta W_{\text{ИСП}}$ ,  $\Delta W_{\text{СБ}}$ ) водогосподарських балансів, які містять в собі інформацію про характер господарської діяльності в межах водозборів. Надалі здійснювався перехід до ймовірнісної форми запису рівняння (5.5), на основі якої виконувалося накладання антропогенних чинників на коливання природного стоку.

Стохастичне моделювання рядів побутового стоку виконувалося роздільно для кожного з досліджуваних факторів антропогенного впливу на основі рівнянь водогосподарських балансів, які враховують реальну фізичну картину взаємодії того або іншого чинника господарської діяльності з річним стоком [9,10,17]:

а) при заборах води для цілей зрошування

$$Y_{\text{ПОБ},P} = Y_{\text{ПР},P} - \frac{M_{0,100-P}}{\eta} f_{\text{ЗР}} ; \quad (5.6)$$

б) за наявності на водозборі штучних водойм з сезонним регулюванням стоку, коли втрати обумовлені додатковим випаровуванням з водної поверхні

$$Y_{\text{ПОБ},P} = Y_{\text{ПР},P}(1 - f_B) - (E_B - X)f_B \quad (5.7)$$

або

$$Y_{\text{ПОБ},P} = Y_{\text{ПР},P}(1 - \alpha_P f_B), \quad \text{де} \quad \alpha_P = \varphi(Y_P); \quad (5.8)$$

в) за наявності скидних вод із зрошуваних масивів

$$Y_{\text{ПОБ},P} = Y_{\text{ПР},P} + \xi \frac{M_{0,100-P}}{\eta} (1 - \eta) f_{3P}, \quad (5.9)$$

де  $Y_{\text{ПОБ},P}, Y_{\text{ПР},P}$  - побутовий і природний стік із заданою забезпеченістю  $P$ , виражений в мм;

$f_{3P}, f_B$  - сумарна площа зрошуваних масивів і водної поверхні штучних водоймищ, виражені в частках від загальної площі водозбору  $F$ ;

$\eta$  - коефіцієнт корисної дії зрошувальної системи;

$\xi$  - коефіцієнт зворотних вод, що утворилися за рахунок втрат стоку при його перекиданні по каналах зрошувальної мережі;

$X$  - опади, які випали на загальну водну поверхню ставків і водосховищ;

$E_B$  - сумарне випаровування з водної поверхні ставків і водосховищ, визначення якого виконувалося по відомих методиках [36,37];

$M_{0,100-P}$  - зрошувальна норма-нетто з ймовірністю перевищення  $100 - P$ ;

$\alpha_P$  - безрозмірний коефіцієнт, який дозволяє врахувати зміну  $(E_B - X)$  залежно від величини річного стоку забезпеченістю  $P$ .

Генерування рядів природного річного стоку виконувалося як моделювання послідовностей випадкових величин, зв'язаних кореляцією марківського нелінійного типу з маргінальним трипараметричним гамарозподілом С.М.Крицького і М.Ф.Менкеля (модель 5) на основі методики, наведеної в розділі 4.

Стохастичне моделювання допускає наявність вхідних даних про статистичні параметри річного стоку в природних умовах його формування. Для визначення параметрів природного річного стоку можуть бути використані як результати спостережень за стоком в природних умовах його формування, так і метеорологічні дані, а також результати

статистичного районування на основі методів багатовимірною статистичного аналізу.

Рівняння водогосподарського балансу вигляду (5.6 – 5.9), які утворюють основу детерміновано-стохастичних моделей, можна описати, використовуючи системний підхід, викладений вище. Звертаючись до (5.3), початковий стан гідрологічної системи  $Y(t_0)$  представляється як її стан в природних умовах формування стоку, тобто у вигляді характеристики річного стоку  $Y_{E,P}$ , тоді як  $L$  - антропогенна складова рівняння водогосподарського балансу, яка відображає динаміку процесу перетворення природного стоку в побутовий. Зовнішні впливи  $\varepsilon$  розглядаються як результат дії кліматичних факторів, які на даному етапі дослідження вважаються незмінними, тобто  $\varepsilon = 0$ .

Рівняння (5.6 - 5.9) враховують взаємозв'язки між ймовірнісними характеристиками стоку (як природного, так і побутового) і ймовірнісними характеристиками антропогенних чинників, що мають стохастичну природу. У них досягнутий певний компроміс між достатньою фізичною обґрунтованістю моделі і компактністю забезпечення її параметрами, які можна визначити на основі даних стандартних спостережень мережі гідрометеорологічних станцій. Таким чином, при створенні детерміновано-стохастичних моделей побутового стоку, крім вирішення задачі побудови стохастичної моделі природного річного стоку, виникає проблема стохастичного опису таких елементів водогосподарських балансів як втрати стоку на додаткове випаровування з водної поверхні ставків і водосховищ, дефіцит водоспоживання води рослинами і т.ін.

### **5.2.1 Обґрунтування стохастичної моделі, що описує вплив штучних водоймищ на річний стік**

Такий вид водогосподарських перетворень в межах водозборів як створення штучних водойм з метою багаторічного і переважно сезонного регулювання стоку одержав інтенсивний розвиток в кінці 50-х та на початку 60-х років минулого сторіччя. Для північно-західного Причорномор'я було характерним створення так званих “копанок”, розташованих в 50-200 м від русел, в які вода закачувалася насосами. Згодом з'ясувалося, що ефективність такого роду споруд невисока. Рівнинна місцевість з ґрунтами, які легко змиваються, примітивні водоскидні споруди привели, кінець кінцем, до замулювання водойм і, отже, до погіршення якості води. В умовах посушливого клімату, з високими теплоенергетичними ресурсами, такі водоймища починають виконувати роль випарників, істотно знижуючих стік річок. Вплив ставків

і водосховищ як штучних об'єктів ландшафту може бути врахований наступним балансовим співвідношенням [17,36,37]

$$Y_{\text{ПОБ}}F = Y_{\text{ПР}}(F - \sum F_B) - (E_B - X)\sum F_B, \quad (5.10)$$

де  $Y_{\text{ПОБ}}, Y_{\text{ПР}}$  - побутовий і природний стік, мм;

$E_B$  - випаровування з водної поверхні, мм;

$X$  - сумарні опади, мм;

$\sum F_B$  - сумарна площа водної поверхні штучних водойм при НПР (нормальному підпірному рівні) в межах водозбору.

Втрати стоку на додаткове випаровування з водної поверхні визначаються різницею  $(E_B - X)$ , яка підкоряється зональному (широтному) закону розподілу і зв'язана з водністю року, а також з сумарною площею водної поверхні  $\sum F_B$ . Безпосередньо для моделювання використовувався вираз вигляду (5.8), де  $\alpha_P$  - безрозмірний коефіцієнт, представлений у вигляді функції водності року

$$\alpha_P = \frac{602}{Y_{\text{ПР},P}^{1,14}} \quad (5.11)$$

У рівнянні (5.8) виконано розділення невинуваткової (тобто директивно заданої) компоненти  $f_B$  і "кліматичної" складової  $\alpha_P = \phi(Y_P)$ , яка підлягає стохастичним законам розподілу. По суті, залежності (5.7) і (5.8) описують реальний фізичний зв'язок між втратами річного стоку на додаткове випаровування з водної поверхні і водністю року, тобто нормою природного річного стоку необхідної забезпеченості, яка задається як  $P$ . Чим вище водність року, тим менше втрати на додаткове випаровування з водної поверхні штучних водойм. Запропонований прийом дозволяє виконувати чисельні експерименти в моделюванні, при яких значення  $f_B$  у кожному випробуванні задається постійним, що дозволяє зберегти стаціонарність ряду побутового стоку, який генерується  $\xi$ .

Певну проблему становить визначення природного стоку південних регіонів України, де відсутні матеріали про стік як у природних, так і порушених господарською діяльністю умовах. Для визначення норм природного стоку в ОДЕКУ розроблена карта норм так званого кліматичного стоку (рис.5.1), розрахованого на основі водно-теплого балансу з використанням метеорологічної інформації. Кліматичний стік ототожнюється з природним стоком річок, на яких підземне живлення не залежить від площі водозбору. Для невеликих та середніх за розміром водозбірних площ річок з нестійким підземним живленням розроблені

перехідні коефіцієнти від норм кліматичного стоку до природного, які залежать від показників факторів підстильної поверхні [17].

Оскільки значні втрати на додаткове випаровування з поверхні штучних водоймищ спостерігаються в зоні недостатнього зволоження, при генерації рядів природного стоку використовувалися норми кліматичного стоку, задані в діапазоні від 50 до 10мм, характерні для південних регіонів України.

Результатом такого роду експериментів є отримані залежності оцінок статистичних параметрів побутового стоку від не випадкового показника  $f_B$ , які для середньобагаторічних величин річного стоку можуть бути описані наступною залежністю

$$\bar{Y}_{ПОВ} = \bar{Y}_K e^{-\alpha f_B}, \quad (5.12)$$

де  $\alpha$  - коефіцієнт, що характеризує інтенсивність зниження норми природного стоку із зростанням відносної площі водної поверхні  $f_B$ .

Величина коефіцієнта  $\alpha$ , у свою чергу, визначається співвідношенням ресурсів тепла і вологи  $\beta_H = \frac{\bar{X}}{\bar{E}_m}$

$$\alpha = -0.622(1 - \beta_H), \quad \text{при } \beta_H \geq 0,45,$$

де  $\bar{X}$  -ресурси зволоженості, представлені у вигляді карти ізоліній норм річних опадів (рис. 5.2);

$\bar{E}_m$  - теплоенергетичні ресурси клімату, представлені у вигляді карти ізоліній норм максимального можливого випаровування (рис.5.3). Максимально можливе випаровування – це шар води, який міг би випаритися з поверхні суші, якщо б на нього були витрачені тільки позитивні складові теплового балансу поверхні суші.

Вплив штучних водойм, що не мають значної регулюючої місткості, в зоні недостатнього зволоження обертається для водних ресурсів не тільки зниженням норми річного стоку, але й збільшенням його багаторічної мінливості і асиметричності розподілу.

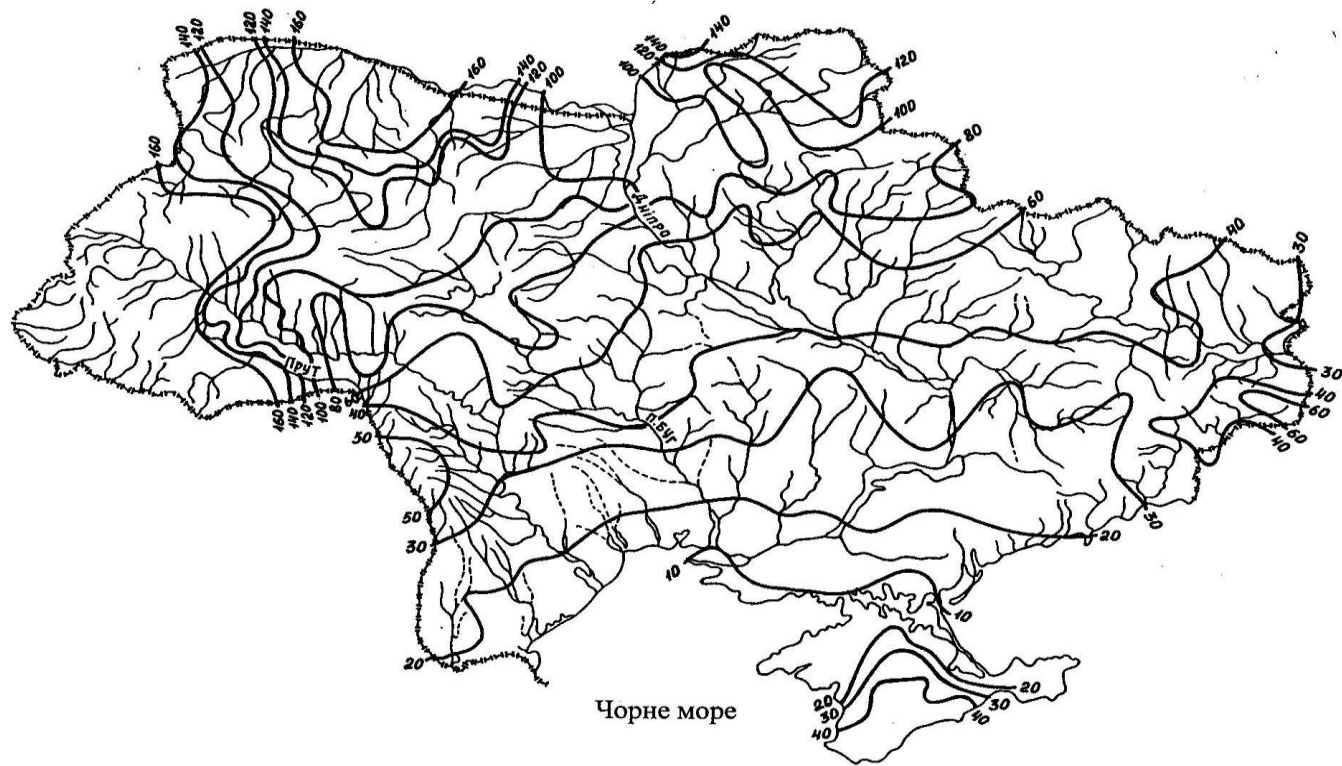


Рис. 5.1 Норма річного кліматичного стоку(мм)

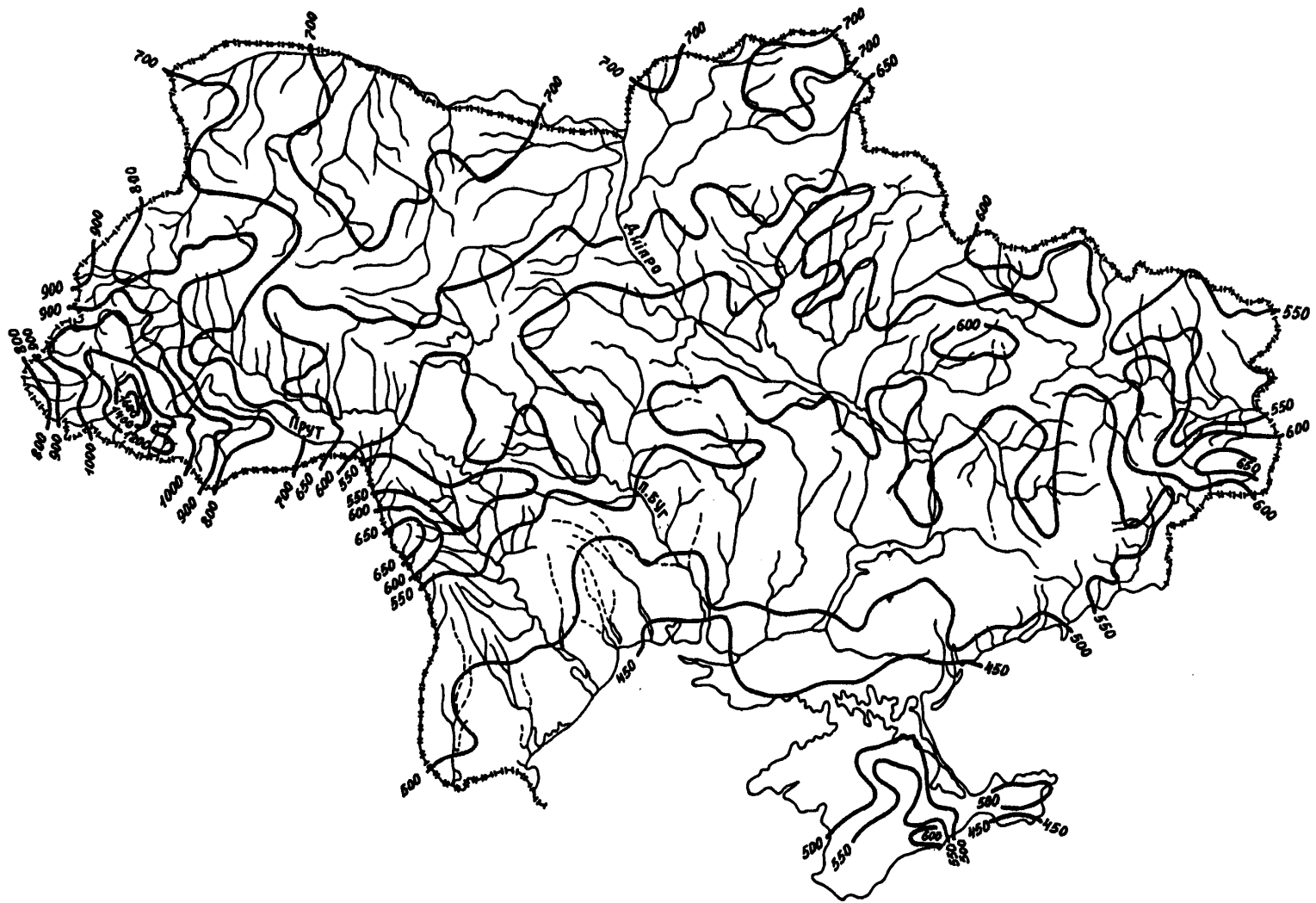


Рисунок 5.2 - Норма річних опадів (мм)

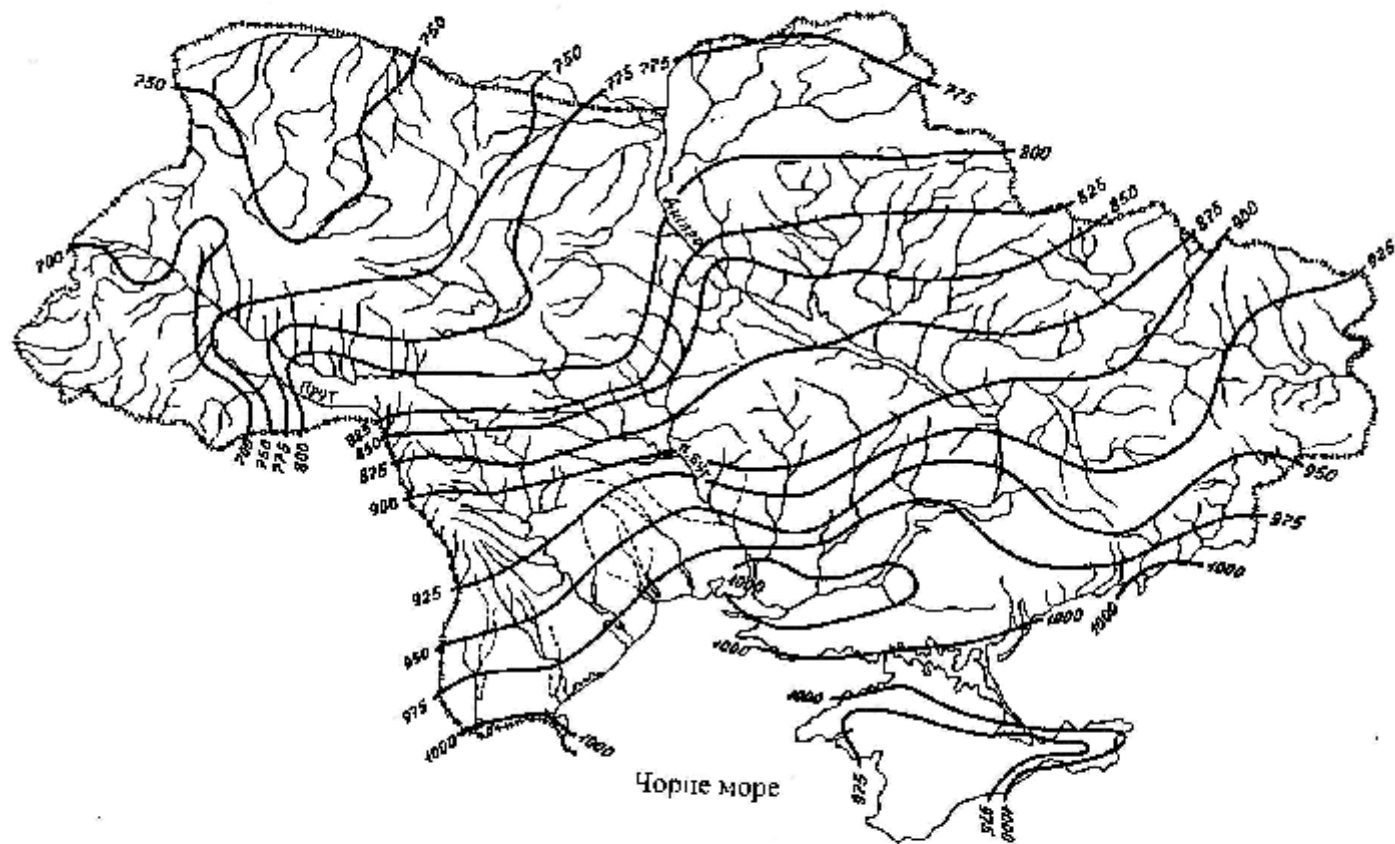


Рисунок 5.3 - Норма максимально можливого випаровування (мм)



Залежності коефіцієнтів варіації і асиметрії побутового стоку від відносної площі водної поверхні штучних водойм можна представити у вигляді регресійних залежностей

$$C_{VB} = 1,50 - 0,0275\bar{Y}_K + 0,0576f_B; R = 0,902, \quad (5.13)$$

$$C_{SB} = 1,80 - 0,0325\bar{Y}_K + 0,1643f_B; R = 0,870, \quad (5.14)$$

де  $C_{VB}, C_{SB}$  - коефіцієнти мінливості і асиметрії річного стоку за наявності на водозборах штучних водойм;

$\bar{Y}_K$  - норма кліматичного стоку, мм;

$f_B$  - відносна площа водної поверхні при НПР, %;

$R$  - коефіцієнт лінійної множинної регресії.

При цьому перший доданок в (5.14) дорівнює граничному максимальному значенню коефіцієнта варіації природного річного стоку  $C_{v, \max} = 1,5$ , яке відповідає мінімальному значенню кліматичного стоку на півдні України -  $\bar{Y}_K = 10 \text{ мм}$ .

Посилення нерівномірності багаторічного розподілу стоку свідчить, перш за все, про зниження стоку маловодних років, яке і в природних умовах є лімітуючим для сільського господарства. Таким чином, інтенсивність змін стоку стає більш вираженою при посиленні посушливості клімату, інтегральною характеристикою якого може бути норма кліматичного стоку.

### 5.2.2 Обґрунтування стохастичної моделі, яка описує вплив зрошувальних меліорацій на річний стік

Рівняння водогосподарського балансу при вилученні вод з поверхневих водотоків для цілей зрошування у варіанті, придатному для стохастичного моделювання, представлено виразом (5.6). Його складовою частиною є зрошувальна норма-нетто, для оцінки якої використовувався підхід, заснований на методі сумісного вирішення рівнянь водного і теплового балансів [9,17]. Необхідна кількість поливної води  $m_0$  визначається за формулою [9]

$$m_0 = v_0^r E_m - X, \quad (5.15)$$

де  $v_0$  - безрозмірна характеристика середнього за весь міжполивний період рівня зволоження ґрунту, при якому розвиток переважаючих на ділянці сільськогосподарських культур відбувається найбільш оптимально;

$r$  - параметр, який залежить від водно-фізичних властивостей ґрунту, для легких ґрунтів -  $r = 1,1 \div 1,5$ ; для середніх ґрунтів -  $r = 1,5 \div 2,0$ ; для важких ґрунтів -  $r = 2,0 \div 3,0$ ;

$E_m$  - максимально можливе випаровування за розрахунковий період, мм;

$X$  - атмосферні опади, які випали на зрошувану ділянку за розрахунковий (вегетаційний) період.

Оптимальна відносна вологість ґрунту  $v_0$  устанавлюється за рівнянням

$$v_0 = \frac{w_0}{w_{HB}}, \quad (5.16)$$

де  $w_0$  - величина оптимального запасу вологи в розрахунковому шарі ґрунту;

$w_{HB}$  - найменша вологомісткість ґрунту цього ж шару.

Різним культурам відповідає різний оптимум зволоження ґрунту. Для таких вологолюбних культур як рис,  $v_0$  звичайно приймається рівним  $1,0 \div 1,2$ , для посухостійких зернових культур -  $v_0 = 0.8$  і т.п.

Зрошувальна норма-нетто розраховується як сума поливних норм за весь період вегетації з розрахунку на середньоструктурний гектар зрошуваних в межах водозбору земель (табл.5.1)

$$M_0 = \sum m_{0i} \quad (5.17)$$

Статистична обробка розрахованих значень зрошувальних норм, виконана по метеорологічних станціях півдня України при різних нормах оптимального зволоження ґрунту, дозволила встановити, що коефіцієнти внутрішньорядної кореляції невеликі і їх порядок відповідає відносній похибці їх визначення. Коефіцієнт варіації зрошувальних норм  $Cv_{M_0}$  може бути оцінений за емпіричною формулою вигляду

$$Cv_{M_0} = \frac{100v_0}{M_0} \quad (5.18)$$

Таблиця 5.1. – Норми зрошування у роки різної водності

Норми річного кліматичного стоку, мм	Норми зрошування, мм			
	Рівень оптимального зволоження ґрунту $V_0$	Маловодні роки	Середні за водністю роки	Багатоводні роки
10	0,8	375	252	9,6
10	0,9	600	454	273
10	1,0	884	727	493
20	0,8	273	149	5,0
20	0,9	479	354	212
20	1,0	741	563	424
30	0,8	260	131	7
30	0,9	453	293	166
30	1,0	710	520	372

У свою чергу, коефіцієнти асиметрії зрошувальних норм близькі до нуля. У зв'язку з цим стохастична генерація рядів зрошувальних норм проводилася як моделювання ряду незалежних випадкових величин, що задовольняють функції безумовного розподілу. Як правило, емпіричні функції розподілу зрошувальних норм апроксимуються двопараметричними функціями, наприклад, логарифмічно-нормального розподілу або гама-розподілу.

Так само, як і при оцінці впливу штучних водоймищ, антропогенну складову  $\Delta W_{OP}$  рівняння (5.5) можна розділити на кліматичну ( $M_0$ ) і антропогенну компоненти ( $f_{OP}, \eta$ ). Базова частина рівняння (5.5) виражається у вигляді зрошувальних норм сільськогосподарських культур  $M_0$ . Закономірності багаторічних коливань останніх задовольняють вимогу стохастичної стаціонарності процесу, а значить до них можуть бути застосовані ті ж математичні схеми, що і при статистичному описі річкового стоку.

Щоб урахувати при моделюванні рядів побутового стоку відповідність стоку і зрошувальних норм по забезпеченості без великих втрат в точності розрахунку, можна використовувати угруповання маловодних (забезпеченість стоку  $P > 75\%$ ), багатоводних ( $P < 25\%$ ) і середніх по водності ( $25 \leq P \leq 75\%$ ) років (табл.5.1). Зокрема, у відповідність величинам річного стоку, які відносяться до групи

маловодних років ( $P > 75\%$ ), ставилися зрошувальні норми  $M_0$ , усереднені в інтервалі забезпеченостей  $P < 25\%$ . І, навпаки, значення річного стоку невеликої забезпеченості ( $P < 25\%$ ) комбінувалися з середніми значеннями дефіцитів водоспоживання, які мають ймовірність перевищення  $P > 75\%$ . Такий підхід виправдовується реальними фізичними процесами, що відбуваються на водозборі: у посушливі роки зростає дефіцит водоспоживання сільськогосподарських культур, а значення річного стоку зменшуються; у багатоводні роки необхідність у великих витратах води на зрошування відпадає.

Результатами моделювання є так звані функції відгуку водозборів на водогосподарські перетворення, представлені у вигляді залежностей статистичних параметрів від рівня водогосподарських перетворень. Характер побудованих по змодельованих рядах кривих  $\bar{Y}_{3P} = \varphi(f_{3P})$ ,  $C_{v_{3P}} = \varphi(f_{3P})$  і  $C_{s_{3P}} = \varphi(f_{3P})$  для території півдня України і Молдови аналогічний тим, які були одержані при оцінці впливу штучних водоймищ на параметри річного стоку. Із зростанням величин  $f_v$  і  $f_{3P}$  відбувається зменшення норми  $\bar{Y}$  і збільшення коефіцієнтів  $C_v$  і  $C_s$ . Причому ефект впливу господарської діяльності багато в чому визначається зволоженістю території. В міру переходу в посушливі райони природні водні ресурси зменшуються і їхня реакція на один і той же рівень антропогенної дії стає більш вираженою. Інтенсивність зміни параметрів річного стоку у міру переходу в посушливі регіони посилюється.

На рис. 5.4 приводяться залежності  $\bar{Y}_{ПОВ}$ ,  $C_{v_{ПОВ}}$ ,  $C_{s_{ПОВ}}$  (індекс «ПОВ» у параметрів відповідає умовам зарегульованості стоку ставками або водосховищами або наявності зрошувальних меліорацій) від відносних площ  $f_v$  і  $f_{3P}$  (для трьох рівнів оптимального зволоження  $v_0 = 0,8; 0,9; 1,0$ ) при початковій нормі природного стоку  $\bar{Y}_{ПР} = 20$  мм.

Чим більше оптимальне зволоження ґрунту, необхідне для вирощування сільськогосподарських культур, тим більша кількість води повинна бути вилучена з місцевого стоку для цілей зрошування, що і викликає посилення інтенсивності зниження стоку.

Процес зрошування, як правило, супроводжується появою зворотних вод.

Зворотні води  $W_{3B}$  виникають за рахунок фільтраційних втрат з водопровідної і водорозподільної мережі, а в рівняннях водогосподарського балансу, які використовуються при моделюванні, враховуються у наступному вигляді

$$W_{3B} = F_{3P} \xi (1 - \eta) \frac{M_0}{\eta} \quad (5.19)$$

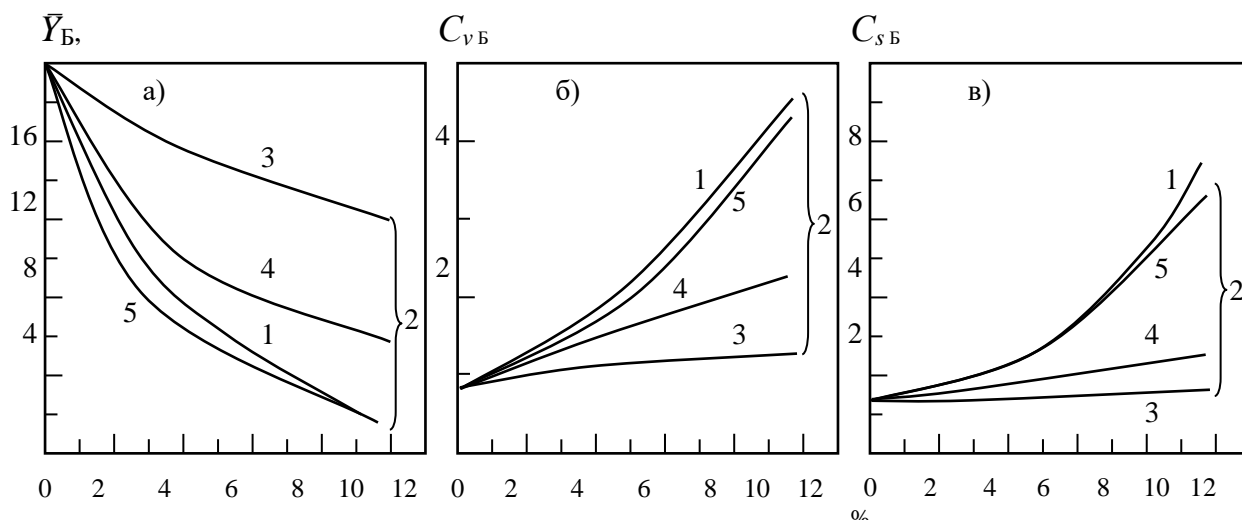


Рисунок 5.4 - Залежності середньобагаторічного шару стоку (а), коефіцієнтів варіації (б) і асиметрії (в) рядів побутового стоку від ступеня зарегульованості водозбору штучними водоймами (1) та площ зрошування (2) при різних рівнях  $v_0$ :  $v_0 = 0.8$  (3);  $v_0 = 0.9$  (4);  $v_0 = 1.0$  (5)

Коефіцієнт  $\xi$  задає частку надходження невикористаних при зрошуванні вод в поверхневі водотоки.

Його значення визначається положенням зрошуваних масивів щодо водоприймача (заплавні, терасні, привододільні масиви), а також наявністю або відсутністю гідрографічного зв'язку зони зрошування з цією річкою [36,37].

Для невеликих річок, які не дренують водоносні горизонти,  $\xi = 0$  і в процесі моделювання враховується тільки водозабір.

Результати чисельних експериментів дозволили встановити, що наявність зворотних вод сприяє збільшенню норм річного стоку і зменшує його мінливість в багаторічному перерізі. Коефіцієнт асиметрії при цьому змінюється незначно.

На рис.5.5 наведені залежності статистичних параметрів від відносної площі зрошуваних земель, отримані для передгірських річок Кримського півострова, які входять до складу зрошувальної системи Північно-Кримського каналу як водоскиди.

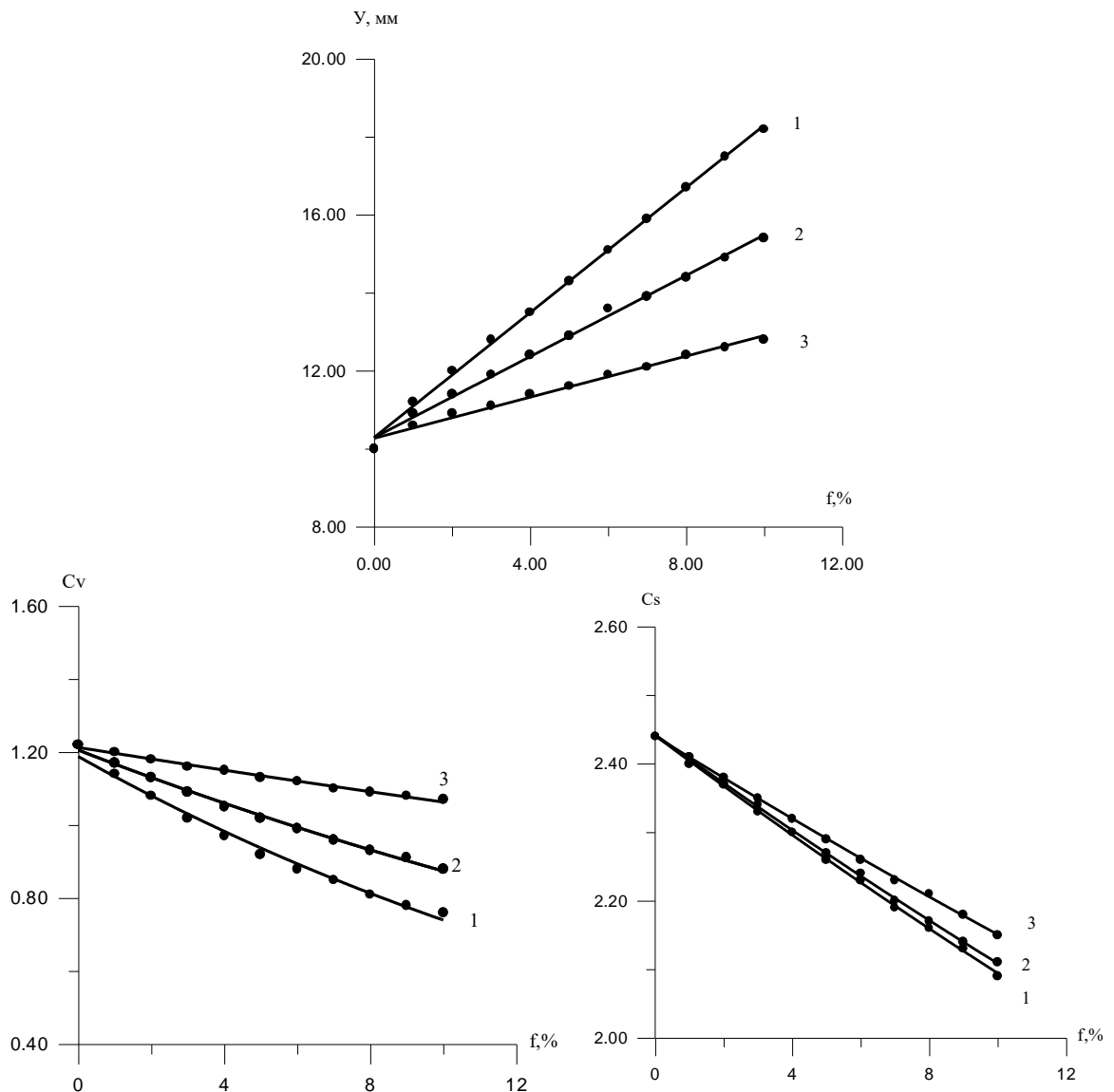


Рисунок 5.5 - Зміна статистичних параметрів річного стоку при наявності на водозборах сільськогосподарських масивів, які зрошуються за рахунок річок-донорів

### 5.2.3 Оцінка характеристик річного стоку в умовах господарської діяльності для річок зони недостатнього зволоження (на основі результатів стохастичного моделювання)

Оскільки в межах однієї і тієї ж річки кількісний і якісний склад антропогенних чинників може змінюватися, для можливості урахування їх

сумісного впливу запропоновано використовувати не самі статистичні параметри, а так звані коефіцієнти антропогенного впливу

$$k = \frac{A_f}{A_0}, \quad (5.20)$$

де  $A_0$  - значення того або іншого статистичного параметра річного стоку в природних умовах;

$A_f$  - значення параметра побутового стоку при заданих показниках антропогенного впливу.

Залежності  $k$  від показників рівня господарського освоєння території аналогічні залежностям для відповідних статистичних параметрів (див. рис.5.4): вплив антропогенних чинників посилюється у міру збільшення площ, які знаходяться під впливом водогосподарських перетворень, а також у міру переходу від зони достатнього зволоження до зони недостатнього. Непрямим показником загального зволоження території може служити норма природного річного стоку.

Результати моделювання представляються в графічному або табличному вигляді, в яких коефіцієнти антропогенного впливу  $k$  є функціями норм природного річного стоку, рівня господарського освоєння водозборів, заданого у вигляді відносних площ, які зазнають антропогенного впливу в сукупності з іншими не випадковими показниками (наприклад,  $\xi, \eta$ ).

Функції антропогенного впливу для передгірських річок Криму в умовах зрошування сільськогосподарських масивів водами Північно-Кримського каналу наведені на рис.5.6, з якого видно, що найбільші зміни стоку відбуваються при рівні оптимального зволоження  $V_0 = 1$ .

Зміни норм стоку, які перевищують границю точності розрахунків (10%), відповідають відносній площі зрошування, яка перевищує 1%. При зрошуванні кормових культур збільшення норм річного стоку, яке перевищує границю точності розрахунків, складає 2% для  $V_0 = 0.9$  и 3% - для  $V_0 = 0.8$ .

Безповоротні зміни водних ресурсів (>70%) спостерігаються при відносних площах зрошування, які перевищують 8% від загальної площі водозбору.

Таким чином, функції антропогенного впливу можуть бути використані для управління водними ресурсами.

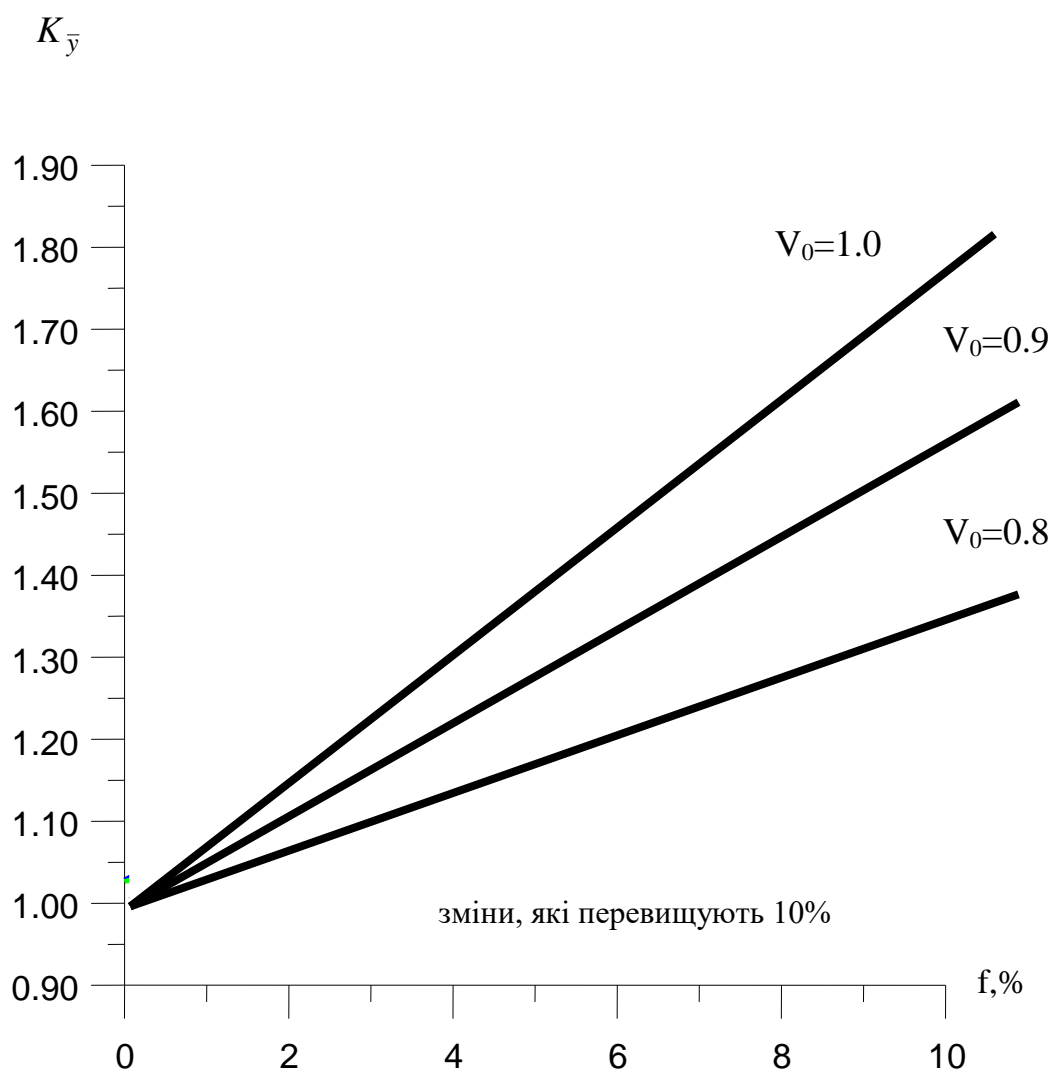


Рисунок 5.6 - Функція антропогенного впливу  $K_{\bar{y}}$ , отримана у результаті моделювання рядів річного стоку при зрошуванні сільськогосподарських масивів водами річки-донора

За наявності трьох вищеописаних чинників антропогенного впливу, найбільш характерних для зони недостатнього зволоження України, сумісний облік впливу чинників господарської діяльності виконувався на основі наступного виразу

$$A_f = A_0(k_1 + k_2 + k_3 - 2) ; \quad (5.21)$$

за наявності двох чинників господарської діяльності -

$$A_f = A_0(k_1 + k_2 - 1), \quad (5.22)$$



де  $k_1, k_2, k_3$  - коефіцієнти, які чисельно враховують антропогенний вплив на статистичні параметри річного стоку;

$k_1$  - коефіцієнт антропогенного впливу, який враховує вплив втрат на додаткове випаровування з водної поверхні штучних водойм;

$k_2$  - коефіцієнт антропогенного впливу, який враховує вплив забору води з місцевих водних ресурсів на зрошування;

$k_3$  - коефіцієнт антропогенного впливу, який враховує вплив скидів зворотних вод з сільськогосподарських масивів, що зрошуються за рахунок річки-донора.

Адекватність розрахункової моделі побутового стоку за даними спостережень була підтверджена відповідністю розрахункових і емпіричних кривих розподілу річного стоку річок з порушеним господарською діяльністю режимом.

Застосована в [17] детерміновано-стохастична модель річного стоку в умовах антропогенного впливу була використана для оцінки параметрів побутового стоку невивчених або маловивчених річок півдня України (зона недостатнього зволоження), де забір води з місцевого стоку для цілей зрошування у верхів'ях поєднується зі скиданням вод з водорозподільних каналів зрошувальних систем, а водосховища виконують роль штучних випарників [36,37].

У табл. 5.2–5.4 наведені результати розрахунків характеристик річного стоку річок північно-західного Причорномор'я в порушених господарською діяльністю умовах. Для визначення коефіцієнтів антропогенного впливу, обумовленого додатковим випаровуванням з водної поверхні штучних водоймищ, необхідно мати в своєму розпорядженні дані про відносну площу штучних водойм  $f_B$ , а також значення норм природного річного стоку, розрахованих за даними про кліматичний стік  $\bar{Y}_K$ . При розрахунках можна використовувати графічні, табличні й аналітичні залежності (5.12), (5.14) і (5.15).

Установлено, що не дивлячись на значне число ставків і водосховищ в південних регіонах, їх відносна площа в межах водозборів невелика ( $<1\%$ ), внаслідок чого поправкові коефіцієнти до норм річного природного стоку близькі до 1 (табл.5.1), а самі зміни знаходяться в межах 10%, тобто в межах точності розрахунків середнього арифметичного значення.

Проте, побутовий стік маловодних років при незначному зменшенні середнього арифметичного значення може виявитися близьким до нуля, що відбувається за рахунок зростання багаторічної мінливості і асиметрії річного стоку при наявності на водозборі штучних водойм (табл.5.3).

Крім забору води на зрошування з місцевих водних ресурсів на території північно-західного Причорномор'я у формуванні річного стоку

беруть участь також зворотні води з масивів, зрошуваних річками – донорами: Дунаєм і Дністром. На зміну водоймам, які виконують сезонне регулювання стоку за рахунок місцевих водних ресурсів, прийшли водосховища, які заповнюються дунайськими і дністровськими водами.

Деякі водотоки цілком постають частинами цих систем. Так, наприклад, р. Фонтанка (права притока р. Когільник) використовується як скидний канал колекторно-дренажних вод. Середньобогаторічна величина побутового стоку цієї річки перевищує 100мм, при нормі кліматичного стоку, що дорівнює 15мм. Результати імітаційних експериментів за даними до початку 90-х років минулого сторіччя (табл.5.4) показали, що скиди вод з масивів, зрошуваних річкою-донором, можуть збільшити удвічі стік малих річок.

Таблиця 5.2 - Зміна норм річного стоку північно-західного Причорномор'я за рахунок втрат на додаткове випаровування з поверхні штучних водойм

Назва річки	Норма природного стоку $\bar{Y}$ , мм	Відносна площа водної поверхні $f_B$ , %	Коефіцієнт впливу додаткового випаровування $k_{1,\bar{Y}} = \frac{\bar{Y}_{ПЮБ}}{\bar{Y}_{ПР}}$
Барабой	5,5	1,25	0,63
Сарата	6,9	0,34	0,88
Хаджидер	6,9	0,92	0,68
Алкалія	7,6	0,51	0,84
Малий Куяльник	8,6	0,38	0,88
Великий Куяльник	9,1	0,26	0,92
Тілігул	9,5	0,24	0,92
Когільник	10,5	0,53	0,95
Ягорлик	21,0	0,63	0,98
Кучурган	10,5	0,35	0,96
Савранка	28,9	0,72	0,99
Кодима	28,9	0,28	0,98
Чичиклея	7,8	0,33	0,98
Гнилий Єланець	10,5	0,52	0,97
Громоклея	6,6	0,67	0,96

Таблиця 5.3 - Статистичні параметри побутового стоку річок північно-західного Причорномор'я в маловодні ( $P = 75\%$ ) і багатоводні ( $P = 25\%$ ) роки при наявності на водозборах штучних водойм

Водозбори	Статистичні параметри побутового стоку			Шар побутового стоку (мм) із забезпеченістю	
	$\bar{Y}_{ПОВ}$ , мм	$C_V$	$C_S$	75%	25%
Барабой	4,2	2,6	3,6	0,0	25,3
Сарата	6,2	1,7	2,9	0,0	27,0
Хаджидер	4,7	2	3,4	0,0	22,9
Алкалія	6,3	1,9	2,8	0,0	30,2
М.Куяльник	1	2	2,2	0,0	5,0
В.Куяльник	11	1,7	2,2	0,0	48,8
Тілігул	12	1,5	2	0,0	48,0
Когільник	10	1,4	2,2	0,3	38,3
Ягорлик	27	0,51	2,2	17,5	54,8
Кучурган	13	1,12	2,1	2,7	42,1
Савранка	35,6	0,6	1	20,0	75,8
Кодима	20	0,9	1,4	6,9	55,1
Чичикля	9,7	0,8	1,3	4,0	24,6
Гнилий Єланець	2,3	0,8	1,3	0,9	5,8
Громокля	8,4	1,6	2,8	0,0	35,3

Додаткове надходження води забезпечує також “вирівнювання” розподілу стоку в маловодні і багатоводні роки, тобто призводить до зниження коефіцієнтів варіації річного стоку. Проте, роль забору води на зрошування з місцевих водних ресурсів не втратила своєї значущості для окремих річок. Так, наприклад, у верхів'ях річок Дунай-Дністровського межиріччя (на території Молдови) переважає зрошування за рахунок місцевих водних ресурсів, а в середній і нижній течіях – за рахунок річок-донорів.

У зв'язку з цим сумарний коефіцієнт антропогенного впливу визначався по (5.21), тобто з урахуванням того чи іншого виду зрошування (за рахунок місцевого стоку або за рахунок річки-донора), а також характеристик зрошувальних систем у верхній і нижній течіях.

Таблиця 5.4 - Зміна статистичних параметрів річного стоку (норми і коефіцієнтів варіації ) річок північно-західного Причорномор'я при зрошуванні водами річки-донора

Назва річки	Параметри природного стоку			Відносні площі зрошуваних масивів	Коефіцієнти впливу зворотних вод	
	$\bar{Y}$ , мм	$C_v$	$C_s$		$f_{зр, \bar{Y}}$ %	$k_{3, \bar{Y}}$
Сарата	6,90	1,7	2,9	6,2	2,8	0,35
Хаджидер	6,90	1,7	2,9	13,9	2,0	0,50
Алкалія	7,60	1,6	2,7	26,8	2,4	0,43
Малий Куяльник	11,0	1,2	2,1	10,6	1,6	0,48
Великий Куяльник	12,0	1,2	2,1	3,4	1,2	0,86
Кучурган	14,0	1,09	1,9	0,32	1,05	0,99
Когільник	14,0	1,09	1,8	1,3	1,2	0,83

Наприклад, для річок Дунай-Дністровського межиріччя в межах Молдови переважає забір води для цілей зрошування з місцевого стоку при коефіцієнті корисної дії зрошувальних систем, що дорівнює 0,9, і переважно привододільному і терасовому положенні зрошуваних масивів, тоді як в нижній течії річок великі площі відведені під зрошування за рахунок річки-донора (р. Дунай), при цьому змінюються і характеристики зрошувальних систем ( $\eta = 0,75$ ), і положення зрошуваних масивів (заплавні).

Приклад розрахунку характеристик побутового стоку для невивченої річки з різними показниками господарської діяльності наведений в табл.5.5.

Таблиця 5.5 – Характеристики річного стоку річки Великий Катлабух при різних чинниках антропогенного впливу

Стан водного об'єкта	Статистичні параметри річного стоку			Відносна площа водної поверхні, %	Відносна площа земель, які зрошуються водами річки-донора, %	Річний стік заданної забезпеченості ( $P\%$ ), мм				
	Y, мм	Cv	Cs			$f_B$	$f_{3P}$	5	25	50
Природний	9.2	1.41	2.40	0.0	0.0	35	13	4.6	0.5	0.0
При наявності штучних водойм	6.1	1.89	3.20	1.3	0.0	29	7.8	1.4	0.0	0.0
При наявності зрошування за рахунок річки-донора	11.0	1.24	2.40	0.0	4.5	39	15	6.1	1.6	0.0
Сумарний результат	7.9	1.72	3.2	1.3	4.5	34	9.9	2.3	0.0	0.0

### 5.3. Математичне моделювання побутового річного стоку на водозборах річок зони надмірного зволоження

Західна частина Українського Полісся відноситься до регіонів з порушеною екосистемою. Широкомасштабна меліорація (осушення боліт) регульованим стоком невеликих річок була розпочата ще в 60-і роки минулого сторіччя і призвела до різкої зміни умов формування складу і якості води в річках, а також до порушення гідрологічного режиму ґрунтових вод на прилеглих територіях, які в 2 рази перевищують поверхню цільового осушування [17].

#### 5.3.1 Огляд методів кількісної оцінки впливу осушування боліт на річний стік річок

Відносно впливу осушування боліт на стік річок існують різні думки (І.І.Жилінський,1899; В.В.Докучаєв,1876; Н.Н.Нікітин,1900; В.В.Оппоков, 1909;А.Г.Булавко,1961; С.М.Новіков,1978; С.І.Харченко,1978;В.Ф.Шебеко, 1983 і ін.). При дослідженні змін стоку на меліорованих водозборах найбільш схожі висновки відносно змін меженного стоку: осушування боліт сприяє вирівнюванню внутрішньорічного розподілу стоку – збільшенню мінімальних витрат стоку за меженні періоди, що пояснюється поліпшенням процесу стікання поверхневих і підземних вод в перетворену водогосподарською діяльністю руслову мережу. Більшість одержаних висновків про вплив осушування боліт заснована на зіставленні спостережень за стоком меліорованих водозборів і водозборів в природних умовах. Достовірність таких результатів визначається тривалістю спостережень за стоком в період природного і порушеного господарською діяльністю водного режиму. Другий напрям робіт по дослідженню впливу осушування боліт на стік річок пов'язаний з диференційованим підходом до чинників, які визначають стік з осушених водозборів. При цьому розглядається рівняння водогосподарського балансу водозбору, згідно з яким зміну об'єму річного стоку під впливом осушування боліт  $\Delta W_{OC}$  можна представити таким чином [36, 37]

$$\Delta W_{OC} = \Delta W_{СП} \pm \Delta W_{ЕБ} , \quad (5.23)$$

де  $\Delta W_{СП}$  - об'єм спрацювання вікових запасів підземних вод;

$\Delta W_{ЕБ}$  - зміни об'єму води за рахунок зміни випаровування на водозборі в результаті осушування боліт.

У методі ДГІ (Російському державному гідрологічному інституті) об'єм спрацювання ( $\text{км}^3$ ) представляється у вигляді

$$\Delta W_{СП} = (Y_{ПОВ,ОС} F_{Б,ОС} + Y_{сух} F_{сух}) \cdot 10^{-3}, \quad (5.24)$$

де  $Y_{ПОВ,ОС}$  - шар спрацюваності болотних вод, м;

$Y_{сух}$  - шар спрацюваності з осушених суходолів, м;

$F_{Б,ОС}$  - площа осушених боліт і заболочених земель,  $\text{км}^2$ ;

$F_{сух}$  - площа суходолів, які примикають до боліт і знаходяться під впливом осушування,  $\text{км}^2$ .

Шар води, який спрацьовується в руслову мережу в результаті зменшення вологості верхнього шару ґрунтів при зниженні рівня води на болоті, визначається за виразом

$$Y_{\Delta H} = \Delta H_{ОС} (w_{np} - w_{oc}); \quad (5.25)$$

$$\Delta H_{ОС} = (H_{ПР} - H_{ОС}), \quad (5.26)$$

де  $H_{ПР}, H_{ОС}$  - середні річні рівні болотних вод відповідно до і після осушування, м;

$\Delta H_{ОС}$  - зміна середнього рівня ґрунтових вод в результаті осушування, м;

$w_{oc}$  - об'ємна вологість торф'яного шару після осушування.

Шар води, який надходить до зрошувальної системи з прилеглих суходолів, по цій же методиці оцінюється таким чином

$$Y_{сух} = \Delta H_{сух} \mu, \quad (5.27)$$

де  $\Delta H_{сух}$  - середнє значення зниження рівня ґрунтових вод на прилеглих ділянках суходолу, м;

$\mu$  - коефіцієнт водовіддачі ґрунтів.

Найбільші труднощі виникають при визначенні  $\Delta H_{сух}$  і  $F_{сух}$ . Площа суходолів, що знаходяться під впливом осушування, перш за все визначається шириною зони впливу меліорацій на навколишні території, яка залежно від гідрогеологічних умов змінюється в широких межах. Узагальнення даних, виконане В.Ф. Шебеко (1983), дозволяє визначити ширину зони впливу осушування при зниженні рівня води на болоті на 1,0-

1,5 м для різних ґрунтів прилеглих територій. Вказані дані можуть бути застосовані для визначення наближених значень  $\Delta H_{\text{сух}}$  і  $F_{\text{сух}}$ .

Об'єм втрат води за рахунок зміни випаровування в басейні визначається за формулою

$$\Delta W_{EB} = (E_{\text{БОЛОТ}} - E_{\text{ОС}}) F_{\text{ОС}} \cdot 10^{-6}, \quad (5.28)$$

де  $E_{\text{БОЛОТ}}$  - випаровування з неосушеного болота, мм;

$E_{\text{ОС}}$  - випаровування з осушеного болота, мм;

$F_{\text{ОС}}$  - площа осушених земель, км<sup>2</sup>.

На ранніх етапах дослідження впливу осушування боліт на стік річок допускалося, що спрацювання запасів підземних вод відбувається тільки в перші роки після проведення осушувальних меліорацій. Як показали дослідження П.О. Кисельова (1964), С.М. Новікова (1976), М.О. Мішустіна (1979), Я.О. Мольчака (1999), за наявності напірних вод стабілізації рівня ґрунтових вод в процесі осушування не настає, оскільки спостерігається додатковий приплив напірних міжпластових вод у верхні водоносні горизонти ґрунтових вод. Саме з цієї причини проведення меліоративних робіт в зоні боліт Поліської низовини призвело до зміни взаємозв'язку річкових і підземних вод в системі річкова мережа - підземний приплив. За даними Я.О. Мольчака (1999), практично на всіх річкових системах з осушувальними меліораціями спостерігається збільшення річного стоку на 21-114%. При цьому домінуюча роль належить підземній складовій.

Роль зміни випаровування  $\Delta W_{EB}$  з поверхні суші при проведенні осушувальних робіт у балансовому рівнянні (5.23) визначається характером освоєння осушених боліт. Безпосередній вплив осушування і освоєння боліт на процеси випаровування пов'язаний із зменшенням вологості торф'яного ґрунту після зниження рівня ґрунтових вод, зміною виду рослинності і зміною здатності транспірації. Доцільно розглянути випаровування на болотах неосушених, осушених, але не освоєних, і на освоєних під посіви сільськогосподарських культур. За величину випаровування з неосушених торф'яників багатьма авторами бралось випаровування з водної поверхні. За весь період вегетації сумарне випаровування з торф'яників і водної поверхні відрізняється мало, але в період посиленої транспірації випаровування з торф'яників вище, ніж випаровування з водної поверхні.

Згідно з дослідженнями В.Ф. Шебеко (1980), тільки весною внаслідок достатнього зволоження ґрунтів і незначного транспірування, випаровування з осушених і неосушених боліт практично не змінюється. У літній період місячне випаровування на осушених, але не освоєних болотах, і його сума за теплий період (квітень – вересень) зменшується в



середньому на 10-15%, а в особливо посушливі місяці і при глибоких ґрунтових водах – на 40%. Це пов'язано не тільки з пониженням ґрунтових вод при осушуванні, але й під впливом зменшення транспірації прорідженого, виродженого травостою. При освоєнні осушених боліт випаровування знову зростає за рахунок транспірації вологи сільськогосподарськими культурами. Слід зазначити, що сільськогосподарські культури при своєму розвитку споживають набагато більше вологи, ніж природна болотна рослинність. Так, для природної болотної рослинності транспірація складає 10-25% від сумарного випаровування, а для культурних трав – 90%. Випаровування на освоєних болотах залежить від ступеня регулювання ґрунтової вологи гідромеліоративними заходами. Якщо система осушування передбачає поповнення запасів вологи зони аерації до максимального рівноважного вологоутримання, то випаровування з поверхні осушених боліт збільшується, в порівнянні з неосушеними болотами, для культур з підвищеним оптимальним зволоженням ґрунту (трави, овес, ячмінь, конопля). Якщо ж оптимальне зволоження, необхідне для розвитку даної культури, невелике (буряк, картопля), то випаровування з осушених і освоєних боліт може виявитися навіть меншим, ніж випаровування з неосушеного болота. Якщо осушування глибоке і додаткового зволоження у вегетаційний період немає, то випаровування з поверхні освоєного болота зменшуватиметься в порівнянні з неосушеним.

Зміна шару річного стоку в результаті осушення боліт, згідно з (5.28), буде дорівнювати різниці між випаровуванням з боліт і випаровуванням з осушеної і освоєної території.

Виходячи з наведених вище міркувань, для урахування впливу заболоченості на величину випаровуваності М.О. Мішустінім (1979) була запропонована формула

$$\bar{Y}_{f_B} = Y_{f_B=0} - (\bar{E}_B - \alpha_1 \bar{E}_C) f_B \quad (5.29)$$

або

$$k_B = 1 - k_{3B} \frac{F_B}{F} \quad (5.30)$$

де  $k_B = \frac{\bar{Y}_{f_B}}{\bar{Y}_{f_B=0}}$  - коефіцієнт, який дорівнює відношенню норми стоку із

заболоченого водозбору з площею боліт  $F_B$ , до норми стоку з незаболоченого басейну  $F$ ;

$f_B$  - відносна площа, зайнята болотами.

$k_{3B}$  - коефіцієнт надмірного випаровування, який дорівнює

$$k_{3B} = \frac{\bar{E}_B - \alpha_1 \bar{E}_C}{\bar{Y}_{f_B=0}} \quad (5.31)$$

$i$  є характеристикою надмірного випаровування з поверхні суші;

$\bar{E}_B$  - середньобогаторічне випаровування з боліт;

$\bar{E}_C$  - середньобогаторічне випаровування з суходолів;

$\alpha_1$  - коефіцієнт, який залежить від сівообороту освоєваної території;

$\bar{Y}_{f_B=0}$  - середньобогаторічний шар стоку з незаболочених територій.

Відповідно до одержаних результатів, М.О. Мішустін дійшов до висновку, що  $k_B$  може бути як більше, так і менше 1, залежно від виду сільськогосподарських культур на освоєних болотах. Одержаний результат підтверджує висновок К.Є.Іванова (1953) про те, що норма випаровування річного стоку в результаті осушувальної меліорації не зміниться при використуванні осушених боліт під інтенсивне сільськогосподарське виробництво з високою агротехнікою і зарегульованим водним режимом.

Описані вище результати і висновки багатьох авторів дозволяють підсумувати, що осушувальні меліорації сприяють збільшенню середньорічного стоку. Це відбувається за рахунок спрацювання запасів ґрунтових вод, збільшення щільності річкової мережі, зменшення тривалості затоплення заплав і деяких інших чинників.

### **5.3.2 Обґрунтування стохастичної моделі, яка описує вплив осушувальних меліорацій на річний стік**

Порівняння величин підземного припливу в річки, розрахованих наближеним методом (як напівсума модулів стоку 30-ти добових мінімумів за літню і зимову межень), і за генетичною формулою А.М.Бефані показало [17], що норми підземного припливу в річки Поліської низовини, визначені генетичним методом, нижчі за величини підземного стоку, розраховані наближеним методом (рис.5.7). Це можна пояснити підвищенням дренажного стоку з водозборів, обумовленого роботою осушувальних систем одностороннього регулювання. На рис.5.8 виразно показаний часовий тренд (направлена зміна в часі) в коливаннях підземного стоку з водозбору р. Вирка – с.Сварині. Початок тренда припадає на 60-і роки минулого сторіччя, коефіцієнт лінійної кореляції підземного стоку з часом дорівнює 0,82.

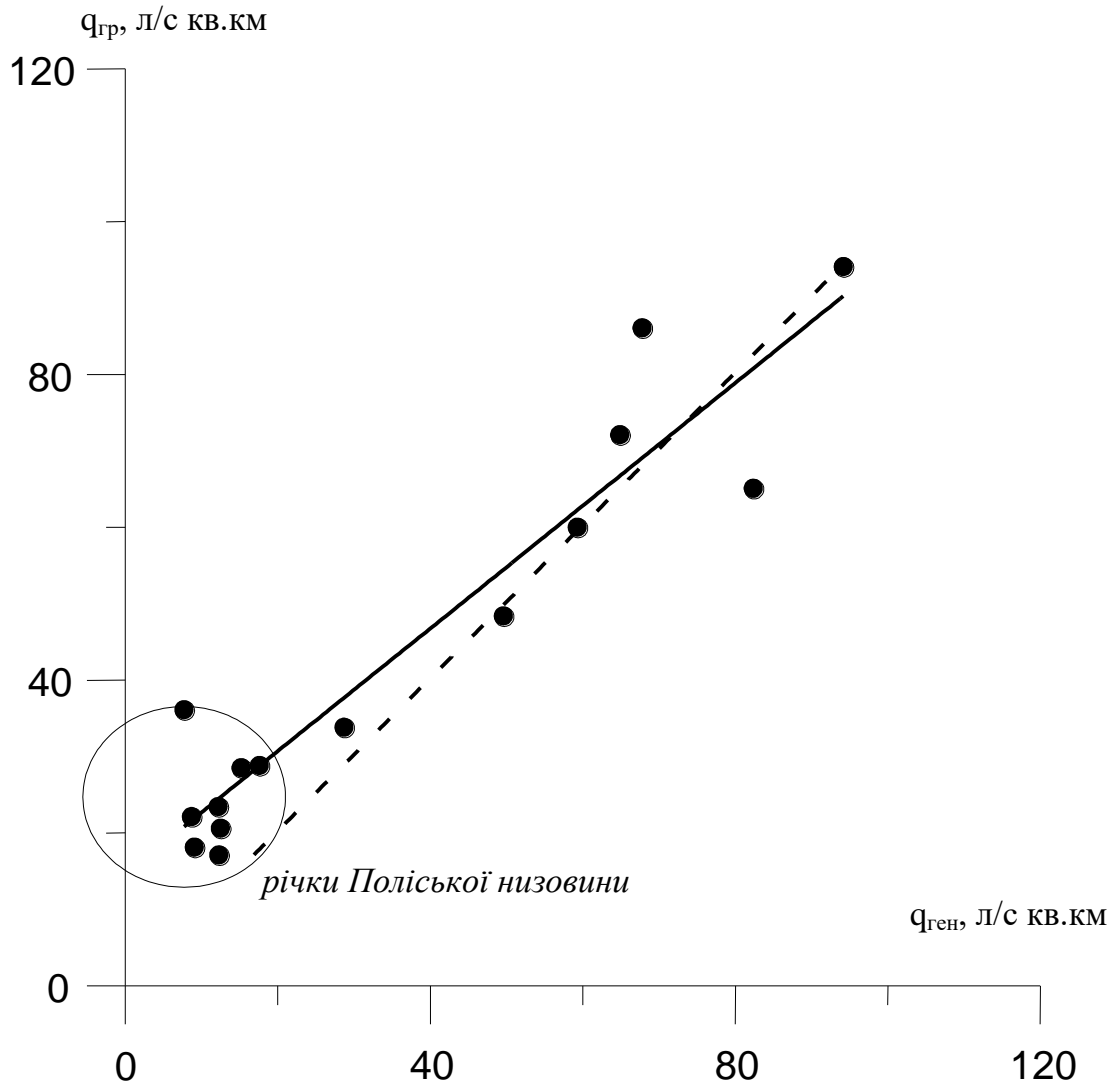


Рисунок 5.7 – Порівняння норм підземного припливу в річки, розрахованих за генетичним ( $q_{ген}$ ) та за наближеним методами ( $q_{гр}$ )  
 - - - лінія рівних значень

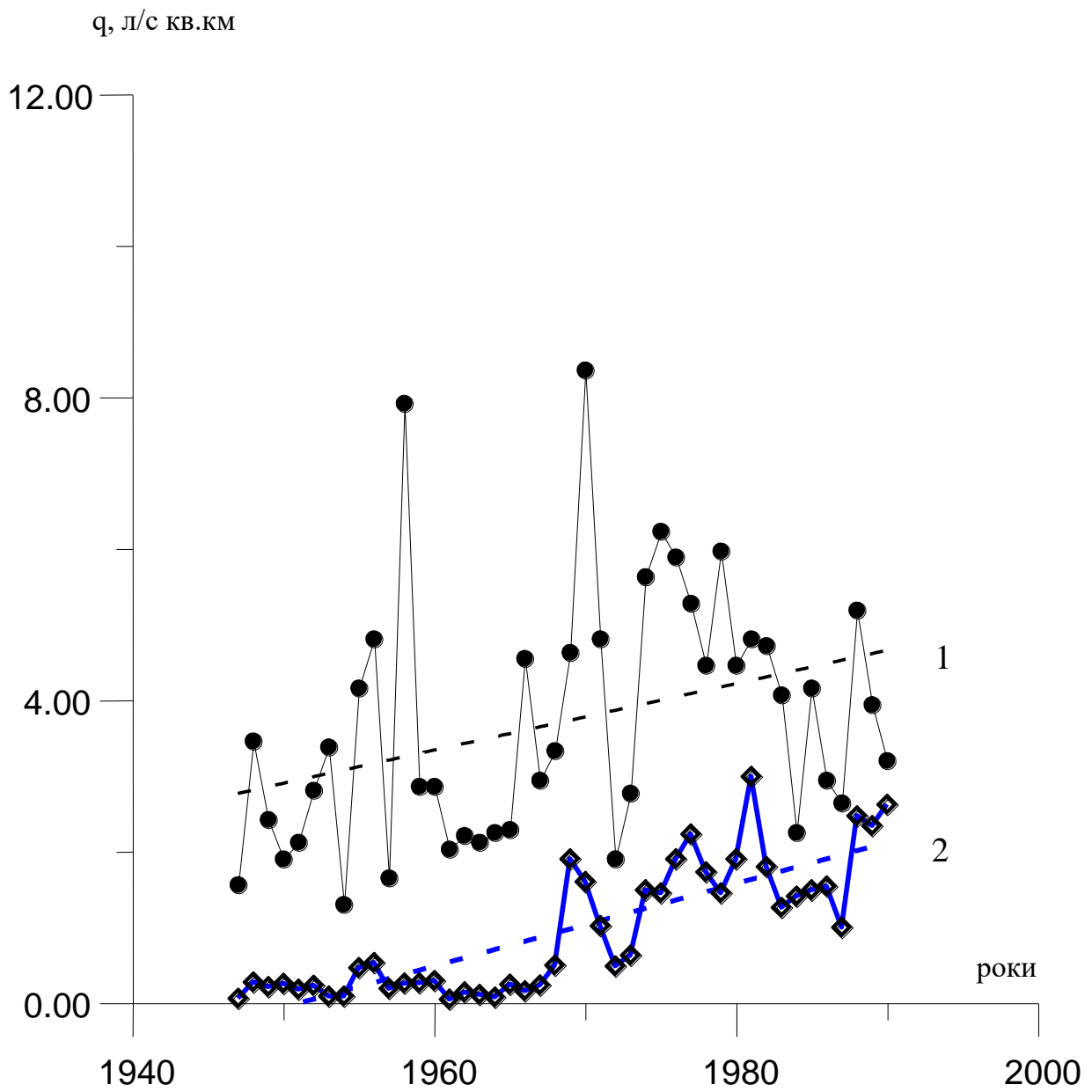


Рисунок 5.8 – Суміщені хронологічні графіки річного (1) і підземного (2) стоку з водозбору р.Вирка – с.Сварині

- - - часовий тренд

За фондовими даними ця річка характеризується наявністю значного дренажного стоку з осушуваних територій і входить до складу меліоративної осушувальної системи. Таким чином, висновок про те, що саме зміна підземного припливу в річки обумовлює зростання сумарного річного стоку річок західної частини Українського Полісся при осушуванні боліт, знайшло своє підтвердження.

Як вже наголошувалося, складову водогосподарського балансу при осушуванні боліт на водозборах  $\Delta W_{CP}$  можна представити у вигляді трьох доданків. Перший обумовлений безпосередньо зниженням рівня ґрунтових вод, другий – втратами вологи за рахунок зменшення вологості верхнього шару покладу при зниженні рівня води на болоті, третій – змінами випаровування на осушених водозборах.

На прикладі західної частини Українського Полісся встановлено, що в зоні надмірного і достатнього зволоження роль додаткових втрат на випаровування з поверхні боліт невелика і в більшості випадків знаходиться в межах точності розрахунку середньобогаторічних величин річного стоку.

У зв'язку з цим прийнято рішення розглядати рівняння водогосподарського балансу тільки з урахуванням перших двох складових.

Рівняння водогосподарського балансу в умовах осушування боліт, покладене в основу детерміновано-стохастичної моделі, може бути представлене у вигляді

$$Y_{OC} = Y_{IP} + \mu \Delta H_{GP} f_{OC} + (w_{IP} - w_{OC}) \Delta H_{GP} f_{OC}, \quad (5.32)$$

де  $Y_{OC}$  - шар стоку з водозбору, на якому проводяться заходи щодо осушування боліт;

$Y_{IP}$  - шар стоку з боліт в непорушених господарською діяльністю умовах (природний);

$\mu$  - коефіцієнт водовіддачі ґрунтів, який оцінюється як різниця між повною і найменшою вологомісткістю, поділеною на товщину даного шару;

$\Delta H_{GP}$  - зміна середнього рівня ґрунтових вод в результаті осушування;

$f_{OC}$  - показник меліорованості території;

$w_{IP}$  - об'ємна вологість торф'яного покладу до осушування;

$w_{OC}$  - об'ємна вологість торф'яного покладу після осушування.

Середньобогаторічна величина  $\Delta H_{GP}$  визначалася відповідно до прийнятих норм осушування мінеральних ґрунтів і торф'яників (табл.5.7,5.8). При цьому під нормою осушування розумівся шар ґрунту,

звільнений від надлишку води і аерований до рівня, який забезпечує нормальний розвиток рослин.

Таблиця 5.7 - Норми осушування мінеральних ґрунтів, мм

Культури	Передпосівний період	Вегетаційний період	
		піщані та супіщані ґрунти	суглинисті та глинисті ґрунти
Лугові трави	0,30-0,40	0,60-0,70	0,70-0,80
Трави на пасовищі	0,40-0,50	0,70-0,75	0,75-0,85
Зернові	0,40-0,50	0,70-0,80	0,80-0,90
Пропашні, овочеві, технічні	0,50-0,60	0,80-0,90	0,90-1,00

Таблиця 5.8 - Норми осушування торф'яників, мм

Культури	Передпосівний період	Вегетаційний період	
		піщані та супіщані ґрунти	суглинисті та глинисті ґрунти
Лугові трави	0,40-0,50	0,70-0,80	0,80-1,00
Трави на пасовищі	0,50-0,60	0,70-0,80	0,55-1,00
Зернові	0,50-0,60	0,80-0,90	1,10-1,20
Пропашні, овочеві, технічні	0,60-0,70	0,90-1,00	1,20-1,30

Слід зазначити, що норми осушування залежать від біологічних особливостей рослин, водно-фізичних властивостей ґрунту, фази розвитку рослин і пори року. Аерований шар ґрунтів приймають таким, щоб коренева система рослини розвивалася на нормальну глибину. Норми осушування визначаються також і загальною зволоженістю конкретного року. У багатоводні роки ( $P < 25\%$ ) норми осушування на 20-30 см більші, ніж в середні, а в посушливі ( $P > 75\%$ ) – на 10-30 см менше. Крім того, норми осушування різні в передпосівний і вегетаційний періоди. Для болотних ґрунтів, особливо торф'яників, які глибоко залягають, норми осушення більші, ніж для мінеральних, характерних для Житомирського Полісся. У свою чергу, для важких мінеральних ґрунтів з високим капілярним рівнем підняття норми осушування вищі, ніж для легких, які не мають цієї властивості. Всі ці обставини повинні бути враховані в розрахунках по (5.32).

Як вже наголошувалося, зміни підземного а, отже, і сумарного річного стоку визначаються також зниженням запасів вологи в ґрунтах  $w_E - w_{OC}$ . По регіону Західного Полісся середня об'ємна вологість боліт і заболочених територій в природних умовах може бути прийнята рівною 88%, а її характеристика після осушування і освоєння боліт визначається видом вирощуваної сільськогосподарської культури. Наприклад, допустима об'ємна вологість ґрунту для зернових культур складає 40-50%, для овочевих - 70-80%, картоплі і коренеплодів – 60-70%, польових багаторічних трав – 70-80%, лугових трав – 80-90% від повної вологомісткості. Переважаюча частина меліорованих земель використовується в сільськогосподарському виробництві, при цьому на потужних і середніх торф'яниках багаторічні трави займають приблизно 50%, а на дрібних торф'яниках - 63-75%, тобто  $w_E - w_{OC} = 3\%$  для більшої частини осушених і освоєних боліт даної території.

На базі цієї схеми створена імітаційна стохастична модель, яка дозволяє визначати ймовірнісні характеристики побутового стоку осушуваних водозборів.

Як вихідні статистичні параметри природного річного стоку бралися результати просторових узагальнень, виконані на базі рівняння водно-теплового балансу [17]. З метою урахування при моделюванні рядів побутового стоку відповідності стоку і норм осушування по забезпеченості, використовувалися групи маловодних (забезпеченість стоку  $P > 75\%$ ), багатоводних ( $P < 25\%$ ) і середніх по водності ( $25 \leq P \leq 75\%$ ) років.

Рівняння водогосподарського балансу (5.32) було представлено у ймовірнісній формі

$$Y_{OC,P} = Y_{PP,P} + \mu \Delta H_{GP,P} f_{OC} + (w_{PP} - w_{OC}) \Delta H_{GP,P} f_{OC}, \quad (5.33)$$

де індекс  $P$  - забезпеченість річного стоку, %.

При цьому приймається, що величини  $Y_E$  і  $\Delta H_{GP}$  підпорядковуються одному і тому ж закону розподілу - трипараметричному гама-розподілу С.М. Крицького – М.Ф. Менкеля.

При постійній динаміці площ осушуваних земель ряди річного і підземного стоку є статистично неоднорідними (див. рис.5.7). Іншими словами, гіпотеза про стаціонарність і ергодичність коливань річного стоку в даному випадку повинна бути знехтувана, а, відповідно, розрахункові характеристики, визначені по вихідних рядах, признаються статистично недостовірними.

Проте, при припущенні, що не випадкові показники антропогенного впливу залишаються незмінними в часі, моделювання побутового стоку на основі рівняння (5.33) дозволяє прийняти гіпотезу про стаціонарність і ергодичність коливань річного стоку, але лише для заданого рівня водогосподарських перетворень на водозборі, який в (5.32) визначається постійними кількісними показниками  $\mu, f_{OC}, w_{PP} - w_{OC}$ . Генерація реалізацій побутового стоку по (5.33) виконувалася для різних показників меліорованості  $f_{OC}$  (у діапазоні від 10 до 50% від загальної площі водозбору) і різних сільськогосподарських культур  $w_{PP} - w_{OC}$ .

В результаті одержані залежності статистичних параметрів побутового стоку від відносних площ осушування для різних видів сільськогосподарських культур, які для практичного застосування представлені у вигляді поправкових коефіцієнтів (коефіцієнтів антропогенного впливу) і дозволяють здійснити перехід від параметрів природного стоку до параметрів побутового (5.20).

В даному випадку розглядався побутовий стік при осушуванні в передпосівний період, оцінюваний поправковим коефіцієнтом  $k_1 = f(f_{OC}, \mu)$ ; побутовий стік при осушуванні у вегетаційний період -  $k_2 = f(f_{OC}, \mu)$  і побутовий стік при зміні вологості ґрунту за рахунок осушування -  $k_3 = f(f_{OC}, w_{PP} - w_{OC})$ . За наявності трьох чинників антропогенного впливу сумісний їх облік виконується за допомогою підсумовування поправкових коефіцієнтів.

Одержані результати показують, що при збільшенні площ осушення відбувається збільшення середньобаторічних величин річного стоку (рис.5.8), а коефіцієнти варіації і асиметрії зменшуються, причому інтенсивність зміни стоку посилюється у міру зниження загальної водності водозбору.



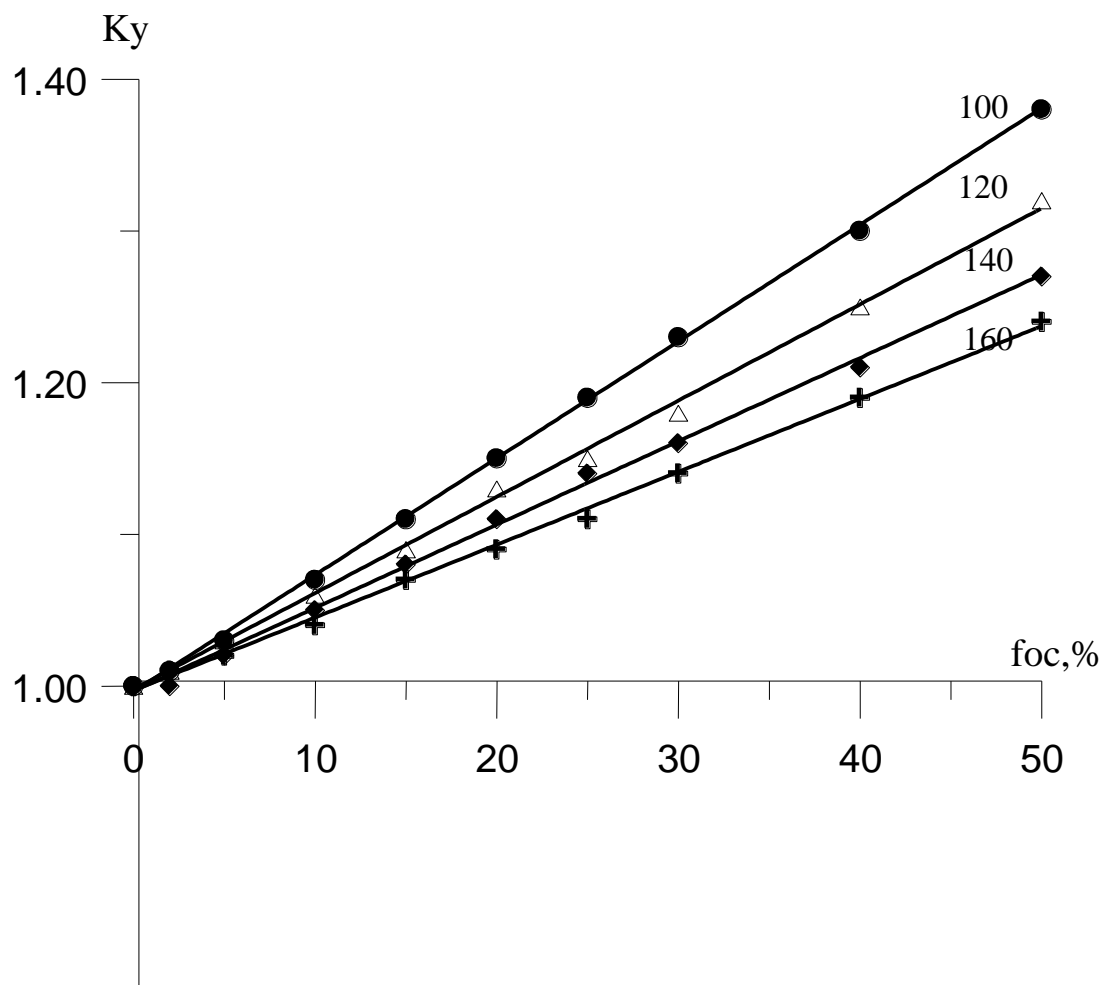


Рисунок 5.9 – Залежності для визначення поправкових коефіцієнтів до норми природного річного стоку, які враховують зниження рівня підземних вод від відносної площі меліорованості водозборів (Житомирське Полісся); біля ліній - норми кліматичного стоку, мм

Дослідження впливу осушувальних меліорацій на водні ресурси західної частини Українського Полісся проводилися багатьма авторами, але в роботі О.З. Ревери (1988) на основі даних спостережень одержані залежності вигляду  $k_{3M} = af_{OC} + b$  (для Житомирського і Волинського Полісся), в яких коефіцієнт зміни стоку  $k_{3M}$  за своїм фізичним значенням відповідає коефіцієнту антропогенного впливу  $K_{\bar{Y},OC}$ , одержаному в результаті стохастичного моделювання.

Зіставлення коефіцієнтів, наведених в літературних джерелах, а також одержаних авторами в результаті моделювання при відповідному показнику меліорованості  $f_{OC}$ , показало задовільні результати (табл.5.9).

Відхилення середніх багаторічних значень стоку в умовах осушувальних меліорацій, розрахованих за результатами моделювання, від  $k_{3M}$  знаходиться в межах  $\pm 10\%$ .

Перевагою запропонованої методики є можливість оцінки впливу осушувальних меліорацій не тільки на середній багаторічний стік, але й на стік різної забезпеченості.

На закінчення можна відзначити, що західна частина Українського Полісся відноситься до регіону із змінними кліматичними чинниками і необхідність осушення земель в значній мірі залежить від виду сільськогосподарських культур, які вирощуються на осушених болотах. Установлено, що основним чинником впливу осушення на річний стік є спрацювання запасів підземних вод, що призводить до збільшення норм сумарного річного стоку в цілому і знижує багаторічну нерівномірність розподілу стоку і асиметрію. Істотний вплив осушування спостерігається при площах, які перебільшують 10%.

Таблиця 5.9 – Порівняння результатів оцінки впливу осушування на річний стік річок за літературними даними та по результатах моделювання побутового стоку

№	Река - пункт	$\bar{Y}_K$ , мм	$f_{OC}$ , %	$k_{\bar{Y},OC}$	$k_{ЗМ}$	$\Delta k$
1	р.Виживка - пгт.Стара Виживка	120	7	1,06	1,07	0,01
2	р.Турія - с.Ягідне	120	9	1,04	1,03	-0,01
3	р.Турія - м.Ковель	115	9	1,09	1,08	-0,01
4	р.Турія - с.Бузаки	100	15	1,15	1,15	0,00
5	р.Стоход - с.Гулівка	120	21	1,20	1,32	0,12
6	р.Стоход - м.Любешов	135	17	1,20	1,30	0,10
7	р.Стир' - с.Шуровці	143	15	1,08	1,10	0,02
8	р.Стир' - м.Луцьк	140	8	1,09	1,03	-0,03
9	р.Стир' - с.Млинок	125	6	1,01	0,95	-0,06
10	р.Іква - с.Радянське (Млиновці)	150	3	0,99	1,01	0,03
11	р.Горинь -м.Ямполь	140	5	0,95	1,02	0,07
12	р.Горинь - с.Ожежин	140	6	1,02	1,06	0,04
13	р.Горинь - с.Деражне	145	7	1,02	1,02	0,00
14	р.Вирка -с.Сварині	155	14	1,12	1,20	0,08
15	р.Бережанка - с. Підлісне (Рудня)	160	11	1,09	1,06	-0,03
16	р.Тня - с.Броніки	92	26	1,33	1,25	-0,08
17	р.Уборть - с.Рудня Иванівська	97	24	1,24	1,34	0,10
18	р.Уборть - с.Перга	125	17	1,15	1,23	0,08
19	р.Уж – м.Коростень	97	9	1,10	1,15	0,05
20	р.Уж -м.Поліське	100	9	1,09	1,15	0,05
21	р.Жерев - с.В'язовка	120	6	1,08	1,05	-0,03
22	р.Грезля - с.Давидки	110	4	1,01	1,08	0,07
23	р.Тетерев - с.Троща	110	1	1,00	1,00	0,00
24	р.Тетерев - м.Житомир	80	7	1,05	1,09	0,04

## Продовження таблиці 5.9

25	р.Гнилоп'ять - с.Головинка	97	6	1,04	1,07	0,03
26	р.Гуйва - с.Городковка	100	2	1,00	1,04	0,04
27	р.Ірша -м. Володарськ Волинський	95	2	1,00	0,98	-0,02
28	р.Ірша – с.Українка	89	3	1,01	1,00	0,01

## 5.4 Приклади

### 5.4.1 Стохастичне моделювання рядів річного стоку

Стохастичне моделювання – це генерація штучних рядів річного стоку на базі статистичних параметрів, визначених за даними спостережень за стоком.

Математичним апаратом, який дозволяє виконувати моделювання, є метод Монте-Карло. За допомогою цього методу можна вирішити задачу оцінки статистичних параметрів у природних і порушених водогосподарською діяльністю умовах шляхом багаторазових випадкових випробувань.

#### Модель I (моделювання сукупності незалежних випадкових величин)

Моделювання здійснюється за одновимірним законом розподілу випадкових величин, виходячи з того, що інтегральна функція є

$$F(x) = \int_{-\infty}^x f(x) dx, \quad (5.34)$$

а забезпеченість

$$P(x) = \int_x^{\infty} f(x) dx. \quad (5.35)$$

Значення інтегральної функції розподілу чи функції забезпеченості випадкової величини прирівнюються до випадкового числа

$$\delta = \int_x^{\infty} f(x) dx \quad \text{або} \quad \delta = \int_{-\infty}^x f(x) dx$$

Задача моделювання полягає у тому, щоб з рівнянь (5.34) і (5.35) знайти випадкову величину  $x$ . Можливі два шляхи:

1. Перший полягає у представленні інтегральної функції розподілу у вигляді полінома.
2. Другий шлях полягає у використанні таблиць тих чи інших законів розподілу.

На практиці частіше за все застосовується другий підхід.

### Алгоритм вирішення поставленої задачі

1. Задаються вихідні параметри генеральної сукупності  $m_x, C_V, C_S$  та відношення  $C_S / C_V$ .
2. Задається закон розподілу (у цьому прикладі це закон Пірсона III).
3. Генерація випадкових чисел виконується за допомогою таблиці випадкових чисел, з якої вони вибираються за будь-якою схемою (наприклад, кожне п'яте число).
4. Обчислюються забезпеченості як  $P_i = \delta_i \cdot 100$ .
5. За таблицею визначаються числа Фостера  $\Phi_{p, C_S}$

$$\Phi_{p, C_S} = \frac{x_p - m_x}{\sigma_x} = \frac{k_p - 1}{C_V}$$

6. Розраховуються значення річного стоку  $x_p$ , які відповідають забезпеченостям  $P_i$ , тобто

$$x_p = k_p \cdot m_x,$$

де  $k_p = C_V \Phi_{p, C_S} + 1$

7. На основі змодельованого ряду розраховуються статистичні параметри та визначається їх відхилення від вхідних параметрів.

### Розрахункова частина

Приклад 1. Моделювання ряду річного стоку за моделлю 1.

- 1)  $m_x = 2$  л/с·км<sup>2</sup>;  $C_V = 0.3$ ;  $C_S = 2C_V$ , тобто  $C_S = 0.6$
- 2) Генеровані випадкові числа  $\delta_i$  наведені у табл. 5.10.
- 3) За виразом  $P_i = \delta_i \cdot 100\%$  розраховуються відповідні забезпеченості
- 4) По таблиці Фостера визначаються значення числа  $\Phi_{p, C_S}$ ;
- 5) Значення модуля річного стоку дорівнює  $q_p = k_p \cdot m_q$ ;
- 6) За результатами моделювання визначаються статистичні параметри

$$\bar{q}_M = 1.81 \text{ л/с·км}^2; \quad C_{VM} = 0.32; \quad C_{SM} = 0.51.$$

Відносні відхилення змодельованих параметрів від параметрів генеральної сукупності визначаються за виразами:

$$\frac{\bar{q}_M - m_q}{m_q} \cdot 100\% = -9.5\%;$$

$$\frac{C_{VM} - C_V}{C_V} \cdot 100\% = 6.67\%;$$

$$\frac{C_{SM} - C_S}{C_S} \cdot 100\% = 15\%.$$

Таблиця 5.10 - Моделювання ряду річного стоку за моделлю 1.

$N$	$\delta_i$	$P_i = \delta_i \cdot 100\%$	$\Phi_{p,C_S}$	$k_p$	$q_p, \text{л/с}\cdot\text{км}^2$
1	0.484	48.4	-0.10	0.97	1.94
2	0.284	28.4	0.484	1.13	2.26
3	0.357	35.7	0.30	1.09	2.18
4	0.927	92.7	-1.4	0.58	1.16
5	0.984	98.4	-1.88	0.44	0.88
6	0.475	47.5	-0.10	0.97	1.94
7	0.925	92.5	0.30	1.09	2.18
8	0.168	16.8	1.06	1.32	2.64
9	0.415	41.5	0.16	1.05	1.10
10	0.567	56.7	0.22	0.93	1.86

Приклад 2. Розглядається штучний ряд довжиною  $N=1000$  із заданими вихідними параметрами  $m_q = 2.0 \text{ л/с}\cdot\text{км}^2$ ;  $C_V = 0.4$ ;  $C_S = 1.2$ .

З тисячолітнього ряду вибираються по 10 вибірок з довжиною, яка відповідає 10, 25, 50 та 100 рокам. Для кожної з цих вибірок розраховуються оцінки статистичних параметрів. Середня квадратична похибка обчислених параметрів, визначених за вибірками заданої довжини  $n$ , розраховується за наступними формулами

$$\sigma_{\bar{q},n} = \sqrt{\frac{\sum_{j=1}^K (\bar{q}_j - m_q)^2}{K-1}};$$

$$\sigma_{C_V, n} = \sqrt{\frac{\sum_{j=1}^K (\hat{C}_{Vj} - C_V)^2}{K-1}},$$

де  $K=10$  через те, що загальна кількість вибірок довжиною  $n$  постійна і дорівнює 10;

$\hat{C}_{Vj}$  - значення коефіцієнта варіації, визначені для  $j$ -ої вибірки довжиною  $n$ ;

$\bar{q}_j$  - середнє арифметичне значення, розраховане для  $j$ -ої вибірки довжиною  $n$ .

Результати обчислень наведені у табл. 5.11, де  $q_{СЕР}$  та  $C_{V,СЕР}$  - це осереднені за  $K$  вибірками заданої довжини  $n$  значення статистичних параметрів.

Таблиця 5.11 – Розрахункові параметри змодельованих рядів при різній довжині вибірки  $n$

Обсяг вибірок $n$	$q_{СЕР}$	$\hat{C}_{V,СЕР}$	$\sigma_{q,n}^2$	$\sigma_{C_V,n}^2$
10	2.22	0.43	0.18	0.014
25	2.13	0.40	0.07	0.012
50	2.02	0.40	0.04	0.005
100	1.98	0.40	0.014	0.0009

Аналіз величин  $\sigma_{q,n}^2$  та  $\sigma_{C_V,n}^2$  показує, що похибки визначення статистичних параметрів зменшуються при збільшенні довжини вибірок  $n$ . Дані статистичних випробувань видаються студентам викладачем.

Задача практичної роботи полягає в аналізі результатів випробувань і побудові графіків  $\sigma_{q,n} = f(n)$  та  $\sigma_{C_V,n} = f(n)$ .

Приклад 3. Розглядаються генеральні сукупності, змодельовані на ПЕОМ ( $N=1000$ ). Математичне сподівання бралось постійним ( $m_q = 2.0$  л/с·км<sup>2</sup>), а змінювався лише коефіцієнт варіації генеральної сукупності. Оскільки співвідношення між  $C_S$  та  $C_V$  бралось постійним ( $C_S = 2C_V$ ), то зі зміною  $C_V$  змінювалося і значення  $C_S$ .

Вихідні параметри штучних сукупностей:

випробування 1:  $m_q = 2.0$  л/с·км<sup>2</sup>;  $C_V = 0.1$ ;  $C_S = 0.2$ ;

випробування 2:  $m_q = 2.0$  л/с·км<sup>2</sup>;  $C_V = 0.2$ ;  $C_S = 0.40$ ;



випробування 3:  $m_q = 2.0 \text{ л/с}\cdot\text{км}^2$ ;  $C_V = 0.3$ ;  $C_S = 0.60$ .

Кожний із модельованих 1000-літніх років розбивався на 40 вибірок довжиною  $n=25$ . Для цих рядів підраховувалися вибіркові статистичні параметри та обчислювалися дисперсії і середньоквадратичні похибки визначення статистичних параметрів. У результаті отримані такі характеристики:

випробування 1:  $\sigma_{q,n=25}^2 = 0.0036$ ;  $\sigma_{C_V,n=25}^2 = 0.0065$  ( $C_V = 0.1$ );

випробування 2:  $\sigma_{q,n=25}^2 = 0.015$ ;  $\sigma_{C_V,n=25}^2 = 0.021$  ( $C_V = 0.2$ );

випробування 3:  $\sigma_{q,n=25}^2 = 0.036$ ;  $\sigma_{C_V,n=25}^2 = 0.050$  ( $C_V = 0.3$ ).

Аналіз похибок  $\sigma_{q,n=25}^2$  та  $\sigma_{C_V,n=25}^2$  показує, що ці похибки зростають із збільшенням вихідної варіації  $C_V$ . Дані статистичних випробувань видаються студентам викладачем.

Задача практичної роботи полягає в аналізі цих даних і побудові залежностей  $\sigma_{q,n=25} = f(C_V)$  та  $\sigma_{C_V,n=25} = f(C_V)$ .

#### 5.4.2 Модель II (модельовання сукупності випадкових величин, зв'язаних кореляцією між суміжними членами рядів)

Сама модель II описана у розділі 4, де умовне математичне сподівання і умовна дисперсія представлені виразами (4.4) та (4.5).

При цьому використовується спрощений спосіб, запропонований у Російському державному гідрологічному інституті. Вважається, що послідовність величин стоку описується законом Пірсона III.

Числа Фостера при одній і тій же самій забезпеченості  $P$  будуть однаковими як для умовного, так і для безумовного розподілу.

Отже,

$$\frac{x_{i+1/i} - \bar{x}_{i+1/i}}{\sigma_{i+1/i}} = \frac{x_i^\circ - \bar{x}}{\sigma_x} \quad (5.36)$$

де  $x_{i+1/i}$  - члени змодельованої сукупності за умови наявності кореляційного зв'язку між суміжними членами ряду;

$\bar{x}_{i+1/i}$  - умовне математичне сподівання, визначене за (4.4), тобто

$$\bar{x}_{i+1/i} = \bar{x}_{i+1} = \bar{x} + r(x_i - \bar{x}); \quad (5.37)$$

$\sigma_{i+1/i}$  - умовне середньоквадратичне відхилення, визначене за (4.5), тобто

$$\sigma_{i+1/i} = \sigma_{i+1} = \sigma_x \sqrt{1-r^2}; \quad (5.38)$$

$x_{i+1/i}^o = x_{i+1}^o$  - значення  $i+1$  члена змодельованої сукупності за умови  $r=0$ ;  
 $r$  - коефіцієнт автокореляції між суміжними членами;  
 $\bar{x}, \sigma_x$  - безумовні математичне сподівання та середньоквадратичне відхилення.

З (5.36), з урахуванням (5.37) і (5.38), можна визначити

$$x_{i+1/i} = \bar{x} + r(x_i - \bar{x}) + (x_{i+1}^o - \bar{x})\sqrt{1-r^2} \quad (5.39)$$

Якщо обидві частини (5.39) поділити на  $\bar{x}$ , отримаємо

$$k_{i+1/i} = 1 + r(k_i - 1) + \Phi_{i+1/i}^o C_V \sqrt{1-r^2}, \quad (5.40)$$

де  $k_{i+1/i} = \frac{x_{i+1/i}}{\bar{x}}$ ;  $k_i = \frac{x_i}{\bar{x}}$ ;

$$\Phi_{i+1/i}^o = \frac{x_{i+1/i}^o - \bar{x}}{\sigma_x} = \frac{x_{i+1/i}^o - \bar{x}}{\bar{x}} C_V. \quad (5.41)$$

Схема моделювання:

1. Обирається перше випадкове число  $\delta_1 = P_1 \cdot 100\%$ .
2. За таблицею Пірсона III визначається  $x(P_1)$ , тобто значення стоку, яке відповідає забезпеченості  $P_1$  і заданим вихідними параметрам ( $\bar{x}$  та  $C_V$ ).
3. Обирається наступне випадкове число, яке береться забезпеченістю  $P_i = \delta_i \cdot 100\%$ .
4. Відповідно забезпеченості  $P_i$  з таблиці розподілу Пірсона III вибирається число  $\Phi_{i+1/i}^o$ ;
5. За (5.40) визначається  $k_{i+1/i}$ , при цьому  $r$  та  $C_V$  - задані вихідні параметри.
6. Шукане значення  $x_{i+1/i}$  обчислюється таким чином

$$x_{i+1/i} = k_{i+1/i} \cdot \bar{x}.$$

Надалі розраховане значення  $x_{i+1/i}$  береться як  $k_i$  і розрахунки повторюються, починаючи з п.3.

Приклад 4. Генеруються чотири штучних ряди з урахуванням наявності внутрішньорядних зв'язків за моделлю II. Вихідні параметри  $m_q = 2.0 \text{ л/с} \cdot \text{км}^2$ ;  $C_V = 0.4$ ;  $C_S = 3C_V$ , тобто  $C_S = 1.2$ .

У ході випробувань змінюється лише коефіцієнт автокореляції  $r$ . Кожен із змодельованих рядів довжиною  $N=1000$  розбивається на вибірки довжиною  $n=10$ . Отже, для кожного штучного ряду із заданим  $r$  отримується 100 вибірок, для яких розраховуються статистичні параметри і обчислюються середні квадратичні значення цих параметрів за формулами

$$\sigma_{q,n=10} = \sqrt{\frac{\sum_{j=1}^K (\bar{q}_j - m_q)^2}{K-1}};$$

$$\sigma_{C_V,n=10} = \sqrt{\frac{\sum_{j=1}^K (C_{Vj} - C_V)^2}{K-1}},$$

де  $K=100$ .

Результати моделювання зводяться до табл. 5.12, де  $q_{СЕР}$  і  $\hat{C}_{V,СЕР}$  - результат осереднення параметрів  $q$  та  $\hat{C}_V$  по 100 вибірках довжиною  $n=10$ .

Таблиця 5.12 – Розрахункові параметри модельованих рядів при різних  $r$

$r$	Параметри вибірки $n=10$			
	$q_{СЕР}$	$\hat{C}_{V,СЕР}$	$\sigma_{q,n=10}^2$	$\sigma_{C_V,n=10}^2$
0.1	2.24	0.415	0.215	$1.21 \cdot 10^{-2}$
0.2	2.27	0.402	0.261	$1.13 \cdot 10^{-2}$
0.3	2.30	0.387	0.318	$1.09 \cdot 10^{-2}$
0.4	2.33	0.371	0.389	$1.08 \cdot 10^{-2}$

Задача полягає в установленні зв'язків між точністю розрахунків параметрів та величинами коефіцієнтів автокореляції  $r$ .

Приклади побудови залежностей показані на рис. 5.10 та 5.11. Аналізуючи приклад, можна зробити висновок, що похибки визначення середнього арифметичного значення зростають при підвищенні тисноти зв'язку між попередніми та наступними значеннями членів ряду, а похибки визначення  $C_V$  зменшуються.

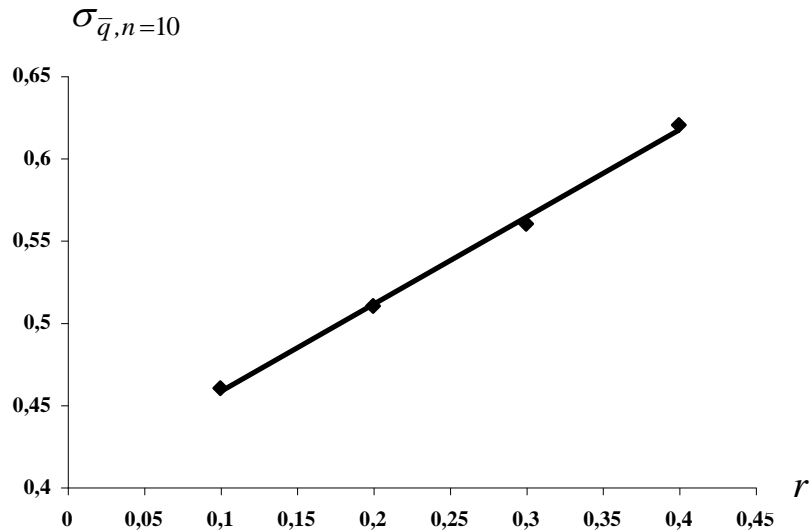


Рисунок 5.10 Залежність між середньоквадратичними похибками  $\sigma_{\bar{q}}$  і коефіцієнтами автокореляції

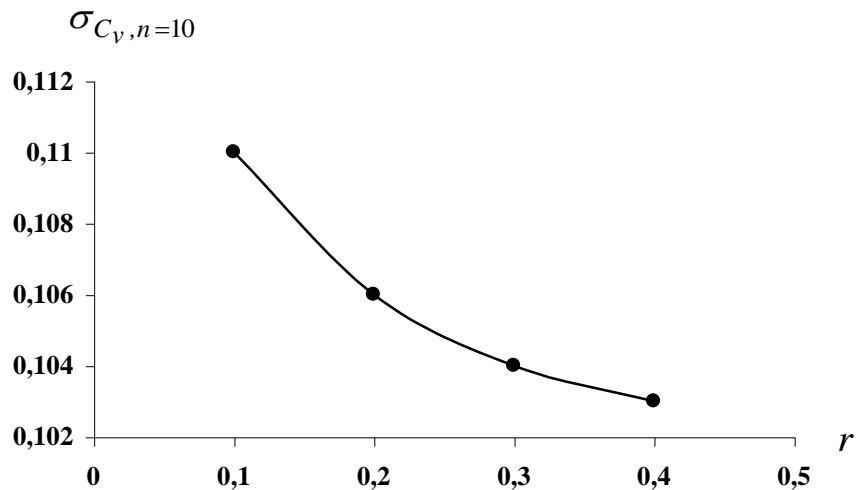


Рисунок 5.11- Залежність між середньоквадратичними похибками  $\sigma_{C_V}$  і коефіцієнтами автокореляції

### 5.4.3 Модель побутового стоку при наявності на водозборі штучних водойм

В основі моделі побутового стоку лежать рівняння, представлені виразами (5.7) та (5.8). Генерація штучних рядів річного побутового стоку відбувається за наступною схемою.

1.Визначаються параметри природного, тобто непорушеного господарською діяльністю стоку.

2.Здійснюється генерація штучних рядів природного стоку за прийнятою моделлю.

3.Виконується генерація штучних рядів побутового стоку при наявності на водозборі штучних водойм при заданому значенні відносної площі водної поверхні  $f_B$ , яка обчислюється як відношення сумарної площі водної поверхні до загальної площі водозбору.

4.Здійснюється визначення статистичних параметрів побутового стоку, отриманого шляхом моделювання.

Виконання практичного завдання передбачає роботу студентів на ПЕОМ.

Приклад 5. Параметри природного річного стоку (у вигляді шарів) задаються такими:  $\bar{Y} = 30$  мм;  $C_V = 0.4$ ;  $C_S = 0.9$ ;  $r = 0.2$ ; показник ступеня антропогенного впливу надається у вигляді відносної площі, зайнятої штучними водоймами  $f_B$ .

В результаті роботи з моделлю студент отримує значення параметрів побутового стоку: середнє арифметичне значення  $\bar{Y}_B$ , коефіцієнт варіації  $C_{VB}$  та коефіцієнт асиметрії  $C_{SB}$ , які зводяться до табл. 5.13.

Таблиця 5.13 – Результати моделювання рядів побутового стоку при різних  $f_B$

$f_B$	$\bar{Y}_B, \text{мм}$	$C_{VB}$	$C_{SB}$
0.01	26.0	0.47	1.07
0.02	22.0	0.56	1.05
0.03	18.0	0.69	1.02

На основі отриманих даних будуються залежності, які являють собою так звані „функції відгуку” статистичних параметрів на створення у межах водозбору штучних водойм. Задача полягає у тому, щоб проаналізувати зміну статистичних параметрів при збільшенні площ водної поверхні (рис.5.12) та визначити відповідні коефіцієнти антропогенного впливу (як відношення параметра побутового стоку до параметра природного стоку).



Коефіцієнти впливу штучних водойм на середньобагаторічну величину стоку становлять:

$$K_{\bar{Y},B} = \frac{\bar{Y}_B}{\bar{Y}} = \frac{26}{30} = 0.87 \quad (\text{при } f_B = 1\%);$$

$$K_{\bar{Y},B} = \frac{\bar{Y}_B}{\bar{Y}} = \frac{22}{30} = 0.73 \quad (\text{при } f_B = 2\%);$$

$$K_{\bar{Y},B} = \frac{\bar{Y}_B}{\bar{Y}} = \frac{18}{30} = 0.60 \quad (\text{при } f_B = 3\%).$$

Отже, можна зробити висновок, що при площі водної поверхні у 3% водні ресурси водозбору зменшаться на 40%.

Коефіцієнти впливу штучних водойм на коефіцієнти варіації становлять:

$$K_{C_v,B} = \frac{C_{v,B}}{C_v} = \frac{0.47}{0.40} = 1.18 \quad (\text{при } f_B = 1\%);$$

$$K_{C_v,B} = \frac{C_{v,B}}{C_v} = \frac{0.56}{0.40} = 1.40 \quad (\text{при } f_B = 2\%);$$

$$K_{C_v,B} = \frac{C_{v,B}}{C_v} = \frac{0.69}{0.40} = 1.72 \quad (\text{при } f_B = 3\%).$$

Коефіцієнти впливу штучних водойм на асиметрію рядів річного стоку визначаються наступним чином

$$K_{C_s,B} = \frac{C_{s,B}}{C_s} = \frac{1.07}{0.9} = 1.19 \quad (\text{при } f_B = 1\%);$$

$$K_{C_s,B} = \frac{C_{s,B}}{C_s} = \frac{1.05}{0.9} = 1.17 \quad (\text{при } f_B = 2\%);$$

$$K_{C_s,B} = \frac{C_{s,B}}{C_s} = \frac{1.02}{0.9} = 1.13 \quad (\text{при } f_B = 3\%).$$

За отриманими даними обчислюються величини річного стоку у природних і побутових умовах заданої забезпеченості  $P$ .

Наприклад, при  $P=75\%$  і параметрах природного стоку  $\bar{Y} = 30$  мм;  $C_v = 0.4$  та  $C_s = 0.9$  параметр  $\Phi_P$  за законом Пірсона III дорівнює 0.73, а  $Y_{P=75\%} = \bar{Y}(\Phi_{P=75\%, C_s=0.9} C_v + 1) = 21.3$  мм, а при параметрах побутового

стоку  $\bar{Y} = 18 \text{ мм}$ ;  $C_V = 0.69$ ;  $C_S = 1.02$  значення  $Y_{P=75\%, C_S=1.02}$  становить лише 9 мм.

Таким чином, у маловодні роки, які відповідають забезпеченості 75%, зменшення річного стоку при наявності штучних водойм з відносною площею 3% становить 70% у порівнянні з природним.

Таке зменшення величин стоку свідчить про суттєве порушення екологічного стану водозбору в цілому і можливість перебоїв у роботі гідротехнічних і гідромеліоративних споруд.



## ПРЕДМЕТНИЙ ПОКАЖЧИК

А  
Антропогенний вплив 131  
Автокореляційна функція 103,114

В  
Водне господарство 131  
Водні ресурси 131  
Водогосподарський баланс 134  
Випадкова величина 13  
Випадкова функція 97

Г  
Генерація випадкових чисел 172

Д  
Дисперсія випадкової величини 25  
Дисперсія випадкової функції 100  
Довірчий інтервал 89

Додаткове випаровування з водної  
поверхні 137, 139

Е  
Ергодичність 111

З  
Закон розподілу випадкової  
величини 19  
Забезпеченість 16  
Зрошення 136, 144  
Зворотні води 137, 147

Й  
Ймовірність 13

К  
Кліматичний стік 139, 141  
Коваріаційний момент 81,87,102

Коваріація 80, 81  
Коефіцієнт антропогенного  
впливу 150  
Кореляція 79  
Коефіцієнт кореляції 81, 88  
Коефіцієнт автокореляції 116

М  
Математичне сподівання випадкової  
величини 20  
Математичне сподівання випадкової  
функції 100  
Меліорація 131  
Метод водно-теплового  
балансу 139,144  
Метод Монте-Карло 30  
Моделювання стоку 121,123,136

Н  
Некорельовані випадкові  
величини 81  
Норма зрошення 144,145,146  
Норма осушування 165

О  
Осушування 157  
Оцінка статистичного параметра 22

П  
Побутовий стік 136  
Природний стік 136  
Простий ланцюг Маркова 12,121,122

С  
Статистичні параметри 22  
Стаціонарна випадкова  
функція 107, 108  
Стохастична залежність 79,80  
Стохастичні моделі річного стоку 172

## У

- Умовний закон розподілу 76
- Умовне математичне сподівання 84
- Умовне середньоквадратичне відхилення 84, 126

## Ф

- Функції антропогенного впливу 150,151

## ІМЕННИЙ ПОКАЖЧИК

### Б

Блохінов Є.Г. 38, 43, 125  
Болгов М.В. 121, 130, 132, 137

### Г

Гопченко Є.Д. 136, 139, 144

### К

Крицький С.М. 7, 8, 38, 40, 126, 132, 137

### Л

Лобода Н.С. 136, 139, 144, 152, 166

### М

Менкель М.Ф. 7, 8, 38, 40, 126, 132, 137, 157

### Р

Раткович Д.Я. 121, 132  
Рождєственський А.В. 132

### С

Сарманов І.О. 121, 125

### Ш

Шелутко В.А. 51, 53, 132  
Шикломанов І.А. 133, 148, 152  
Школьний Є.П. 86

## СПИСОК ЛІТЕРАТУРИ

1. Блохинов Е.Г. Распределение вероятностей величин речного стока.-М.: Наука, 1974.-169 с.
2. Блохинов Е.Г., Сарманов О.В. Гамма-корреляция и ее использование при расчетах многолетнего регулирования речного стока. // Тр.ГГИ.- Вып. 143 Л., Гидрометеиздат.-1968.- С.52-75.
3. Блохинов Е.Г., Сотникова Л.Ф. Об оценке параметров распределения вероятностей годового стока рек СССР. // Тр.ГГИ.-Вып. 180.-Л. Гидрометеиздат.-1970.- С. 85-113.
4. Болгов М.В., Сарманов И.О. Построение двумерной плотности для гидрологических и гидрогеологических приложений // Водные ресурсы. – 1991. - №3.- С.24-29.
5. Болгов М.В. , Лобода Н.С., Николаевич Н.Н. Пространственное обобщение параметров внутрирядной связности рядов годового стока // Метеорология и гидрология. - 1993. - №7. - С. 83 - 91.
6. Болгов М.В., Лобода Н.С., Николаевич Н.Н. Пространственное обобщение коэффициентов автокорреляции годового стока Украины // Труды УкрНИГМИ. – 1993. - Вып. 245. - С. 22 - 29.
7. Болгов М.В. , Лобода Н.С., Николаевич Н. Н., Сарманов И.О. О свойствах выборочных оценок параметров моделей марковских процессов с линейной гамма-корреляцией смежных членов // Метеор. климат. и гидр. - 1993. - Вып. 29. - С. 110 - 122.
8. Водноэнергетические расчеты методом Монте-Карло / Под ред. А.Ш.Резниковского. – М.:Энергия. – 1969. – 302 с.
9. Гопченко Е.Д., Лобода Н.С. Применение методов статистического моделирования при оценке изменений годового стока рек под влиянием орошения // Метеорология и гидрология. – 1986. - №9. – С. 79-84.
10. Гопченко Є., Лобода Н. Динаміко - стохастична модель стоку зрошуваних річок Північно-Західного Причорномор'я в умовах глобального потепління // Гідрологія, гідрохімія і гідроекологія: Науковий збірник. - т.1. – Київ: Ніка - центр, 2000. - С. 154 - 158.
11. Казакевич Д.И. Основы теории случайных функций и ее применение в гидрометеорологии. – Л.: Гидрометеиздат, 1971. – 230 с.
12. Картвелиашвили Н.А. Стохастическая гидрология. - Л.: Гидрометеиздат, 1981. – 167 с.
13. Коваленко В.В. Гидрометрическое оценивание речного стока с элементами стохастического подхода. - Л.: ЛПИ, - 1986. – 61 с.
14. Коваленко В.В. Динамические и стохастические модели гидрологического цикла. – Л.: ЛПИ, - 1988. – 34 с.

15. Крицкий С.Н., Менкель М.Ф. Гидрологические основы управления речным стоком.- М. Наука,1981. - 235с.
16. Крицкий С.Н., Менкель М.Ф. Гидрологические основы управления водохозяйственными системами.- М. Наука,1982. – 271 с.
17. Лобода Н.С. Расчеты и обобщения характеристик годового стока рек Украины в условиях антропогенного влияния. - Монографія. – Одеса: Екологія, 2005. -208с.
18. Лобода Н.С., Шаменкова О.И. Обоснование стохастических моделей годового суммарного и подземного стока в условиях осушительных мелиораций (по материалам западного Полесья) // Міжвід. наук. зб. України. - Метеорологія, кліматологія та гідрологія. - Одеса. - 2002. – Вип. 46. - С. 279 – 293.
19. Музылев С.В., Привальский В.Е., Раткович Д.Я. Стохастические модели в инженерной гидрологии. – М.: Наука. – 1982. – 184 с.
20. Пособие по определению расчетных гидрологических характеристик. - Л.: Гидрометеиздат,- 1984. - 447с.
21. Пространственно-временные колебания стока рек СССР / Под ред. А.В. Рождественского. – Л.: Гидрометеиздат,1988. – 376 с.
22. Раткович Д.Я. Многолетние колебания речного стока. - Л.: Гидрометеиздат. - 1976. - 255 с.
23. Раткович Д.Я. Гидрологические основы водообеспечения.- М.: РАН ИВП. - 1995. - 428с.
24. Раткович Д.Я., Болгов М.В. Стохастические модели колебаний составляющих водного баланса речного бассейна. - М.: РАН ИВП. - 1997. - 262 с.
25. Рождественский А.В., Чеботарев А.И. Статистические методы в гидрологии. - Л.: Гидрометеиздат, 1974. - 424с.
26. Рождественский А.В., Ежов А.В., Сахарюк А.В. Оценка точности гидрологических расчетов. - Л.: Гидрометеиздат, 1990. - 276 с.
27. Сарманов И.О. Построение корреляции между равномерно-распределенными случайными величинами // Тр. ГГИ.-1968.-вып. 160.- С.81-89.
28. Сарманов И.О. О корреляции между функциями зависимых случайных величин, имеющих гидрологическое приложение // Проблемы регулирования и использования водных ресурсов. - М.: Наука, 1973.- С. 87-103.
29. Сарманов И.О. Новые формы корреляции для гидрологических приложений // Водные ресурсы. - 1975. №2. С. 18-31.
30. Сарманов И.О. Точное распределение членов ранжированного ряда и применение корреляции между ними для моделирования гидрологического ряда.// Тр.V Всесоюз. гидрологического съезда.-Л.: Гидрометеиздат, 1989.- С.184-190.

31. Сарманов О.В. Псевдонормальная корреляция и различные ее обобщения // Доклады АН СССР.-т. 132.-1960.-№2.- С. 299-302.
32. Сарманов О.В. Исследование стационарных марковских процессов методом разложения по собственным функциям // Тр. МИАН.-1961.- С.239-261.
33. Сарманов О.В., Сарманов И.О. Основные типы корреляции, применяемые в гидрологии.-М.: Наука, 1983.-200 с.
34. Шелутко В.А. Статистические модели и методы исследования многолетних колебаний годового стока. -Л.: Гидрометеиздат,1984.- 160 с.
35. Шелутко В.А. Численные методы в гидрологии. -Л.: Гидрометеиздат,1991.- 238 с.
36. Шикломанов И.А. Антропогенные изменения водности рек. - Л.: Гидрометеиздат, 1979. - 302 с.
37. Шикломанов И.А. Влияние хозяйственной деятельности на речной сток. - Л.: Гидрометеиздат, 1989. - 335 с.
38. Школьный Є.П., Лоева І.Д., Гончарова Л.Д. Обробка та аналіз гідрометеорологічної інформації: навчальний підручник. - К.: Міносвіти України, 1999. - 600 с.
39. Loboda N.S. Stochastic model of long-term annual flow fluctuations under anthropogenic influences (in Ukraine)// Special Issue "Stochastic models of hydro-logical processes and their applications to problems of environmental preservation". ARW Series.-1998.-Vol.1.-P.110-118.
40. Loboda N.S. The assessment of present and future Ukrainian water resources on meteorological evidence // Proceedings of The Second International Conference on Climate and Water. - vol.3. - Espoo, Finland: Edita Ltd, Helsinki. - 1998. - p.1486-1494.
41. Loboda N.S. New Stochastic and Fractal Multisystems models of Non-linear Hydrological Systems and Applications to description of Hydrological processes and Enviromental preservation// Ecology of Siberia, the Far East and the Arctic.-2001.-Vol.1- P.84-86.

## Додаток

### Таблиця випадкових чисел Дж. В. Вільямса

11	16	43	63	18		75	06	13	76	74		40	60	31	61	52		83	23	53	73	61	
21	21	59	17	91		76	83	15	86	78		40	94	15	35	85		69	96	86	09	16	
10	43	84	44	82		66	55	83	76	49		73	50	58	34	72		55	95	31	79	57	
36	79	22	62	36		33	26	66	65	83		39	41	21	60	13		11	44	28	93	20	
73	94	40	47	73		12	03	25	14	14		57	99	47	67	48		54	62	74	85	11	
49	56	31	28	72		14	06	39	31	04		61	83	45	91	99		15	46	98	22	85	
64	20	84	82	37		41	70	17	31	17		91	40	27	72	27		79	51	62	10	07	
51	48	67	28	75		38	60	52	93	41		58	99	98	38	80		20	12	51	07	94	
99	75	62	63	60		64	51	61	79	71		40	68	49	99	48		33	88	07	64	13	
71	32	55	52	17		13	01	57	29	07		75	97	86	42	98		08	07	46	20	55	
65	28	59	71	98		12	13	85	30	10		34	55	63	98	61		88	26	77	60	61	
17	26	45	73	27		38	22	42	93	01		65	99	05	70	48		25	06	77	75	71	
95	63	99	97	54		31	19	99	25	58		16	38	11	50	69		25	41	68	78	75	
61	55	57	64	04		86	21	01	18	08		52	45	88	88	80		78	35	26	79	13	
78	13	79	87	68		04	68	98	71	30		33	00	78	56	07		92	00	84	48	97	
62	49	09	92	15		84	98	72	87	59		38	71	23	15	12		08	58	86	14	90	
24	21	66	34	44		21	28	30	70	44		58	72	20	36	78		19	18	66	96	02	
16	97	59	54	28		33	22	65	59	03		26	18	86	94	94		51	35	14	77	99	
59	13	83	95	42		71	16	85	76	09		12	89	35	40	48		07	25	58	61	49	
29	47	85	96	52		50	41	43	19	66		33	18	68	13	46		85	09	53	72	82	
96	15	59	50	09		27	42	97	29	18		79	89	32	94	48		88	39	25	42	11	
29	62	16	65	83		62	96	61	24	68		48	44	91	51	02		44	12	61	94	38	
12	63	97	52	91		71	02	01	72	65		94	20	50	42	59		68	98	35	05	61	
14	54	43	71	34		54	71	40	24	01		38	64	80	92	78		81	31	37	74	00	
83	40	38	88	27		09	83	41	13	33		04	29	24	60	28		75	66	62	69	54	
67	64	20	52	04		30	69	74	48	06		17	02	64	97	37		85	87	51	21	39	
64	04	19	90	11		61	04	02	73	09		48	07	07	68	48		02	53	19	77	37	
17	04	89	45	23		97	44	45	99	04		30	15	99	54	50		83	77	84	61	15	
93	03	98	94	16		52	79	51	06	31		12	14	89	22	31		31	36	16	06	50	
82	24	43	43	92		96	60	71	72	20		73	83	87	70	67		24	86	39	75	76	
96	99	05	52	44		70	69	32	52	55		73	54	74	37	59		95	63	23	95	55	
09	11	97	48	03		97	30	38	87	01		07	27	79	32	17		79	42	12	17	69	
57	66	64	12	04		47	58	97	83	64		65	12	84	83	34		07	49	32	80	98	
46	49	26	15	94		26	72	95	82	72		38	71	66	13	80		60	21	20	50	99	
08	43	31	91	72		08	32	02	08	39		31	92	17	64	58		73	72	00	86	57	
10	01	17	50	04		86	05	44	11	97		57	23	82	74	64		61	48	75	23	29	
92	42	06	54	31		16	53	00	55	40		24	21	94	10	90		08	53	16	15	78	
35	54	25	58	65		07	30	44	70	10		31	30	94	93	87		02	33	00	24	76	
86	59	52	62	47		18	55	22	94	91		20	75	09	70	24		72	61	96	66	28	
72	11	53	49	85		58	03	69	91	37		28	53	78	43	95		26	65	43	78	51	
07	42	85	88	63		96	02	38	89	36		97	92	94	12	20		86	43	19	44	85	
35	37	92	79	22		28	90	65	50	13		40	56	83	32	22		40	48	69	11	22	
10	98	22	28	07		10	92	02	62	99		41	48	39	29	35		17	06	17	82	52	
90	12	73	33	41		77	80	61	24	46		93	04	06	64	76		24	99	04	10	99	
63	00	21	29	90		23	51	06	87	74		76	86	93	93	00		84	97	80	75	04	
40	77	98	63	82		48	45	46	52	69		02	98	25	79	91		50	76	59	19	30	

Продовження таблиці

43	21	61	26	08		18	16	78	46	31		94	47	97	65	00		39	17	00	66	29	
96	16	76	43	75		74	10	89	36	43		52	29	17	58	22		95	96	69	09	47	
70	97	56	26	93		35	68	47	26	07		03	68	40	36	00		52	83	15	53	81	
85	81	26	18	75		23	57	07	57	54		58	93	92	83	66		86	76	56	74	65	
37	10	06	24	92		63	64	24	76	38		54	72	35	65	27		53	07	63	82	35	
53	40	61	38	55		38	51	92	95	00		84	82	88	12	48		25	54	83	40	75	
55	17	28	15	56		18	85	65	90	43		65	79	90	19	14		81	36	30	51	73	
40	35	38	48	07		47	76	74	68	90		87	91	73	85	49		48	21	37	17	08	
18	89	90	96	12		77	54	15	76	75		26	90	78	81	73		71	18	92	83	77	
68	14	12	53	40		92	55	11	13	26		68	05	26	54	22		88	46	00	63	52	
51	55	99	11	59		81	31	06	32	51		42	58	76	81	49		88	14	79	97	00	
92	21	43	33	86		73	45	97	93	59		97	17	65	54	16		67	64	20	50	51	
15	08	95	05	57		33	16	68	70	94		53	29	58	71	33		38	26	49	47	08	
96	46	10	06	04		11	12	02	22	54		23	01	19	41	08		29	19	66	51	87	
28	17	74	41	11		15	70	57	38	35		75	76	84	95	49		24	54	36	32	85	
66	95	34	47	37		81	12	70	74	93		86	66	87	03	41		66	46	07	56	48	
19	71	22	72	63		84	57	54	98	20		56	72	77	20	36		50	34	73	35	21	
68	75	66	47	57		19	98	79	22	22		27	93	67	80	10		09	61	70	44	08	
75	02	26	53	32		98	60	62	94	51		31	99	46	90	72		37	35	49	30	25	
11	32	37	00	69		90	26	98	92	66		02	98	59	53	03		15	18	25	01	66	
55	20	86	34	70		18	15	82	52	83		89	96	51	02	06		95	83	09	54	06	
11	47	40	87	86		05	59	46	70	45		45	58	72	96	11		98	57	94	24	81	
81	42	28	68	42		60	99	77	96	69		01	07	10	85	30		74	30	57	75	09	
21	77	17	59	63		23	15	19	02	74		90	20	96	85	21		14	29	33	91	94	
42	27	81	21	60		32	57	61	42	78		04	98	26	84	70		27	87	51	54	80	
17	69	76	01	14		63	24	73	20	96		19	74	02	46	37		97	37	73	21	12	
05	68	63	02	43		34	12	40	29	36		50	19	77	98	69		86	49	76	87	09	
52	99	24	66	50		89	91	05	73	95		46	95	46	75	36		28	96	88	19	36	
94	51	89	39	84		81	47	86	77	50		82	54	96	26	76		31	12	34	98	99	
00	18	47	21	86		78	90	67	54	80		61	79	88	16	00		80	01	88	47	42	
87	46	26	31	65		79	81	66	16	30		57	66	62	90	55		46	51	80	14	87	
88	69	25	87	16		12	27	34	81	76		29	80	56	49	94		66	87	26	22	30	
20	09	44	29	62		41	38	21	67	68		06	71	13	49	39		19	59	97	62	47	
60	93	58	15	04		50	52	08	21	53		13	93	44	68	85		58	31	58	83	60	
51	39	28	59	36		43	89	85	05	96		25	54	99	83	27		99	94	32	53	77	
54	23	94	19	18		79	52	64	62	74		40	87	16	18	03		25	76	75	54	84	
57	89	27	33	94		07	16	09	02	62		47	70	43	83	55		71	70	88	01	17	
02	33	07	47	36		53	27	44	44	68		62	61	11	96	98		09	30	42	92	65	
76	11	52	92	47		55	34	25	12	99		03	04	78	39	81		11	91	60	92	67	
63	31	28	18	86		29	08	52	01	01		26	46	05	05	01		31	73	11	89	38	
27	63	22	15	70		34	27	45	64	26		01	76	42	59	59		69	29	38	98	75	
06	33	56	21	41		44	01	45	25	67		11	76	25	48	06		02	65	15	29	12	
64	14	28	76	76		21	35	88	87	73		31	73	63	16	95		11	52	36	42	13	
28	43	62	54	68		75	23	57	53	70		97	15	54	87	06		52	23	92	18	31	
09	52	28	38	55		85	97	31	58	88		31	18	14	96	72		17	23	70	40	24	
93	71	41	54	14		93	71	20	27	42		32	11	58	26	83		67	18	28	90	30	
15	68	15	35	99		58	18	57	38	40		07	06	87	59	47		71	74	36	92	85	
77	71	22	39	14		08	90	74	37	68		26	62	27	41	84		75	16	69	67	48	
78	45	35	48	44		61	50	90	12	45		02	80	55	26	76		22	51	94	78	48	
24	86	06	82	84		19	36	72	90	73		32	30	15	87	01		04	19	33	01	42	
37	28	40	68	44		78	88	75	72	76		26	33	95	69	09		39	33	14	21	01	



Продовження таблиці

35	48	85	24	73		37	63	43	25	69		95	27	40	95	08		81	01	24	24	13
51	59	55	99	09		35	22	34	49	91		24	27	53	96	32		05	77	79	88	00
90	66	03	51	71		30	02	19	11	20		36	11	64	21	28		69	40	19	41	99
47	50	50	20	08		20	30	08	71	88		96	19	50	70	59		13	26	63	13	89
13	35	00	84	14		64	04	99	43	77		22	40	89	49	58		19	09	55	80	35
33	00	69	26	90		69	24	89	74	43		53	89	62	35	08		16	22	75	69	29
55	21	66	38	86		06	80	41	18	61		22	56	50	24	75		00	25	87	90	18
21	99	12	62	28		14	80	11	91	92		49	43	82	07	72		60	84	66	97	32
71	02	52	82	12		10	47	42	75	22		65	62	03	46	84		00	21	00	48	63
65	52	21	52	42		84	55	47	45	60		20	24	62	69	41		41	29	80	47	63
27	97	55	49	23		90	65	00	61	70		09	43	30	91	67		35	16	63	27	31
07	30	00	97	04		36	09	96	15	77		95	55	27	34	56		16	57	88	81	40
54	35	71	36	89		19	56	90	38	14		76	05	30	51	50		99	12	56	94	42
00	97	70	44	81		42	04	40	86	49		34	82	23	58	43		78	46	88	23	80
13	92	07	87	61		12	31	19	28	08		07	75	30	40	73		58	52	08	00	22
08	39	53	70	43		37	88	03	41	72		04	20	49	44	34		62	79	88	19	02
46	16	66	72	06		01	61	94	37	69		96	77	01	94	40		29	70	04	20	93
87	76	77	76	07		03	74	20	16	13		65	98	96	28	43		10	91	73	44	58
29	88	09	52	88		21	64	44	65	87		06	64	49	47	84		66	99	56	18	12
36	24	83	66	66		14	89	45	92	73		88	95	04	60	77		34	65	11	20	38
12	38	62	96	56		30	47	42	59	64		21	48	29	54	22		02	00	23	36	71
52	06	87	38	01		52	18	81	94	91		55	13	76	10	39		02	00	66	99	13
41	72	75	21	71		56	71	90	60	54		98	44	18	15	29		59	60	76	52	25
49	31	97	45	80		57	47	01	46	00		57	16	83	04	58		23	89	20	78	25
84	78	67	69	63		12	12	72	50	14		71	88	66	53	34		38	01	30	93	79
64	86	69	52	02		43	98	37	26	55		40	41	85	95	04		52	38	30	72	32
11	84	92	64	82		20	46	19	94	50		28	83	37	66	61		47	27	79	29	35
54	96	61	75	94		57	39	37	32	67		37	88	36	21	24		62	19	94	95	42
10	95	93	33	49		80	71	99	67	51		44	88	23	35	92		66	23	41	38	21
22	78	40	77	83		35	90	30	00	91		19	08	21	38	73		07	18	42	15	66
86	03	76	17	91		33	81	56	39	68		45	31	62	92	83		89	31	85	58	06
80	03	76	50	89		85	91	97	43	91		22	78	85	54	33		31	18	87	48	82
72	75	18	43	59		15	76	91	36	15		08	29	38	61	93		05	02	62	12	55
18	53	20	38	74		66	22	07	90	50		29	22	37	05	41		67	11	58	45	84
22	93	62	20	58		49	17	11	10	27		22	68	18	01	10		31	59	50	92	46
66	39	77	65	10		81	15	00	07	04		74	58	09	03	54		43	74	42	21	78
89	73	02	32	72		65	42	03	50	91		69	09	37	13	64		08	10	79	69	52
81	82	17	53	23		96	06	89	17	24		40	45	69	12	34		58	09	06	53	42
94	37	78	25	54		53	58	61	14	32		72	92	76	73	49		83	96	25	89	12
68	48	54	99	91		53	16	51	98	65		61	86	93	30	93		81	12	90	64	81
07	33	00	71	84		86	78	86	45	77		40	04	81	65	20		07	63	81	07	97
10	99	31	49	30		35	07	23	64	29		68	77	39	76	69		28	65	68	99	38
20	80	11	51	78		64	45	38	33	57		09	77	43	07	51		49	74	01	13	85
79	24	13	53	47		66	85	17	92	47		46	13	93	66	89		82	58	71	35	86
43	59	33	95	55		97	34	55	84	94		26	56	69	53	23		32	99	38	99	88
29	52	26	27	13		33	70	11	71	86		06	76	55	71	41		48	61	71	82	82
88	83	64	72	90		67	27	47	83	62		35	38	49	03	80		12	31	78	97	02
65	90	56	62	53		91	48	23	06	89		49	33	37	84	82		36	19	91	13	55
44	79	86	93	71		07	86	58	17	56		45	59	51	40	44		56	80	69	91	26
35	51	09	91	39		32	03	12	79	25		79	81	91	50	54		76	17	41	22	06
50	12	59	32	23		64	20	94	97	14		11	97	16	22	34		74	85	74	64	01

Продовження таблиці

25	17	39	00	38		63	87	14	04	18		11	45	28	93	18		53	08	42	19	93	
68	45	99	00	94		44	99	59	37	18		38	74	68	12	71		96	26	09	81	37	
93	36	91	30	44		69	68	67	81	62		66	37	80	29	19		34	01	25	00	80	
19	36	05	50	49		94	95	17	63	41		84	01	93	06	90		25	65	67	29	96	
47	79	88	98	90		06	89	36	54	83		17	70	12	12	92		14	88	01	53	86	
69	22	33	20	07		03	51	36	11	49		32	54	69	20	72		62	52	22	15	04	
34	51	15	07	21		84	85	03	41	59		97	13	86	19	19		97	78	92	85	75	
54	03	15	93	29		58	96	35	22	20		35	29	22	79	24		55	46	74	30	36	
66	72	28	55	15		04	72	39	24	11		02	73	70	81	68		30	04	36	34	50	
71	05	90	74	96		38	40	41	81	26		28	26	13	78	44		12	54	31	43	98	
45	47	88	60	66		31	13	53	32	43		80	57	33	06	06		48	64	45	30	08	
97	24	69	11	21		89	43	72	03	93		77	15	38	85	52		26	84	31	28	44	
22	98	22	59	36		96	41	73	48	45		85	14	95	75	04		15	05	93	68	49	
48	24	36	29	93		47	13	28	52	48		35	22	97	28	37		36	75	27	16	55	
93	51	41	49	15		67	96	08	22	03		40	11	72	43	46		32	18	08	70	74	
69	70	79	83	03		93	06	91	62	16		60	87	59	75	45		68	65	29	21	60	
87	46	79	17	94		70	81	41	27	43		03	76	93	25	51		74	80	14	16	92	
81	00	68	14	98		59	37	53	05	02		94	07	79	22	09		31	50	66	96	06	
15	45	88	34	81		50	18	74	33	75		94	37	60	06	66		94	14	52	23	99	
33	46	91	25	10		23	09	54	80	16		42	35	41	13	47		90	92	00	38	64	
67	19	80	71	76		65	99	61	83	17		81	14	94	32	91		10	81	74	43	48	
58	03	79	22	61		85	50	45	56	90		10	63	17	82	38		00	15	74	62	59	
84	98	36	83	12		25	51	95	61	58		86	30	00	76	89		14	00	67	77	53	
35	55	40	29	35		72	88	96	87	72		19	85	03	96	50		65	22	21	55	63	
04	36	81	76	32		50	96	27	19	08		94	46	46	64	32		62	24	31	36	74	
81	31	16	04	79		69	98	53	09	52		23	92	14	97	30		21	71	89	23	14	
03	82	38	98	87		55	82	87	44	52		72	77	52	37	16		42	85	37	47	93	
80	42	26	54	37		38	79	75	62	61		27	81	64	67	04		82	73	50	33	39	
61	30	74	94	68		43	34	44	37	00		20	20	77	70	88		17	16	72	45	31	
83	87	38	25	57		10	00	28	00	93		59	28	30	44	94		60	72	52	14	31	
38	11	01	68	55		28	92	29	37	58		88	73	13	63	76		51	38	35	76	19	
43	89	29	11	89		87	22	65	69	35		84	76	26	79	96		75	00	00	17	45	
93	68	30	96	64		53	92	74	98	85		20	75	49	23	55		57	95	51	09	40	
32	74	80	21	21		11	97	29	69	14		28	06	56	95	64		06	83	55	68	45	
49	21	19	29	63		38	62	56	53	12		62	17	57	33	53		84	97	21	77	26	
63	36	56	42	24		69	47	55	75	12		11	04	45	04	83		68	82	19	74	26	
63	57	62	63	73		44	61	04	37	48		00	33	16	34	22		99	62	27	67	57	
41	07	84	70	36		65	52	46	84	66		67	15	72	64	19		37	97	81	65	11	
70	84	68	95	58		64	17	31	53	81		87	71	35	08	41		46	27	02	65	08	
68	80	06	44	92		20	16	23	27	07		10	28	18	25	25		74	15	58	67	49	
44	97	78	95	25		51	26	96	37	47		91	36	77	40	33		67	02	06	90	92	
79	35	46	38	47		24	39	55	36	79		40	56	03	69	14		69	17	63	19	18	
14	95	42	22	99		40	15	65	26	85		29	22	33	83	83		30	31	57	09	99	
01	71	19	84	39		09	44	63	39	37		49	09	54	02	38		81	69	71	24	74	
62	32	85	53	28		45	73	89	39	40		27	46	62	69	27		53	34	51	13	79	
73	00	46	21	09		81	90	77	10	77		57	46	37	00	45		65	12	34	90	70	
	21	88	94	45		05	60	95	23	36		50	55	89	22	42		52	73	28	15	02	
99	15	90	19	68		45	88	68	68	75		28	41	39	59	18		44	15	64	69	59	
92	85	82	99	49		15	81	79	33	72		56	65	74	31	93		58	13	05	42	73	
27	39	69	74	77		65	55	47	16	01		13	12	16	88	67		95	76	35	96	67	
37	10	34	53	09		30	12	94	33	80		96	99	68	93	56		22	78	46	01	84	

Продовження таблиці

57	34	79	70	12		48	42	82	06	06		60	74	22	22	26		89	99	32	45	97	
29	41	72	09	72		84	13	42	91	66		71	99	34	28	19		28	79	91	30	95	
70	18	50	54	60		07	35	38	58	55		11	83	80	22	91		89	88	24	16	13	
06	39	35	20	17		21	32	62	05	30		14	30	05	53	19		47	90	83	92	32	
42	60	59	24	74		09	36	14	24	04		33	72	24	77	22		12	54	72	97	59	
26	48	34	88	73		84	26	03	32	85		51	27	58	48	21		87	44	30	24	91	
93	60	97	74	25		57	26	86	93	20		13	56	80	08	25		80	83	58	36	39	
19	79	00	32	65		68	47	42	96	78		08	45	60	14	75		40	31	77	20	77	
63	43	46	72	51		20	51	08	06	27		62	61	86	40	27		64	47	75	04	32	
76	63	64	27	71		12	58	35	63	29		59	51	22	52	49		42	18	50	96	22	
51	25	87	43	30		47	75	15	26	72		10	99	59	19	22		50	21	71	41	87	
13	26	01	42	38		88	39	15	08	58		60	73	55	89	65		58	49	20	96	66	
31	81	15	09	35		89	51	01	55	66		22	09	62	63	38		30	24	00	43	43	
48	83	26	95	24		32	11	34	66	73		24	96	59	71	07		41	85	68	52	96	
32	36	71	84	57		03	31	22	55	64		75	10	99	31	42		91	95	16	88	94	
90	34	61	74	76		71	03	92	14	97		64	00	15	42	64		71	66	79	42	62	
98	51	30	86	79		69	41	38	20	67		68	25	07	45	01		48	84	60	91	00	
08	48	41	56	17		82	79	00	83	24		95	09	12	70	82		73	94	95	95	33	
60	04	28	75	62		49	72	82	64	00		10	11	21	06	30		82	87	33	93	12	
87	65	38	64	09		09	27	73	83	79		09	93	45	75	87		44	00	52	10	90	
35	71	10	03	98		30	49	45	09	23		33	30	85	14	41		13	80	43	26	09	
46	48	33	27	22		74	79	96	02	77		09	29	94	26	17		94	65	41	13	15	
60	85	80	63	37		20	22	03	49	52		98	21	86	97	10		03	42	79	92	55	
70	78	30	14	52		87	66	12	40	28		41	83	04	04	72		46	04	10	89	52	
65	34	64	30	42		62	13	43	22	85		35	20	18	56	72		66	68	85	39	48	
96	34	95	21	64		23	29	87	22	17		41	64	06	30	18		74	28	09	16	70	
63	46	91	34	59		11	09	19	66	46		37	84	93	43	89		55	97	49	94	43	
60	57	97	54	98		97	81	58	61	51		60	23	67	57	05		83	33	98	49	23	
31	01	87	40	81		01	49	40	61	65		70	80	28	81	89		11	59	41	34	47	
00	04	73	38	02		73	34	25	43	38		43	35	49	39	50		54	76	53	75	53	
55	75	48	69	00		87	93	10	12	38		15	44	92	88	47		82	63	56	85	43	
38	05	83	03	46		35	72	27	88	49		92	82	54	05	36		76	78	04	42	64	
06	90	44	32	00		89	09	23	09	74		00	70	16	65	89		04	62	25	13	54	
48	43	07	91	89		59	69	09	38	30		59	93	99	51	37		65	79	94	11	47	
78	55	65	87	76		35	37	02	47	74		57	49	08	58	05		04	16	80	74	03	
07	66	20	77	12		70	41	09	94	28		76	51	36	93	39		33	70	13	21	84	
98	74	51	23	38		11	08	80	54	54		06	87	15	01	32		50	01	02	40	89	
52	92	38	13	43		14	90	82	38	96		01	52	90	94	99		57	66	00	23	09	
35	16	59	31	08		32	06	82	15	47		54	73	97	37	96		14	28	37	69	59	
65	22	11	22	07		66	07	95	94	34		94	94	61	72	41		82	75	41	80	73	
74	45	35	59	75		85	46	39	45	67		49	80	03	00	54		85	78	02	30	94	
03	49	66	06	56		69	73	55	50	53		15	39	38	25	29		15	76	08	97	88	
64	25	28	09	87		44	41	63	14	38		26	05	53	00	56		15	22	61	10	81	
49	22	51	16	47		69	73	37	99	39		86	30	53	50	43		66	21	97	48	31	
62	47	61	93	65		16	56	86	25	09		89	14	62	53	24		22	75	93	56	01	
21	11	93	11	66		78	38	51	48	75		86	85	79	94	04		91	54	77	49	34	
35	08	65	00	39		71	25	14	57	69		27	90	84	89	44		09	01	75	43	97	
33	33	07	87	01		88	53	64	13	98		86	41	32	07	09		41	58	96	36	41	
68	34	64	53	55		37	07	24	08	32		33	91	55	38	90		19	88	73	41	08	
55	52	63	17	47		57	30	90	92	38		78	74	26	99	71		30	14	87	26	47	
86	02	05	57	48		08	63	66	37	73		07	22	80	50	82		72	82	19	10	26	

*Навчальне видання*

**ЛОБОДА** Наталія Степанівна  
**ГОПЧЕНКО** Євген Дмитрович

**СТОХАСТИЧНІ МОДЕЛІ У ГІДРОЛОГІЧНИХ  
РОЗРАХУНКАХ**

*Навчальний посібник*

Підп. до друку                      Формат 60x84/16 Папір офс.  
Умовн. друк. арк.                      Тираж                      Зам. №  
Надруковано з готового оригінал-макета

---

Одеський державний екологічний університет  
65016, Одеса, вул.Львівська, 15

---

Видавництво і друкарня “Екологія”

**Лобода Н.С.  
Гопченко Є.Д.**

Л68 Стохастичні моделі у гідрологічних розрахунках: Навчальний посібник. –  
Одеса: Екологія, 2006. –194с.

ISBN

В навчальному посібнику викладені теоретичні та практичні основи стохастичного моделювання характеристик стоку у природних та порушених водогосподарською діяльністю умовах. Розглянуті питання визначення статистичних параметрів стоку за даними спостережень, установлення параметрів умовного закону розподілу. Представлені стохастичні моделі стоку у виді простого ланцюга Маркова. Посібник розрахований на студентів, магістрів, аспірантів гідрологічних та гідромеліоративних спеціальностей вищих навчальних закладів.

**ББК 26.22  
УДК 556.16.01**