КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА

МЕТОДИЧНІ ВКАЗІВКИ ДО ВИКОНАННЯ ПРАКТИЧНИХ РОБІТ З ДИСЦИПЛІНИ «**МАТЕМАТИЧНІ МЕТОДИ В ГІДРОМЕТЕОРОЛОГІЇ**»

ДЛЯ СТУДЕНТІВ ГЕОГРАФІЧНОГО ФАКУЛЬТЕТУ

НАПРЯМ: 6.040105 ГІДРОМЕТЕОРОЛОГІЯ
СПЕЦІАЛЬНІСТЬ: ГІДРОЛОГІЯ ТА ГІДРОХІМІЯ

Рецензент:

канд.геогр. наук Гребінь В.В.

*Рекомендовано до друку вченою радою
географічного університету
(протокол № 8 від 12 жовтня 2009 року)*

Методичні вказівки до виконання практичних робіт з дисципліни «Математичні методи в гідрометеорології» для студентів географічного факультету / Упорядник О.І.Лук'янець. – К.: Видавничо-поліграфічний центр «Київський університет», 2010.

ВСТУП

Практичні заняття з дисципліни «Математичні методи в гідрометеорології» повинні сприяти закріпленню у студентів основних теоретичних положень математичних методів, які використовуються в гідрометеорологічних дослідженнях, і засвоєнню навичок їх практичного застосування.

Математичні методи в гідрометеорології подані в основному математичною статистикою, яка побудована на теорії ймовірності. Якщо остання є однією з математичних наук, яка досліджує абстрактні взаємовідношення ймовірностей випадкових подій, то математична статистика має більш прикладний характер. Вона може бути визначена як розділ математики, який присвячений методам збору, обробки, систематизації, інтерпретації та використання даних для інженерних та наукових висновків. В першу чергу це стосується кількісних даних. Звідси можна зробити важливий висновок, що математична статистика зі всіма своїми теоретичними положеннями, визначеннями та методами має застосування у будь-якої галузі науки й в інших галузях людської діяльності. Вона широко застосовується й у гідрометеорології, і, зокрема, у гідрології. Завжди необхідна розумна перевірка істинності того, що статистичні методи придатні і пристосовані для рішення кожного конкретного гідрологічного завдання.

У центрі уваги буде поняття «ймовірність». Ймовірність, у загальному випадку, – числова характеристика можливості настання якої-небудь визначеної події.

Основним завданням застосування статистичних методів було і залишається завдання побудови кривих розподілу гідрометеорологічних величин. Традиційно це робиться за наступною схемою:

1. Хронологічний ряд спостережень за випадковою змінною величиною розташовується (упорядковується) у спадному порядку, тобто компонується варіаційний ряд (послідовність порядкових статистик).
2. Кожному члену варіаційного ряду ставиться у відповідність (назначається, розраховується) емпірична ймовірність (відносна частота появи).
3. За даними спостережень обчислюються статистичні характеристики емпіричного розподілу та функції, які використовуються для наближеного визначення значень параметрів аналітичного закону розподілу (правдоподібності, моментні, квантильні). Виходячи з характеру емпіричної функції розподілу вибирається відповідний вид апроксимуючої аналітичної функції розподілу.
4. Зрештою оцінюється закон розподілу (вибір аналітичного виразу та оцінка параметрів), а також проводиться оцінка графічного порівняння емпіричної та аналітичної функцій розподілу.

Застосування методів математичної статистики та теорії ймовірності у гідрометеорології має ряд особливостей, які обумовлені специфічністю гідрологічних і метеорологічних явищ та процесів, і ускладнюється супутніми додатковими проблемами:

- обмеженістю інформації (недостатністю довжини рядів спостережень), яка істотно не може бути збільшена;
- ряди спостережень неоднорідні як у просторі, так і в часі. Тому перед початком статистичного оброблення необхідно провести ретельний аналіз вихідної інформації на його генетичну й статистичну однорідність;
- для гідрологічних і метеорологічних характеристик має місце внутрішньорядна зв'язаність. Наявність такої зв'язаності порушує принцип випадковості відбору, призводить до зменшення обсягу незалежної інформації, що збільшує зміщуваність та нестійкість оцінок досліджуваних рядів.

Гідрометеорологічні процеси та явища є результатом взаємодії великої кількості факторів, різних за своїм походженням, часом впливу й часткою участі у формуванні досліджуваних процесів і явищ, тобто вони багатофакторні. У зв'язку із цим між ними й зумовлюючими їх факторами існують не функціональні, а статистичні зв'язки, серед яких особливе місце належить кореляційним зв'язкам як частковій формі статистичних. Кореляційні зв'язки (лінійна кореляція між двома змінними, множинна кореляція, автокореляція й т.д.) виражаються у вигляді кореляційних рівнянь або рівнянь регресії. Показником тісноти таких зв’язків є коефіцієнт кореляції (для лінійних зв’язків) або кореляційне відношення (для нелінійних).

Склад та зміст практичних робіт за цією дисципліною визначені виходячи з того, що статистична обробка даних та теорія ймовірності більш детально й повно розглядаються у спеціальних лекційних курсах «Математичні методи гідрологічних і гідрохімічних досліджень», «Річковий стік та гідрологічні розрахунки», «Гідрологічні прогнози». Тут ставиться завдання познайомити студентів з початковими прийомами математичної обробки гідрометеорологічної інформації, з методами математичної статистики та теорії ймовірності, з послідовністю статистичного дослідження.

Студенти повинні добре уявляти сутність гідрологічних явищ, взаємопов’язаність і взаємообумовленість всіх природних факторів і процесів, які відбуваються на водозборах, і вміти аналізувати отримані результати розрахунків,вірно застосовувати різноманітній і потужній апарат математичної статистики та теорії імовірності в гідрометеорологічних дослідженнях.

Методичні вказівки вміщують рекомендації по виконанню 8 практичних завдань, кожне з яких включає опис необхідних вихідних матеріалів і вимоги для виконання завдання, пояснення, методичні рекомендації та порядок виконання.

До завдань додаються форми та зразки розрахункових таблиць, необхідні графіки, рисунки, а також необхідні довідкові таблиці.

Рекомендована література до всіх завдань подана у кінці посібника. Вихідні матеріали студенти виписують самостійно з видань Водного кадастру або отримують від викладача.

ПРАКТИЧНА РОБОТА №1
Способи узагальнення статистичних даних.
Побудова гістограми та емпіричної кривої розподілу

Завдання: 1) вибрати, проаналізувати, скомпонувати необхідні для виконання практичної роботи дані, користуючись матеріалами видань Державного Водного кадастру; 2) обчислити характеристики емпіричного розподілу для інтервалів значень змінної величини; 3) побудувати гістограму розподілу та криву накопичених частот.

Пояснення. В основі статистичних висновків, які можна зробити про закономірності поведінки комплексу випадкових гідрологічних і метеорологічних величин, лежать ряди спостережень. При статистичній обробці рядів спостережень слід розрізняти два випадки. У першому випадку – відомі всі можливі значення змінної величини. Прикладом може слугувати обробка щоденних рівнів або витрат води за конкретний рік. У другому випадку, який найбільш часто зустрічається у гідрометеорологічних дослідженнях, – відома лише частина можливих значень змінної величини. У математичній статистиці такий ряд спостережень називають випадковою вибіркою значень змінної величини з генеральної сукупності, яка являє собою сукупність всіх членів ряду, кількість яких наближається до нескінченості ($N \to \infty$). Очевидно, що випадкова вибірка не може вичерпно охарактеризувати всю генеральну сукупність. Наприклад, ряд спостережень за річним стоком довжиною 20 або навіть у 100 років ймовірно не буде включати екстремальні значення стоку, які могли мати місце за період у декілька століть і тим більш у декілька тисячоліть.

У першому випадку можна розрахувати істинні, дійсні статистичні параметри (характеристики) ряду спостережень, у другому – лише наближені статистичні оцінки цих параметрів.

Порядок виконання та методичні указівки. Для виконання завдань практичних робіт 1-7 необхідно сформувати репрезентативні, однорідні, безперервні послідовності значень змінної величини. Це, наприклад, можуть бути для обраної річки у замикальному створі середні річні витрати води, значення максимальних витрат води під час весняного водопілля, значення максимальних модулів стоку за періоди дощових паводків, середні витрати води під час літньої чи зимової межені, мінімальний стік та ін. Хоча за бажанням студента і після консультації з викладачем можна прийняти для статистичної обробки випадкову вибірку значень інших гідрологічних, гідрохімічних або метеорологічних величин. Всі дані у хронологічному порядку компонуються за формою таблиці 1.1.

Для зручності вибрані значення змінної величини розташовують у спадному порядку за формою таблиці 1.2.

З таблиці 1.2 вибирають найбільше $(Q_{íàéá.})$ і найменше $(Q_{íàéì\ ..})$ значення максимальної витрати води під час весняного водопілля та знаходять за формулою 1.1 між ними різницю ($\grave{A}_Q$), яка називається амплітудою коливання змінної величини або розмахом варіювання.

$$A_Q = Q_{íàéá.} - Q_{íàêì}\ . \tag{1.1}$$

Таблиця 1.1. Вихідні дані у хронологічному порядку.
Максимальні витрати води під час весняного водопілля
Річка ……….. – пост…………

№ пп.	Рік	Максимальні витрати води, м^3/с

Таблиця 1.2. Максимальні витрати води
під час весняного водопілля у спадному порядку
Річка ……….. – пост…………

№ пп.	Максимальні витрати води, м^3/с

Якщо для статистичної обробки були взяті ні одна величина, а два або три різних за характеристикою елемента метеорологічного чи гідрологічного режиму, то подальші розрахунки здійснюються окремо для кожної генетично однорідної послідовності значень змінної величини.

Отриману амплітуду коливань змінної величини ділять на інтервали (або градації) і підраховують кількість попадань ознаки варіювання у кожний з них. Частіше використовують рівні за величиною інтервали, кількість яких призначають в залежності від об’єму вибірки змінної величини і призначають так, щоб відобразити її типові риси. При цьому зі збільшенням довжини інтервалу кількість попадань змінної в кожний інтервал буде зростати, що збільшує статистичну надійність наданого матеріалу. Але при невеликому об’ємі спостережень і завеликій довжині інтервалу кількість градацій буде невеликою, й тоді може виявитися знівельованими типові риси того чи іншого ряду спостережень. При зменшенні довжини інтервалу кількість попадань у них буде зменшуватися і виникає небезпека появи закономірностей, які не властиві даному статистичному ряду. Для наближеної оцінки кількості інтервалів (градацій) використовується формула

$$n_x \le 5 \lg N\ , \tag{1.2}$$

де n_x – кількість інтервалів, N – кількість членів ряду (загальний об’єм вибірки).

Задані межі інтервалів повинні бути такими, щоб одне й те ж значення ряду змінної величини не попадало одразу у два інтервали. Для зручності всі розрахунки проводять у табличній формі (табл.1.3).

Таблиця 1.3. Згруповані дані максимальних витрат води під час весняного водопілля

Річка ……….. – пост…………

Характеристика	Інтервали								Примітка
	400-351	350-301	300-251	250-201	200-151	150-101	100-51	50-1	
1.Абсолютна частота (в кількості випадків) n									$\sum n = N$
2.Відносна частота (%)									$\sum \% = 100\%$
3.Абсолютна накопичена частота (в кількості випадків $\sum n$)									
4.Відносна накопичена частота ($\sum \%$)									
5. Абсолютна щільність									
6. Відносна щільність									

З ряду спостережень вибирають значення максимальних витрат води під час весняного водопілля, які попадають у той чи іншій інтервал, підраховують їх кількість у кожній градації і тим самим визначають абсолютну частоту появи змінної величини в інтервалі (у кількості випадків) або повторюваність. Ці показники заносять у перший рядок таблиці 1.3. Загальна сума абсолютних частот повинна дорівнювати кількості членів вибірки. Якщо виразити абсолютні частоти у відсотках від загальної кількості випадків, прийнятих за 100%, отримаємо розподіл відносних частот (рядок 2 таблиці 1.3).

Шляхом послідовного додавання абсолютних або відносних частот знаходять відповідно абсолютні (в кількості випадків) або відносні (у %) накопичені частоти – забезпеченість (рядки 3 та 4 таблиці 1.3). Як бачимо, накопичені частоти показують частоту появи значень досліджуваної величини у межах визначеного інтервалу та вище за нього.

Якщо поділити відносну або абсолютну частоту появи певної характеристики на довжину інтервалу, отримують відповідно відносну або абсолютну щільність розподілу (рядки 5 та 6 таблиці 1.3).

Складена таким чином таблиця називається таблицею емпіричного розподілу. Результати розрахунків таблиці 1.3, можна представити у графічному вигляді (рисунок 1.1), де по осі абсцис відкладають абсолютні та відносні частоти появи в інтервалі, а також накопичені частоти, а по осі ординат – прийняті градації максимальних витрат води (значення меж інтервалів).

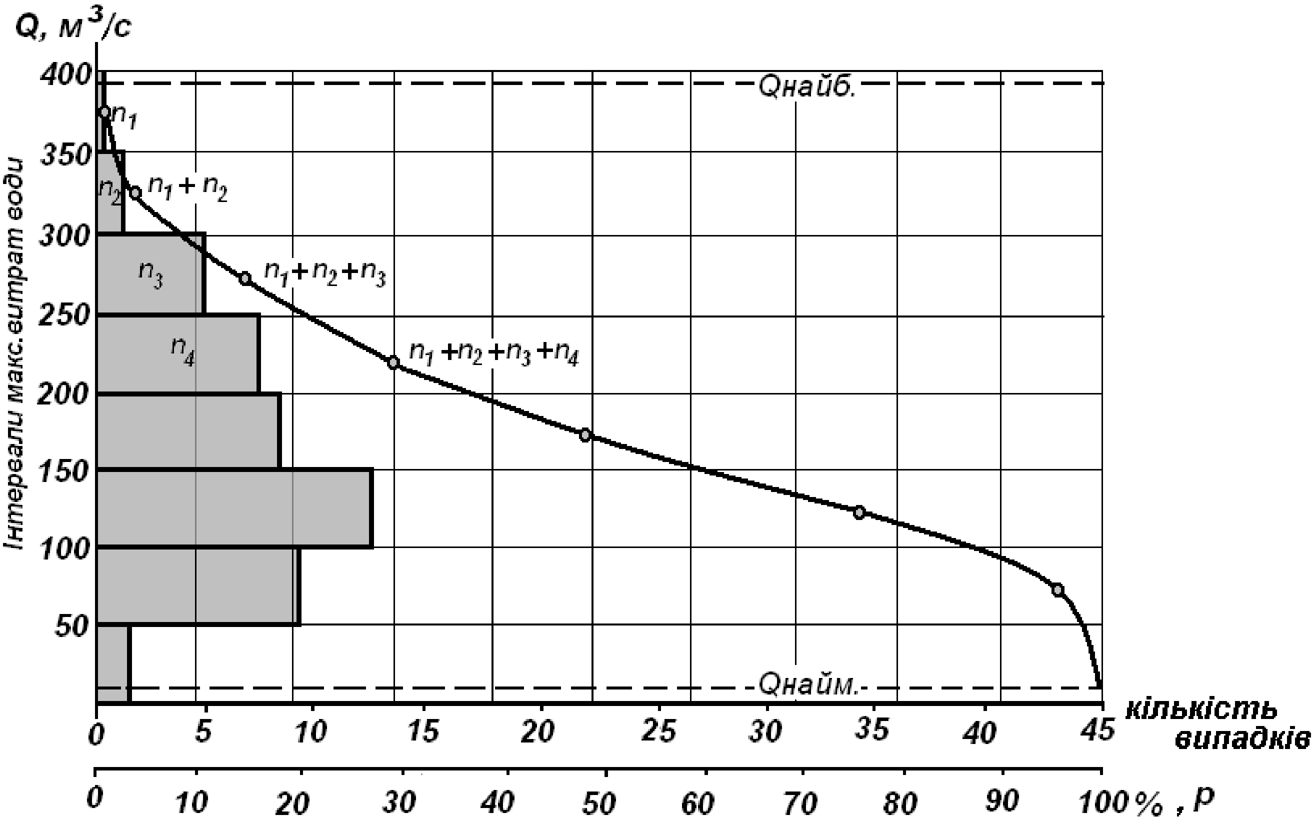


Рисунок 1.1. Приклад побудови гістограми розподілу та кривої накопичених частот (емпіричної кривої розподілу) максимальних витрат води під час весняної повені

Річка ………… – пост………….

Ступенева діаграма, яка представлена у вигляді суміжних прямокутників і побудована на одній прямій, називається гістограмою розподілу. Вона показує абсолютну або відносну частоту появи значень змінної величини в інтервалі. Крива абсолютних та відносних накопичених частот, яка представлена на графіку у вигляді плавної лінії, називається емпіричною кривою розподілу.

ПРАКТИЧНА РОБОТА №2
Визначення основних параметрів статистичного ряду

Завдання: 1) розрахувати основні статистичні параметри ряду значень змінної величини; 2) дати оцінку визначених параметрів, проаналізувати з результатами практичної роботи № 1.

Пояснення. Прості способи опрацювання статистичних даних, побудова гістограми та емпіричної кривої розподілу (практична робота №1) показують, що вже елементарні узагальнення дозволяють представити вихідну сукупність значень змінної величини в більш наочній та компактній формі. Одночасно можна відмітити, що розглянуті форми узагальнення статистичного матеріалу істотно різних метеорологічних та гідрологічних величин дозволяють виявити лише загальні статистичні закономірності і, разом з цим, розподіли деякіх з них мають свої індивідуальні особливості, які можуть бути описані за допомогою параметрів статистичного ряду. В якості характеристик, які описують статистичні закономірності рядів гідрометеорологічних даних і відповідно аналітичні криві розподілу, використовуються параметри статистичного ряду випадкової змінної величини, які характеризують її центр розподілу и ступінь її змінності.

Параметри статистичного ряду діляться:

- *на центри групування (розподілу),* до яких відносяться середня арифметична величина (норма) $\bar{x}$**,** медіана M_e**,** мода M_o**,** середня геометрична величина φ**,** середня гармонічна величина H;
- *міри розсіяння (варіювання)*, які включають амплітуду коливання змінної величини (розмах варіювання) A, середнє абсолютне відхилення d, дисперсію σ_x^2, середнє квадратичне відхилення (стандарт) σ_x, коефіцієнт варіації C_V, коефіцієнт асиметрії C_S.

Порядок виконання та методичні указівки. Якщо $x_1, x_2, x_3, ..., x_n$ – ряд спостережень випадкової величини x_i, то складена з них послідовність у спадному порядку називається *варіаційним рядом*, а кожний член цього ряду – *порядковою статистикою*.

<u>Центри групування (розподілу).</u> *Середня арифметична величина* – це центр, відносно якого розподіляються члени статистичної сукупності і розраховується як середнє значення з усіх членів ряду за формулою:

$$\bar{x} = \frac{1}{n}(x_1 + x_2 + ... + x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i \, , \tag{2.1}$$

де n - кількість членів ряду (від 1 до n), x_i - значення змінної величини.

Середнє арифметичне значення будь-якої гідрологічної характеристики, яке розраховане за тривалий період спостережень, називається *нормою***.** Середнє арифметичне є найбільш точною статистичною оцінкою істинного значення центра розподілу всієї генеральної сукупності – математичного сподівання.

Медіана – значення порядкової статистики варіаційного ряду, що займає серединне (центральне) положення. За непарної кількості членів ($n = 2m+1$) медіаною є x_{m+1}- член варіаційного ряду, тобто

$$M_e = x_{m+1} \, . \tag{2.2}$$

При парної кількості членів ($n = 2m$) медіаною є середнє значення між двома центральними значеннями варіаційного ряду:

$$M_e = \frac{1}{2}(x_m + x_{m+1}) \, . \tag{2.3}$$

Мода – це найбільш ймовірне значення випадкової величини ряду, що відповідає максимуму щільності розподілу. Функції щільності розподілів можуть бути одномодальними і полімодальними (мультимодальними, багатомодальными). Для помірно асиметричних рядів та одномодальних розподілів мода може бути розрахована за формулою К.Пірсона

$$M_o = \bar{x} + 3(M_e - \bar{x}). \tag{2.4}$$

Для симетричних розподілів – $\bar{x} = M_o = M_e$.

Норма $\bar{x}$, медіана M_e та мода M_o– три числові характеристики положення центра групування значень випадкової величини x_i.

Середня геометрична – це

$$\varphi = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \ldots x_n}\ . \tag{2.5}$$

Для її визначення виконують логарифмування виразу (2.5) та отримують

$$\ell g\varphi = \frac{1}{n}(\ell gx_1 + \ell gx_2 + \ldots\ldots + \ell gx_i) = \frac{1}{n}\sum_{i-1}^{n}\ell gx_i\ . \tag{2.6}$$

Звичайно середня геометрична менше середньої арифметичної ряду.

Як однією з форм трансформації величин вихідного ряду використовують перетворення сукупності величин $x_1, x_2, x_3, \ldots, x_n$ у ряд такого вигляду $\frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{x_3}, \ldots, \frac{1}{x_n}$. Середня арифметична трансформованого ряду відповідно дорівнює

$$\frac{1}{H} = \frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + \ldots \frac{1}{x_n}\right) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}. \tag{2.7}$$

Тоді $H = n\Big/\sum_{i=1}^{n}\frac{1}{x_i}$.

Величина H, обернене значення якої дорівнює середній арифметичній із обернених значень змінної x_i, називається *середньою гармонічною* варіаційного ряду.

Міри розсіяння (варіювання). Найбільш простою мірою варіювання статистичного ряду змінної величини x_i є її *амплітуда* коливання (або *розмах варіювання*), яка визначається за формулою (1.1).

Наступні параметри відносяться до параметрів, які відображають мінливість значень випадкової величини щодо центра групування.

Середнє абсолютне відхилення випадкової величини – це середнє арифметичне з модулів відхилень змінної величини від її норми. Воно розраховується за формулою

$$d = \frac{\sum_{i=1}^{n}\left|(x_i - \bar{x})\right|}{n}, \tag{2.8}$$

де n – кількість членів ряду x_i (від 1 до n), $\bar{x}$ - норма варіаційного ряду.

Дисперсія випадкової величини – це середнє арифметичне значення з суми квадратів відхилень змінної величини x_i від її норми $\bar{x}$ (при цьому $(x_i - \bar{x}) = \Delta x_i$)

$$\sigma_x^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n}\Delta x_i^2}{n}. \qquad (2.9)$$

Для характеристики ступеня мінливості змінної величини використовують *середнє квадратичне відхилення випадкової величини*. Цей параметр у статистиці ще називають *стандартом* і він розраховується за формулою:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}\Delta x_i^2}{n}}. \qquad (2.10)$$

Якщо ряд спостережень менше 30 років, то при визначенні дисперсії та середнього квадратичного відхилення випадкової величини в формулах (2.9) і (2.10) у знаменнику від кількості значень змінної величини n слід відняти одиницю, тобто відповідно формули мають вигляд

$$\sigma_x^2 = \frac{\sum_{i=1}^{n}\Delta x_i^2}{n-1}, \qquad (2.11)$$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n}\Delta x_i^2}{n-1}}. \qquad (2.12)$$

Цим самим досягається найбільше наближення цих параметрів ряду змінної величини до генеральної сукупності (усувається зміщення). Дисперсію и середнє квадратичне відхилення зручно розраховувати в табличній формі (таблиця 2.1). Якщо для статистичної обробки були взяті декілька різних за характеристикою елемента гідрологічного режиму, то подальші розрахунки здійснюється для кожної генетично однорідної послідовності окремо.

Таблиця 2.1. Розрахунок дисперсії та середнього квадратичного відхилення (стандарту) максимальних витрати води під час весняного водопілля

Річка ……….. – пост…………

№ пп.	*Рік*	Максимальні витрати води під час весняного водопілля, $Q_{i,ìàêñ}$	$\Delta Q = Q_{i,ìàêñ} - \overline{Q}_{ìàêñ}$	ΔQ^2

У розрахунках часто виникає необхідність порівняння ступеня мінливості рядів, які утворені з істотно різних характеристик. У цьому випадку середні квадратичні відхилення, хоча й характеризують варіювання конкретного ряду, але виявляються такими, які не можна порівняти, зіставити для різних рядів значень змінних величин. Співставлення мінливості подібних рядів здійснюється за допомогою *коефіцієнту варіації* (C_V), який являє собою відношення середнього квадратичного відхилення значень змінної величини до середнього арифметичного (тобто відношення стандарту до норми):

$$C_V = \frac{\sigma_x}{\overline{x}}. \qquad (2.13)$$

Коефіцієнт варіації є безрозмірним параметром мінливості статистичного ряду. Іноді коефіцієнт варіації виражають у відсотках від норми $C_V = \frac{\sigma_x}{\overline{x}} \cdot 100\%$.

Коефіцієнт варіації можна визначити й через модульні коефіцієнти за наступною формулою

$$C_V = \sqrt{\frac{\sum_{i=1}^{n}(K_i - 1)^2}{n}}, \qquad (2.14)$$

де $K_i = \frac{x_i}{\overline{x}}$ - модульні коефіцієнти значень змінної величини.

Як міра розсіяння та показник симетричності (асиметричності) ряду приймається середнє значення кубів відхилень значень змінної величини ряду від норми:

$$\mu_3 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{n} = \frac{\sum_{i=1}^{n}\Delta x_i^3}{n}. \qquad (2.15)$$

Коли члени ряду змінної величини розташовуються симетрично відносно норми, різні за величиною додатні та від’ємні відхилення від середнього повторюються однаково часто. Тобто, якщо ряд симетричний, то відхилення, які виходять з різними знаками, взаємно урівноважуються та їх сума дорівнює нулю. В інших випадках додатні відхилення (наприклад, роки з дуже високими максимальними витратами води під час весняного водопілля) можуть повторюватися рідше, чим від’ємні, в цьому випадку значення моди змінної величини виявляється менше норми. Це випадок додатної асиметрії. У протилежному випадку можна спостерігати від’ємну асиметрію і результат буде мати від’ємний знак.

Щоб отримати безрозмірний вираз для характеристики асиметрії ряду, середнє значення кубів відхилень ділять на куб стандарту. Це відношення називається *коефіцієнтом асиметрії* та обчислюється за формулою

$$C_S = \frac{\sum_{i=1}^{n}\Delta x_i^3}{n{\sigma_x}^3}. \qquad (2.16)$$

Якщо у виразі (2.16) замінити $\sigma_x = C_V \cdot \overline{x}$, отримаємо

$$C_S = \frac{\sum_{i=1}^{n}(K_i - 1)^3}{nC_V^3}. \qquad (2.17)$$

Найбільш зручно розраховувати коефіцієнти варіації та асиметрії за формою таблиці 2.2.

Таблиця 2.2. Розрахунок коефіцієнтів варіації та асиметрії максимальних витрат води під час весняного водопілля

Річка ……….. – пост…………

№ пп.	Максимальні витрати води під час весняного водопілля ($Q_{i,\text{ìàêñ}}$) у спадному порядку	$K_i = Q_{i,\text{ìàêñ}} / \overline{Q}_{\text{ìàêñ}}$	$K_i - 1$	$(K_i - 1)^2$	$(K_i - 1)^3$

Всі отримані параметри статистичного ряду значень змінної величини бажано об’єднати в одній таблиці (таблиця 2.3).

Таблиця 2.3.Основні статистичні параметри максимальних витрат води під час весняного водопілля

Річка ……….. – пост………….., період спостережень – з … по … рр.

Параметри статистичного ряду	Значення статистичних параметрів максимальних витрат води під час весняного водопілля
Кількість членів варіаційного ряду (n)	
Найбільше значення ряду	
Найменше значення ряду	
Розмах варіювання (A)	
Середня арифметична величина (норма) ($\overline{x}$)	
Медіана (M_e)	
Мода (M_o)	
Середня геометрична (φ)	
Середня гармонічна (H)	
Середнє абсолютне відхилення (d)	
Дисперсія (σ_x^2)	
Середнє квадратичне відхилення (стандарт) (σ_x)	
Коефіцієнт варіації (C_V)	
Коефіцієнт асиметрії (C_S)	

ПРАКТИЧНА РОБОТА №3

Визначення емпіричної ймовірності перевищення (забезпеченості) членів варіаційного ряду за формулами різних авторів

Завдання: 1) обчислити емпіричні ймовірності перевищення (забезпеченості) кожного члена варіаційного ряду за формулами різних авторів; 2) скласти узагальнену таблицю розрахунків, проаналізувати результати, обґрунтувати вибір формули визначення емпіричної ймовірності перевищення (забезпеченості) кожного члена варіаційного ряду для подальшого дослідження.

Пояснення. Емпірична крива розподілу (практична робота №1, табл.1.3) була отримана шляхом послідовного сумування відносних частот або, що теж саме, емпіричних ймовірностей. Однак подібні побудови (рис.1.1) можливі лише у випадку наявності статистичної сукупності достатньо великого об’єму і вони дають уявлення про розподіл досліджуваних значень змінної величини в межах визначених інтервалів. Якщо сукупність випадкових величин включає не більше кілька десятків членів, групування їх по градаціям є практично неможливим завданням. Тому при узагальненні рядів такого об’єму використовується інший прийом розрахунку та побудови емпіричної кривої. При цьому члени статистичної сукупності розташовуються у спадному порядку – $x_1 > x_2 > x_3 .... > x_m > .... > x_n$, де m змінюється від 1 до n. Тоді теоретична ймовірність перевищення кожного члена варіаційного ряду при $n \to \infty$ виражається за формулою $P = \ell im(\frac{m}{n})_{n\to\infty}$, яка, звичайно, невідома, тому що відсутні вибірки з генеральної сукупності великого об’єму для гідрометеорологічних даних, коли би $n \to \infty$.

Тому при побудові емпіричної кривої використовують емпіричну ймовірність перевищення кожного значення змінної величини варіаційного ряду P_m, яка у першому наближенні розраховується за формулою

$$P_m = \frac{m}{n} \cdot 100\%, \tag{3.1}$$

де m – порядковий номер члена варіаційного ряду, n – кількість членів цієї вибірки. У цьому випадку при кінцевому значенні n отримуємо оцінку ймовірності перевищення з деякою систематичною похибкою. Формула емпіричної забезпеченості (3.1) дає прийнятні результати при не дуже малому n. Для членів сукупності, які займають останнє місце у варіаційному ряду випадкової змінної, при будь-якому кінцевому значенні n завжди $Pm = 100\,\%$, а для першого члену ряду $Pm = 1/n$, що звичайно є досить грубою оцінкою.

Однак є набагато більше повноцінний спосіб оцінки емпіричного розподілу при цьому рекомендується для використання та же формула (3.1), але тільки після наступного перетворення варіаційного ряду:

вихідний ряд: $x_1 \geq x_2 \geq x_3 \geq ... x_{n-1} \geq x_n$

перетворений ряд: $x_{1,2} \geq x_{2,3} \geq x_{3,4} \geq ... x_{n-1,n}$,

де $x_{1,2} = (x_1 + x_2)/2$, $x_{2,3} = (x_2 + x_3)/2 \ldots$.

Варто мати на увазі, що в перетвореному ряду кількість його членів на одиницю менше, ніж у вихідному. Але можливо (і доцільно) додавання двох додаткових членів – максимального x_1 і мінімального x_n з точними індивідуальними оцінками їхніх ймовірностей:

$$P_1 = 1 - \exp(-\frac{\ln 2}{n}), \tag{3.2}$$

$$P_n = \exp(-\frac{\ln 2}{n}). \tag{3.3}$$

Порядок виконання та методичні указівки. Для отримання найбільшого наближення емпіричної оцінки ймовірності перевищення до аналітичного його значення з врахуванням мінливості змінних випадкових величин варіаційного ряду запропоновано декілька формул розрахунку емпіричного розподілу членів варіаційного ряду:

Формула Хазена

$$P_m = \frac{m - 0,5}{n} \cdot 100\%. \tag{3.4}$$

Формула С.Н.Крицького та М.Ф.Менкеля

$$P_m = \frac{m}{n+1} \cdot 100\%. \tag{3.5}$$

Формула Н.Н.Чагодаєва

$$P_m = \frac{m - 0,3}{n + 0,4} \cdot 100\%. \tag{3.6}$$

Формула Е.Г.Блохинова

$$P_m = \frac{m - 0,4}{n + 0,2} \cdot 100\%. \tag{3.7}$$

При об’ємі вибірки 20÷70 членів і при $C_S = 2C_V$ рекомендується застосовувати формулу (3.6), при $C_S < 2C_V$ – формулу (3.7), при $C_S > 2C_V$ – формулу (3.4).

У розрахунках максимальних значень випадкових величин при будь-яких співвідношеннях C_S / C_V слід використовувати формулу (3.5), тому що вона заснована на уявленні про те, що обраний період визначається серед генеральної сукупності підвищеною ймовірністю високих значень і пониженою ймовірністю низьких значень певної величини.

Розраховані за формулами різних авторів значення ймовірностей для подальшого аналізу слід занести до таблиці 3.1.

Таблиця 3.1. Емпіричні ймовірності перевищення максимальних витрат води під час весняного водопілля, розрахованих за формулами різних авторів

Річка ……….. – пост………….

№ пп.	Модульні коефіцієнти K_p максимальні витрати води під час весняного водопілля у спадному порядку	Ймовірності перевищення (забезпеченість)				
		$P_m = \frac{m}{n} \cdot 100\%.$	Формула Хазена $P_m = \frac{m-0,5}{n} \cdot 100\%$	Формула Крицького та Менкеля $P_m = \frac{m}{n+1} \cdot 100\%$	Формула Чагодаєва $P_m = \frac{m-0,3}{n+0,4} \cdot 100\%$	Формула Блохинова $P_m = \frac{m-0,4}{n+0,2} \cdot 100\%$

ПРАКТИЧНА РОБОТА №4
Методи розрахунку параметрів аналітичних кривих розподілу при різних значеннях коефіцієнтів варіації і асиметрії (C_v і C_s)

Завдання: 1) визначити основні статистичні параметри для послідовності випадкових величин за методами моментів, найбільшої правдоподібності, графоаналітичним; 2) розрахувати ординати аналітичних кривих розподілу ймовірностей.

Пояснення. Емпірична функція розподілу як реалізований природою й виміряний людиною прообраз істинного розподілу щодо генеральної сукупності, який має у своїй основі варіаційний ряд як упорядковану послідовність спостережень за деякою випадковою величиною, має максимально можливу інформацію, що міститься в цих спостереженнях. Найбільш елементарним, але в той же час і фундаментальним завданням математичної статистики є знаходження для емпіричного розподілу відповідної йому аналітичної кривої (функції) розподілу.

Для розрахункових цілей емпіричний розподіл необхідно апроксимувати підходящими аналітичними виразами. Останні – аналітичні криві (функції) розподілу – є сполучною ланкою між самими загальними уявленнями про гідрометеорологічні процеси і явища та властивостями емпіричних даних.

Процедура використання аналітичних розподілів тривіальна. Її звичайно йменують згладжуванням або підбором аналітичної кривої до спостережених частот. В основі загальноприйнятої схеми оцінки параметрів аналітичної кривої розподілу лежить ідея зіставлення числових характеристик величин рядів спостережень і аналітичних виражень цих же характеристик. Розходження в методах оцінки – лише в сутності останніх і в способах їхнього отримання. Кількість обраних характеристик звичайно приймається рівним числу параметрів аналітичної кривої розподілу.

Порядок виконання та методичні указівки. Для побудови аналітичної кривої розподілу випадкових величин, яка б відповідала емпіричній, необхідно визначити три основних параметра: середнє арифметичне ряду, коефіцієнти варіації і асиметрії. Ці параметри можуть розраховуватися за *методом моментів* (при $C_V \le 0{,}5$), *найбільшої правдоподібності* (при $C_V \ge 0{,}5$) та *графоаналітичним (через квантилі)* за будь-яких значень коефіцієнта варіації.

Метод моментів. При невеликій мінливості ряду найбільше поширення у практиці отримав метод моментів, суть якого полягає в тому, що добирається така аналітична крива, у якої другий та третій центральні моменти дорівнюють відповідним моментам емпіричного розподілу. Параметри цієї кривої – $\bar{x}$, C_V, C_S (середнє арифметичне ряду, коефіцієнти варіації та асиметрії) визначаються за вихідними даними таблиці 2.2 та за формулами (2.1), (2.14), (2.17).

У зв'язку з тим, що оцінювання параметрів здійснюється за обмеженими рядами спостережених величин, які є вибірками з генеральної сукупності, необхідно знати точність розрахованих характеристик. Похибки у визначенні параметрів емпіричного розподілу обчислюються за формулами:

$$E_{\bar{x}} = \pm \frac{C_V}{\sqrt{n}} \cdot 100\%, \tag{4.1}$$

$$Ec_V = \pm \sqrt{\frac{1 + {C_V}^2}{2n}} \cdot 100\%, \tag{4.2}$$

де $E_{\bar{x}}$, Ec_V - відносні середні квадратичні похибки відповідно середнього арифметичного значення ряду (формула 4.1) та коефіцієнта варіації (формула 4.2), C_V - коефіцієнт варіації вибірки, n – кількість членів ряду.

Відносна середня квадратична похибка коефіцієнта варіації є показником репрезентативності ряду спостережень і, якщо вона $\le$ 10-15%, то тривалість ряду достатня для визначення C_V.

При визначенні коефіцієнта асиметрії за короткими рядами спостережень можливі великі похибки, тому частіше C_S визначається шляхом підбору його значення, виходячи з найкращої відповідності емпіричних точок аналітичній кривій розподілу.

Для визначення ординат аналітичної кривої розподілу $x_p = f(p)$ за методом моментів використовують біноміальну асиметричну криву розподілу, при цьому встановлюються нормовані відхилення від середнього значення ординат розподілу Пірсона III типу при C_v=1,0 (таблиця А.1, додаток А):

$$\phi(x) = f(C_S, P) = \frac{K_p - 1}{C_V},$$

Звідси отримуємо

$$K_p = C_V \phi + 1, \tag{4.3}$$

де ϕ – відхилення ординат кривої забезпеченості від середини при даному значенні C_V в залежності від величини C_S; K_p – модульний коефіцієнт заданої ймовірності перевищення, C_V - коефіцієнт варіації.

Розрахункові значення величин x_p заданих ймовірностей перевищення (забезпеченостей) визначаємо за формулою

$$x_p = \bar{x} K_p, \tag{4.4}$$

де $\bar{x}$ – середнє арифметичне ряду.

Розрахунок ординат аналітичної кривої розподілу ведеться в табличній формі (таблиця 4.1).

Таблиця 4.1. Розрахунок ординат аналітичної кривої розподілу максимальних весняних витрат води за методом моментів

Річка ……….. – пост…………

при $\bar{x} =$ $C_V =$ $C_S =$

Порядок розрахунку	Ймовірність перевищення (забезпеченість), %													
	0.01	0.1	1.0	3.0	5.0	10	25	50	75	90	95	97	99	99.9
ϕ														
$C_V \phi$														
$K_p = C_V \phi + 1$														
$x_p = \bar{x} K_p$														

Метод найбільшої правдоподібності. Метод найбільшої правдоподібності полягає в тому, що за оцінку невідомого параметру приймають таке значення, при якому функція правдоподібності (добуток ймовірностей спостережених величин) досягає найбільшого можливого значення. За цим методом розрахункові значення коефіцієнтів варіації C_V та асиметрії C_S визначаються у залежності від статистик λ_2 та λ_3 за допомогою номограм ([4], додаток 1). На осі абсцис номограм відображена статистика λ_3, на осі ординат – λ_2. Точка перетину статистик й визначає параметри C_V та C_S. Значення самих статистик λ_2 та λ_3 обчислюються за формулами:

$$\lambda_2 = \frac{\sum_{i=1}^{n} \lg K_i}{n-1}, \tag{4.5}$$

$$\lambda_3 = \frac{\sum_{i=1}^{n} K_i \lg K_i}{n-1}, \tag{4.6}$$

Розрахунок статистик ведеться за формою таблицею 4.2.

Таблиця 4.2. Розрахунок статистик λ_2 та λ_3 для визначення коефіцієнтів варіації C_V та асиметрії C_S за методом найбільшої правдоподібності

№ пп.	Рік	Q_i (у спадному порядку)	$K_i = Q_i / \overline{Q}$	$\lg K_i$	$K_i \lg K_i$
			$\sum K_i = n$		

Відносна середня квадратична похибка коефіцієнта варіації, визначеного методом найбільшої правдоподібності, розраховується за формулою:

$$E_{C_V} = \pm\sqrt{\frac{3}{2n(3 + C_V{}^2)}} \cdot 100\%. \tag{4.7}$$

Згідно отриманим значенням статистик λ_2 і λ_3 та визначеним за допомогою номограм значень коефіцієнтів варіації C_V та асиметрії C_S, ординати теоретичної кривої забезпеченості визначаються безпосередньо за значенням C_V і співвідношенням C_S / C_V за таблицею ординат кривих трипараметричного гамма-розподілу (таблиця А.2, додаток А).

Ряди вважаються репрезентативними, а результати оцінювання можуть бути прийнятними як розрахункові характеристики, якщо середня квадратична похибка визначення середньої арифметичної знаходиться у межах 5÷10%, а коефіцієнта варіації – 10÷15%.

Графоаналітичний метод. Графоаналітичний метод дозволяє встановлювати параметри аналітичної кривої розподілу досліджуваних характеристик безпосередньо за згладженою емпіричною кривою. При використанні графоаналітичного методу в якості основної вихідної приймається умова збігу теоретичної кривої з емпіричною щонайменше у трьох точках.

За значеннями модульних коефіцієнтів та їх емпіричними ймовірностями (таблиці 2.2 або 3.1) на клітчатці ймовірностей будують згладжену емпіричну криву забезпеченості. На ній визначають три опорні точки – у верхній частині кривої точку із ймовірністю перевищення 5%, у середній –50%, в нижній – 95%. Тобто, це будуть ординати $X_{5\%}, X_{50\%}, X_{95\%}$. Для кожної з ординат, через яку повинна проходить аналітична крива розподілу, використовуючи формули (2.13) та (4.4), можна записати наступні три рівняння

$$X_{5\%} = \overline{x} + \sigma\phi_{5\%}, \tag{4.8}$$

$$X_{50\%} = \overline{x} + \sigma\phi_{50\%}, \tag{4.9}$$

$$X_{95\%} = \bar{x} + \sigma\phi_{95\%}. \qquad (4.10)$$

У цих рівняннях невідомими є три параметри – норма $\bar{x}$, середнє квадратичне відхилення (стандарт) σ та коефіцієнт асиметрії C_S, які треба визначити. У подальшому розрахункові параметри позначимо – $\bar{x}^*$, σ^*, C_S^*.

Для визначення коефіцієнта асиметрії C_S^* скористуємося величиною коефіцієнта скісності (S). Коефіцієнт асиметрії входить у ці рівняння через величину $\phi = f(C_s)$. Значення ординат $X_{5\%}, X_{50\%}, X_{95\%}$ оцінюються безпосередньо за згладженою емпіричною кривою забезпеченості, яка достатньо добре повинна відповідати розташуванню емпіричних точок. За цими ординатами знаходять коефіцієнт скісності скошеності $S = f(C_S)$.

$$S = \frac{X_{5\%} + X_{95\%} - 2X_{50\%}}{X_{5\%} - X_{95\%}}. \qquad (4.11)$$

За розрахованим значенням S знаходимо значення C_S^* за таблицею А.3 (додаток А).

У зв'язку з тим, що норма і середнє квадратичне відхилення (стандарт) мають розмірність вихідного ряду спостережень, то для їх визначення у нашому випадку використовуємо значення не модульних коефіцієнтів, а максимальних витрат води зазначених ймовірностей перевищення, які можна розрахувати за формулою (4.4). Віднімаючи з рівняння (4.8) рівняння (4.10) та вирішуючи різницю відносно σ, маємо

$$\sigma* = \frac{X_{5\%} - X_{95\%}}{\phi_{5\%} - \phi_{95\%}}. \qquad (4.12)$$

Різницю нормованих відхилень ординат кривої забезпеченості від середини ($\phi_{5\%}$– $\phi_{95\%}$) (5% та 95% ймовірності перевищення) визначаємо за таблицею А.3 (додаток А).

Уточнене значення середнього арифметичного $\bar{x}^*$ знаходять за рівнянням (4.9).

$$\bar{x}^* = X_{50\%} - \sigma^*\phi_{50\%} \qquad (4.13)$$

У формулах (4.12) і (4.13) $\phi_{5\%}$, $\phi_{50\%}$, $\phi_{95\%}$ – нормовані ординати біноміальної кривої розподілу, які відповідають встановленому значенню коефіцієнта S.

На основі отриманих значень σ^* та $\bar{x}^*$ оцінюють також й коефіцієнт варіації C_V^* за рівнянням (2.13). Маючи обчислені параметри аналітичної кривої ($\bar{x}^*, C_V^*$, C_S^*) за схемою таблиці 4.1, знаходимо її ординати. У графоаналітичному методі аналітична крива розподілу достатньою мірою відповідає емпіричному, якщо виконується нерівність

$$|\,\bar{x} - \bar{x}^*| \le 0{,}02\bar{x}, \qquad (4.14)$$

де $\bar{x}$ – середнє арифметичне ряду, яке отримано за рівнянням (2.1), а $\bar{x}^*$ – по рівнянню (4.13).

Якщо в результаті розрахунків будь-яким методом отримано від'ємне значення коефіцієнта асиметрії C_S<0 , що часто буває при характеристиці найбільших рівнів або витрат води різного походження, то використовують таблицю нормованих відхилень ординат біноміальної кривої забезпеченості від середини у дзеркальному відображенні (таблиця А.1, додаток А)., тобто приведені в таблиці значення відповідають ймовірностям, які доповнюються до 100%, та беруться зі зворотним знаком. Так, ордината, яка відповідає ймовірності перевищення 1%, при додатної асиметрії стає ординатою ймовірності 99% та приймається зі зворотнім (від'ємним) знаком.

ПРАКТИЧНА РОБОТА №5
Побудова емпіричної і аналітичних кривих розподілу. Визначення розрахункових характеристик заданої ймовірності перевищення

Завдання: 1) нанести на клітчатку ймовірностей перевищення емпіричний розподіл у вигляді системи точок; 2) побудувати на клітчатці разом з емпіричними точками відповідні їм апроксимуючі аналітичні криві розподілу, 3) візуально проаналізувати узгодженість аналітичних кривих розподілу до обрису емпіричного; 4) вибрати та обґрунтувати розрахункову аналітичну криву розподілу.

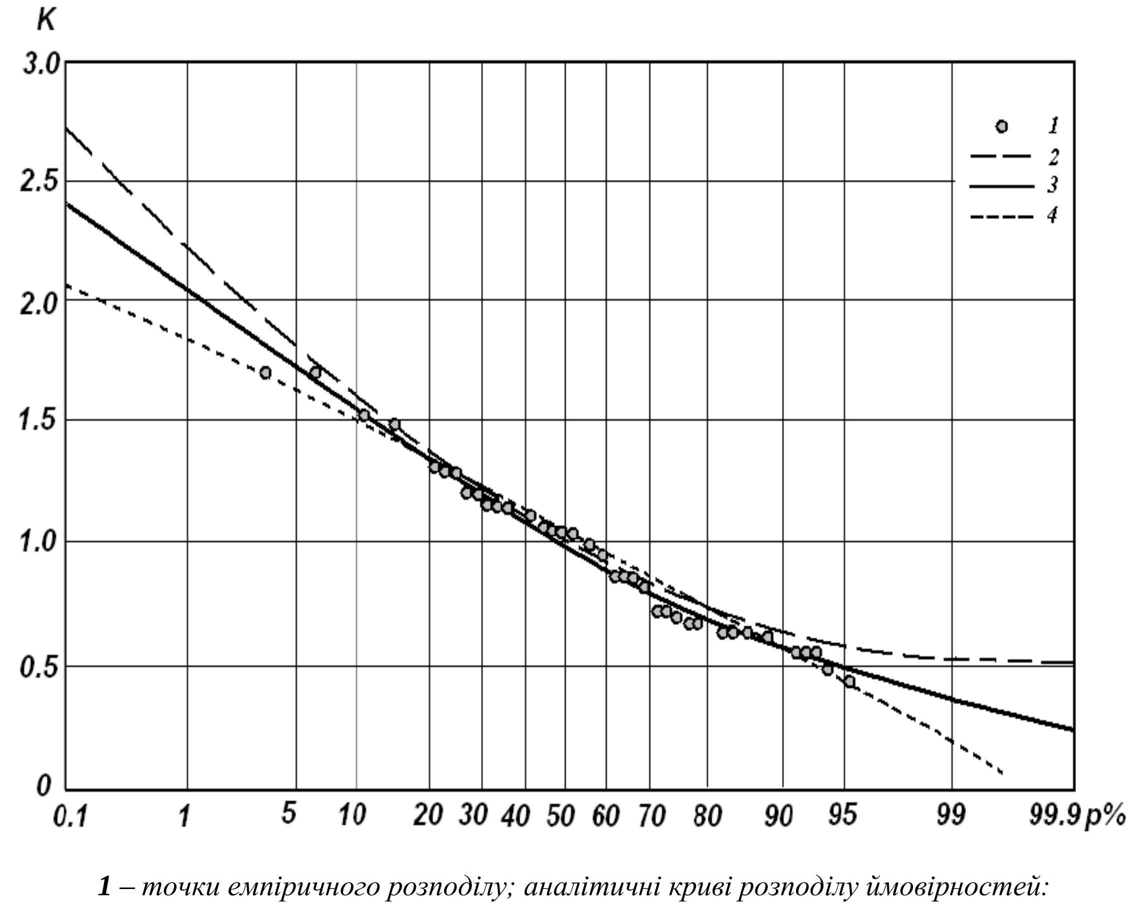
Пояснення. Найбільш елементарним, але в той же час і фундаментальним завданням математичної статистики є знаходження для емпіричної функції розподілу відповідної їй аналітичної функції. У реальності це зводиться до проведення останньої в найбільшій близькості до системи «точок» емпіричного розподілу. Рішення цього завдання ділиться на три етапи: 1) вибір підходящої за своїми межами та графічними обрисами аналітичної функції розподілу; 2) безпосередня оцінка параметрів останньої; 3) візуальна або чисельна оцінка задовільності рішення поставленого завдання – порівняння емпіричної й аналітичної функцій розподілу.

Емпіричні та аналітичні криві розподілу ймовірностей, які побудовані в прямокутній системі координат, мають складні випукло-вигнуті обриси. На кінцях цих кривих незначний приріст ймовірності викликає різке збільшення функції розподілу, що завдає труднощів або навіть робить практично неможливим графічне згладжування та особливо екстраполяцію цих кривих у межі малих та великих ймовірностей, які не охоплені даними фактичних спостережень. Для усунення цих технічних вад використовують спеціальні клітчатки ймовірностей перевищення: для кривих з помірною асиметричністю (шкала абсцис – логарифмічна, а ординат – рівномірна); для кривих зі значною асиметрією (обидві шкали – логарифмічні).

Побудовані на таких клітчатках криві легко апроксимуються у верхній та нижній частинах розподілу до заданих ймовірностей перевищення та мають плавні обриси (окреслення): за додатної асиметрії ($C_S > 0$) вони плавно-вигнуті, за від'ємної ($C_S < 0$) – плавно-випуклі, при $C_S = 0$ на напівлогарифмічній клітчатці крива перетворюється в пряму лінію.

Порядок виконання та методичні указівки. При виконанні практичної роботи №4 були визначені ординати аналітичних кривих розподілу за різними методами у вигляді модульних коефіцієнтів певної ймовірності перевищення. На клітчатку ймовірності (за помірної асиметрії – напівлогарифмічну, за значної – логарифмічну) по осі ординат відкладаються значення модульних коефіцієнтів, а по осі абсцис – значення ймовірностей перевищення. Вертикальний масштаб добирається таким чином, щоб увесь інтервал модульних коефіцієнтів, включно з рідкісною повторюваністю, розташовувався на одному аркуші клітчатки. У такому масштабі на цієї же клітчатці наносяться точки, які відповідають емпіричному розподілу модульних коефіцієнтів, отриманих за фактичними даними.

Всі розподіли й емпіричний, й аналітичні будуються на одній клітчатці, але позначаються різними умовними знаками (рисунок 5.1).



__1__ – точки емпіричного розподілу; аналітичні криві розподілу ймовірностей:

2 – C_V =0,35 C_S =1,4; ***3*** – C_V =0,35 C_S =2 C_V; ***4*** – C_V =0,35 C_S =0

Рисунок 5.1. Емпіричний та аналітичні криві розподілу ймовірностей

Це дає можливість краще провести візуальну оцінку графічного зіставлення емпіричної й аналітичної функцій розподілу. Графічна інтерполяція статистичних сукупностей на клітчатках дозволяє наочно пересвідчитися у відповідності прийнятої аналітичної моделі розподілу щодо емпіричному матеріалу, оцінити вплив окремих точок, які відхиляються від певної закономірності, на загальний вигляд розподілу.

Користуючись аналітичними кривими розподілу, легко обчислити розрахункову характеристику заданої ймовірності перевищення. Якщо значення модульного коефіцієнта певної ймовірності помножити на норму характеристики $\overline{x}$, отримаємо значення цієї характеристики певної ймовірності перевищення (формула 4.4).

Якщо ні одна з аналітичних кривих, ординати яких розраховані за різними методами, не співпадають з характером розподілу емпіричних точок, то приймають розрахункову криву, обриси якої найкращим чином відповідають емпіричному розподілу. Слід зауважити, що зі збільшенням співвідношення C_S/C_V збільшується прогин кривої, тому що збільшуються її крайні ординати та зменшуються ординати середньої частини. Чим більше C_S, тим крутіше верхня частина та більш полога нижня. При $C_S < 0$ крива має випуклу форму.

ПРАКТИЧНА РОБОТА №6
Оцінювання однорідності рядів спостережень за допомогою різних критеріїв

Завдання: 1) провести аналіз однорідності рядів спостережень (формулювання нульової і альтернативної гіпотез, визначення рівня значимості, вибір критичної області, бракування або прийняття нульової гіпотези); 2).визначити і оцінити однорідність рядів спостережень за допомогою різних критеріїв; 2) проаналізувати результати оцінки однорідності за параметричними та непараметричними критеріями.

Пояснення. Застосування аналітичних кривих розподілу для опису статистичних сукупностей можливо тільки у тому випадку, якщо ці сукупності сформовані з якісно однорідних та незалежних елементів. З’ясування статистичної однорідності сукупностей є важливим елементом оцінки достовірності статистичних узагальнень, тому що при використанні аналітичних кривих розподілу необхідно достатньо чітко уявляти наскільки прийнята теоретична схема узгоджується з емпіричним матеріалом. При відсутності надійної інформації про час та джерела, які порушують однорідність рядів спостережень, статистичні методи є єдиними для такої оцінки.

Під генетичною однорідністю будь-якої характеристики стоку розуміють не тільки те, що вона відноситься до однорідних фаз водного режиму, а також відносна незмінність на протязі всього часу, за який вона аналізується, фізико-географічних умов, водогосподарського використання, агролісомеліоративних заходів, методик спостережень та розрахунків гідрометеорологічних елементів. Насамперед при дослідженні однорідності рядів спостережень треба прийняти умову, що результати спостережень будемо вважати однорідними тоді, коли вони належать до однієї генеральної сукупності. При цьому всі спостереження будемо вважати незалежними як усередині рядів спостережень (інакше кажучи дотримана умова випадковості відбору), так й між рядами спостережень, що досліджуються.

У теорії ймовірностей відомо багато критеріїв однорідності, використовуючи які можна визначити однорідність вибіркових значень параметрів розподілу, зокрема середніх значень, дисперсій, або безпосередньо встановити належність декілька вибірок до однієї генеральної сукупності. Критерії однорідності діляться на дві групи – параметричні, які потребують знання закону розподілу (критерії Стьюдента, Фішера, Бартлета та ін.), та непараметричні (критерії Вількоксона, Ван-дер Вандера, Фішера-Йєтса, Клотца, критерій ω^2 та ін.).

Порядок виконання та методичні указівки. Статистичний аналіз однорідності рядів спостережень включає наступні основні етапи: формулювання нульової і альтернативної гіпотез, визначення рівня значимості, вибір критичної області, бракування або прийняття нульової гіпотези. Треба зауважити, що при аналізі однорідності рядів гідрологічних величин, які переважно мають асиметричний розподіл, як правило, застосовують непараметричні критерії. При аналізі метеорологічних величин, асиметричність яких незначна, застосовуються в основному параметричні критерії.

Критерій Z. Оцінку однорідності здійснюють для рядів, де вибіркові середні значення розподілені за нормальним законом або за наявності тривалих рядів спостережень, коли збільшення чи зменшення ряду на декілька членів істотно не змінить середнього арифметичного значення.

Одним з найбільш простих критеріїв є критерій Z. Для оцінювання на однорідність використовують два ряди: першій ряд, який містить n_x членів та характеризується середнім арифметичним $\bar{x}$; другий ряд, який містить n_y членів та характеризується середнім арифметичним ряду $\bar{y}$.

Нульова гіпотеза полягає у припущенні, що середні значення дорівняють одне одному $(\bar{x} = \bar{y})$, а альтернативна – в нерівності цих значень $(\bar{x} \neq \bar{y})$. За критичну область приймають область великих за абсолютним значенням відхилень. Рівень значимості дорівнює 5%.

Критерій однорідності визначають за формулою

$$Z = \left| \frac{\bar{y} - \bar{x}}{\sigma_{(\bar{y} - \bar{x})}} \right|, \tag{6.1}$$

де

$$\sigma_{(\bar{y}-\bar{x})} = \sqrt{\frac{\sigma_y^2}{n_y} + \frac{\sigma_x^2}{n_x}}\ , \qquad (6.2)$$

σ_y і σ_x - середні квадратичні відхилення рядів, що розглядаються.

Отримавши критерій однорідності для двох вибірок, знаходять область допустимих відхилень при даному рівні значимості, використовуючи таблицю нормального закону розподілу (таблиця Б.1, додаток Б). Якщо отримані значення критерію однорідності Z більше верхньої межі критичної області, то нульова гіпотеза відхиляється та приймається альтернативна.

Критерій Стьюдента. За коротких рядів спостережень (n < 30 років) при аналізі однорідності застосовується критерій t–розподілу Стьюдента. Його застосування правомірно для порівняння двох середніх значень з нормально розподілених сукупностей у тому випадку, якщо припустити, що дисперсії статистичних рядів, які розглядаються, дорівнюють одна одної $\sigma_x^2 = \sigma_y^2 = \sigma^2$, хоча невідомі. Критерій Стьюдента подається через статистику t у вигляді

$$t = \frac{\bar{y} - \bar{x}}{\sqrt{n_x \sigma_x{}^2 + n_y \sigma_y{}^2}} \cdot \sqrt{\frac{n_x n_y (n_x + n_y - 2)}{n_x + n_y}}, \qquad (6.3)$$

де $\bar{x}$ і $\bar{y}$ – середні значення вибірок, σ_x^2 та σ_y^2 – дисперсії вибірок, n_x і n_y – об'єми вибірок.

Цей критерій підкоряється розподілу Стьюдента за кількістю степенів вільності

$$K = n_x + n_y - 2. \qquad (6.4)$$

Як й у попередньому критерії, приймається нульова та альтернативна гіпотези. Критична область t–критерію встановлюється за рівнем значимості f та кількістю ступенів вільності K. Для перевірки за допомогою критерію Стьюдента визначається область при двосторонньому 5%-вом рівні значимості. Величина ($t_{f,K}$) добирається з таблиці розподілів Стьюдента (таблиця Б.2, додаток Б). Якщо при співставленні $|\,t\,| \le \left|t_{f,K}\right|$, то нульова гіпотеза приймається, а альтернативна відхиляється.

Критерії Z і Стьюдента є параметричними критеріями, тому що їх використання пов'язано з необхідністю прийняття для цієї конкретної вибірки нормального закону розподілу.

Критерії Вількоксона. Із непараметричних критеріїв однорідності широке застосування отримали критерії Вількоксона, які застосовуються для перевірки однорідності двох вибірок. Один з них, який ґрунтується на кількості інверсій, відомий як статистика u, другій базується на визначенні рангів і називається статистикою w.

Розрахунок *статистики u критерію Вількоксона* полягає у підрахунку кількості інверсій, які виявляються у результаті наступних процедур. Спостереження, які складають дві вибірки розташовуються в загальній послідовності в спадному порядку їх значень у вигляді

$$y_1 \quad x_1 \quad x_2 \quad y_2 \quad y_3 \quad y_4 \quad x_3 \quad y_5 \quad y_6 \quad x_4 \quad y_7 \quad y_8 \quad x_5 \quad x_6\,,$$

де $x_1,\ldots x_6$ – члени, які належать до першої вибірки, $y_1,\ldots y_8$ – члени другої вибірки.

Якщо будь-якому значенню x передує деяке значення y, то ця пара утворює інверсію. Так, у нашому випадку

x_1, x_2 утворюють по одній інверсії з y_1 ,

x_3 утворює чотири інверсії з y_1, y_2, y_3, y_4,

x_4 утворює шість інверсій з $y_1, y_2, y_3, y_4, y_5, y_6$,

x_5, x_6 утворюють по вісім інверсій з $y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8$.

Усього в цьому випадку буде $1+1+4+6+8+8=28$ інверсій.

Теоретично доведено, що при довжині рядів, які порівнюються, не менше ніж 10 членів, кількість інверсій розподілено приблизно за нормальним законом з їх математичним сподіванням

$$M_u = \frac{m \cdot n}{2} \qquad (6.5$$

і дисперсією

$$D_u = \frac{m \cdot n}{12}(n + m + 1), \qquad (6.6)$$

де m та n - відповідно кількість членів у вибірках, які порівнюються.

Рівень значимості приймають двосторонній і той же, що й у попередніх випадках – 5%.

Критичною областю для нульової гіпотези є область великих за абсолютною величиною відхилень з довірчими межами

$$u_H \le M_u - tp\sigma_u, \qquad (6.7)$$

$$u_B \ge M_u + tp\sigma_u, \qquad (6.8)$$

де u_H та u_B – відповідно нижнє та верхнє критичні значення, tp – нормоване відхилення за прийнятого рівня значимості (його значення для двосторонніх рівнів значимості 0.1, 1.0, 5.0, та 10.0% відповідно дорівнюють 3.29, 2.58, 1.96, і 1.64), $\sigma_u = \sqrt{D_u}$ - середнє квадратичне відхилення кількості інверсій.

Якщо порахована загальна кількість інверсій попадає у область допустимих значень, то приймається нульова гіпотеза (вибірки однорідні). Але якщо загальна кількість інверсій перевищує або менше нижнього критичного значення, то гіпотеза про однорідність вибірок з ймовірністю 95% відхиляється і приймається альтернативна (вибірки неоднорідні).

Статистика W *критерію Вількоксона***.** Ранговий критерій, як й статистика u**,** застосовується для перевірки гіпотези про належність двох вибірок до однієї сукупності.

Нехай треба визначити однорідність вибірок значень річного стоку річки за два періоди з різними умовами формування стоку (таблиця 6.1) тобто є дві вибірки, всі члени яких розташовуються в один загальний ряд у зростаючому порядку їхніх значень. Об’єм першої вибірки (X)– m =11, другої (Y)– n =15. Порядкові номера членів такого ряду називаються рангами (R). Сума рангів меншої за об'ємом вибірки, а при їхньої рівності першої з них, й буде статистикою w (в нашому випадку це сума рангів членів вибірки X).

Таблиця 6.1.Приклад оцінки однорідності ряду по $w-$ статистиці критерію Вількоксона

Рік	Річний стік Q (м^3/с) у хронологічному порядку	Вибірки				Ранги членів вибірок	
		Q_x	Q_y	у зростаючому порядку		R_x	R_y
				Q_x	Q_y		
1968	*Y1*	*X1*			*Y5*		*1*
1969	*Y2*	*X2*			*Y8*		*2*
1970	*Y3*	*X3*			*Y2*		*3*
1971	*Y4*	*X4*			*Y3*		*4*
1972	*Y5*	*X5*			*Y14*		*5*
1973	*Y6*	*X6*			*Y1*		*6*
1974	*Y7*	*X7*			*Y15*		*7*
1975	*Y8*	*X8*		*X4*		*8*	
1976	*Y9*	*X9*		*X8*		*9*	
1977	*Y10*	*X10*			*Y10*		*10*
1978	*Y11*	*X11*			*Y4*		*11*
1979	*Y12*		*Y1*		*Y12*		*12*
1980	*Y13*		*Y2*	*X3*		*13*	
1981	*Y14*		*Y3*		*Y7*		*14*
1982	*Y15*		*Y4*		*Y11*		*15*
1983	*X1*		*Y5*		*Y6*		*16*
1984	*X2*		*Y6*	*X11*		*17*	
1985	*X3*		*Y7*		*Y13*		*18*
1986	*X4*		*Y8*		*Y9*		*19*
1987	*X5*		*Y9*	*X1*		*20*	
1988	*X6*		*Y10*	*X9*		*21*	
1989	*X7*		*Y11*	*X6*		*22*	
1990	*X8*		*Y12*	*X2*		*23*	
1991	*X9*		*Y13*	*X7*		*24*	
1992	*X10*		*Y14*	*X10*		*25*	
1993	*X11*		*Y15*	*X5*		*26*	
		m=11	*n=15*			$w_x = 208$	$w_y = 143$

Підрахунок статистик w_x та w_y ведеться за формулами:

$$w_x = R_{x1} + R_{x2} + R_{x3} .... + R_{xm} \tag{6.9}$$

$$w_y = R_{y1} + R_{y2} + R_{y3} .... + R_{yn} \tag{6.10}$$

Контроль за правильністю підрахунку рангів статистик w_x та w_y виконується за формулою:

$$w_x + w_y = \frac{(m+n)\cdot(m+n+1)}{2} \tag{6.11}$$

Якщо розподіл випадкової величини w симетричний відносно математичного сподівання, то верхнє критичне значення w_B пов’язано з нижнім критичним значенням w_H співвідношенням

$$w_B = 2M_w - w_H. \tag{6.12}$$

Нижнє критичне значення w_H для $m \le 25$, $n \le 25$ для рівнів значимості (двосторонніх) 1; 5 та 10% наведені у таблиці Б.3 (додаток Б). У останньому стовбці вказано подвоєне математичне сподівання $2M_w$ як функція від m і n.

Якщо хоча б одна з об’ємів вибірок перевищує 25, то для визначення нижніх критичних значень w_H використовують формулу

$$w_H = M_w - tp \cdot \sigma_w, \tag{6.13}$$

де tp – нормоване відхилення за прийнятим рівнем значимості (1, 5 та 10% дорівнюють 2,58; 1,96 та 1,64 відповідно). Математичне сподівання M_w та середнє квадратичне відхилення σ_w статистики w для формули (6.13) визначається за формулами:

$$M_w = \frac{m\cdot(m+n+1)}{2}, \tag{6.14}$$

$$\sigma_w = \sqrt{\frac{m\cdot n(m+n+1)}{12}}. \tag{6.15}$$

ПРАКТИЧНА РОБОТА №7
Оцінювання узгодження емпіричних і аналітичних функцій розподілу за допомогою критерію χ^2

Завдання: 1) розрахувати критерій згоди емпіричної і аналітичної кривих розподілу; 2) порівняти розраховане та теоретичне значення критерію згоди при прийнятому рівні значимості; 3) прийняти нульову або альтернативну гіпотезу щодо узгодження емпіричної і аналітичної кривих розподілу.

Пояснення. Велике різноманітність аналітичних схем розподілу, які використовуються для опису гідрологічних характеристик, свідчать про те, що емпіричні дані гідрологічних спостережень у загальному непогано відповідають аналітичним кривим розподілу. Встановити відповідність між ними можна, якщо побудувати їх на відповідній клітцатці ймовірностей (практична робота №5), але це тільки візуальна оцінка. Такій прийом узгодження емпіричних даних і аналітичної функції хоча володіє низкою переваг, але має суттєвий недолік, який полягає у суб’єктивізмі і, як слід, у неоднозначності отриманого рішення. Існують критерії згоди емпіричних і аналітичних кривих розподілу. Розглянемо один з них, а саме критерій χ^2.

Порядок виконання та методичні указівки. В якості міри розходження в даному критерії згоди використовується співставлення згрупованих емпіричних ймовірностей з ймовірностями прийнятого аналітичного закону розподілу. Тому використання критерію χ^2 при малому об'ємі статистичної сукупності (n<30) недоцільно. Особливо корисно використовувати критерій χ^2, коли перевіряються декілька аналітичних функцій розподілу щодо розподілу емпіричних точок. Головна перевага критерію χ^2 полягає у тому, що він може використовуватися для оцінки узгодженості емпіричних даних з будь-яким законом розподілу.

При використанні цього критерію рекомендується приймати нерівномірні інтервали змінної випадкової величини так, щоб аналітична ймовірність попадання цієї змінної у кожний інтервал була постійною. Кількість інтервалів можна прийняти рівним 10. Тоді межі інтервалів змінної випадкової величини легко визначаються виходячи з прийнятої аналітичної кривої розподілу. Для цього розбиваємо весь інтервал ймовірностей перевищення (100%) на десять рівних частин. За прийнятою аналітичною кривою розподілу для ймовірностей перевищення 10, 20, 30,......., 90% визначають відповідно межі інтервалів змінної випадкової величини. Для першого інтервалу з ймовірністю перевищення від 0 до 10% верхня межа інтервалу змінної буде дорівнювати нескінченості (∞), тому що функція розподілу безмежна зверху. Ці дані заносять до таблиці 7.1. Якщо випадкова змінна попадає на межу інтервалу, то її відносять до верхнього.

Таблиця 7.1. Вихідні дані для розрахунку критерію згоди χ^2 емпіричної і аналітичної функцій розподілу

№ інтервалу,	Інтервали ймовірностей перевищення, %	Межі інтервалів змінної випадкової величини за прийнятою теоретичною кривою розподілу	Кількість попадань емпіричних точок змінної випадкової величини у інтервал теоретичної кривої розподілу, n_i^*	$(n_i^*)^2$
1	0 -10	∞-…		
2	10-20			
3	20-30			
4	30-40			
5	40-50			
6	50-60			
7	60-70			
8	70-80			
9	80-90			
10	90-100			

Теоретична ж кількість випадків попадання у кожний інтервал для конкретної послідовності дорівнює

$$n_{\text{òåîð.}} = \frac{N}{K}, \tag{7.1}$$

де N - кількість членів вибірки, K - кількість інтервалів, яке дорівнює в нашому випадку 10.

Значення критерію $\chi^2_{\text{ðîçð.}}$ для конкретного випадку, який аналізується щодо узгодження емпіричного і аналітичного розподілу, розраховують за формулою

$$\chi^2_{\text{ðîçð.}} = \frac{1}{n_{\text{òåîð.}}} \sum (n_i^{*})^2 - N. \tag{7.2}$$

За таблицею розподілу χ^2 (таблиця В.1, додаток В) визначають теоретичне значення параметра $\chi^2_{\text{òåîð.}}$ при рівні значимості 5% і за кількістю степенів вільності

$$l = K - r - 1, \tag{7.3}$$

де l – число степенів вільності, K – кількість інтервалів, r – кількість вільно призначених параметрів, у нашому випадку їх три $(\overline{x}, C_V, C_S)$. Тоді $l = 6$.

Потім порівнюють теоретичне та розрахункове значення критерію χ^2. Якщо розрахований критерій χ^2_{ϕ} буде за значенням меншим за теоретичний $\chi^2_{\text{òåîð.}}$ ($\chi^2_{\text{ðîçð.}} < \chi^2_{\text{òåîð.}}$), то це свідчить про те, що χ^2_{ϕ} попадає в область допустимих значень і ,як слід, при прийнятому рівні значимості нульова гіпотеза приймається. Це говорить про те, що прийнята гіпотеза відповідності емпіричних точок аналітичної кривої розподілу на цьому етапі дослідження вважається правильною.

ПРАКТИЧНА РОБОТА №8
Дослідження кореляційних зв’язків та складання рівнянь регресії

Завдання: 1) користуючись матеріалами гідрометеорологічних спостережень скомпонувати необхідні дані для дослідження кореляційних зв’язків; 2) розрахувати коефіцієнт парної кореляції та скласти рівняння регресії; 3) провести розрахунки та скласти рівняння регресії множинної кореляції, визначити ймовірну похибка коефіцієнта множинної кореляції; 2).визначити і оцінити автокореляційні зв’язки.

Пояснення. Гідрологічні явища, як правило, обумовлені досить великою кількостю факторів, повне врахування яких практично неможливо і більшості випадків недоцільно. Тому при встановленні різного роду причиннонаслідкових зв'язків в аналіз включають лише ті фактори, які на підставі загальних фізичних міркувань про закономірності та особливості утворення стоку можуть розглядатися як головні, що вносять основний вклад у процеси формування гідрологічних характеристик. Ці головні фактори й визначають основний вид зв'язку, а менш впливові та менш суттєві причини утворюють поле розсіяння.

Важливою обставиною, яка дозволяє використовувати статистичні зв'язки для розрахунку та прогнозу характеристик гідрологічного режиму є прийняття гіпотези сталості (або стаціонарності) комплексу умов, які обумовлюють ці зв'язки. При вивченні статистичних зв'язків взагалі та між гідрометеорологічними змінними зокрема звичайно застосовують так званні кореляційні зв'язки, які представляють собою зв'язки між визначеними фіксованими значеннями однієї величини (аргументу) і відповідними їм умовними середніми значеннями іншої (функції). Очевидно, що кореляційні зв'язки являють собою часткову форму вираження статистичних зв'язків. Вони полягають в становленні імовірнісних залежностей між двома або більше змінними, які у даному процесі або зв'язані як причина і наслідок, або є наслідком загальної причини. Параметри часу і простору являються тією ланкою, яка пов'язує змінні, але звичайно вони виключаються у процесі побудови зв'язку. У результаті цього одна змінна (Y) виражається в залежності від іншої ($X1$) або інших ($X1, X2, X3......$). Змінна Y умовно називається залежною змінною, в той час як інші – незалежними. Іноді змінна, яка повинна бути передбачена на основі зв'язку, називається предиктантом (звичайно це змінна Y або її функція), а вихідні незалежні змінні ($X1, X2, X3......$) – предикторами.

Порядок виконання та методичні указівки. Кореляційні зв'язки виражаються у формі кореляційних рівнянь або рівнянь регресії, які бувають лінійними і нелінійними.

Найбільше застосування у практиці гідрологічних розрахунків та прогнозів отримала парна кореляція (лінійна кореляція між двома змінними).

Парна кореляція може бути подана у вигляді графічної залежності або визначена аналітично. Виявлення зв'язку між гідрометеорологічними характеристиками здійснюється на основі співставлення двох сукупностей. Наприклад, вихідними даними для розрахунку коефіцієнта кореляції можуть бути суми добових опадів (x), які випали попередньо до проходження максимальних витрат води у річці, та максимальні витрати води під час паводку (y).

Вихідна інформація подається у вигляді таблиці 8.1.

Коефіцієнт парної кореляції розраховується за рівнянням

$$r = \frac{\sum(\Delta x \cdot \Delta y)}{\sqrt{\sum \Delta x^2 \cdot \sum \Delta y^2}}. \qquad (8.1)$$

Середні квадратичні відхилення –

$$\sigma_x = \sqrt{\frac{\sum \Delta x^2}{n}}, \tag{8.2}$$

$$\sigma_y = \sqrt{\frac{\sum \Delta y^2}{n}}. \tag{8.3}$$

Таблиця 8.1. Вихідні дані для розрахунку коефіцієнта парної кореляції між сумами добових опадів та максимальними витратами води під час паводку р. ….. п. …..

№ пп.	Рік	x_i	y_i	$\Delta x = x_i - \overline{x}$	$\Delta y = y_i - \overline{y}$	Δx^2	Δy^2	$\Delta x \cdot \Delta y$

Ймовірна похибка коефіцієнта кореляції –

$$\sigma_r = \frac{1 - r^2}{\sqrt{n-1}}. \tag{8.4}$$

Приведені формули (8.1-8.4) справедливі у тому випадку, якщо ряд спостережень більше ніж 20 значень. За малої кількості членів ряду (n<20) коефіцієнт кореляції може виявитися недостовірним (випадковим).

Оцінку достовірності (невипадковості) коефіцієнта кореляції визначають за допомогою коефіцієнта достовірності

$$K_f = \frac{|r| \sqrt{n-1}}{1 - r^2}, \tag{8.5}$$

де $|r|$ - абсолютна величина коефіцієнта кореляції, n - кількість членів ряду.

Якщо $K_f < 1$, то зв'язок відсутній; при $1 < K_f < 3$ є тенденція зв'язку; при $K_f > 3$ – зв'язок достовірний.

Середня квадратична похибка коефіцієнта кореляції обчислюється за формулою В.І. Романовського

$$\sigma_r = \frac{1 - r^2}{\sqrt{n-1}} \cdot \sqrt{1 + \frac{11r^2}{2n} + \frac{75r^2 - 13}{2n^2}}. \tag{8.6}$$

Якщо n>20 , формула (61) спрощується

$$\sigma_r \approx \frac{1 - r^2}{\sqrt{n-1}}. \tag{8.7}$$

Якщо значення коефіцієнта кореляції $r > 0.8$, то зв'язок вважається добрим; при $0.6 < r < 0.8 -$ зв'язок задовільний; при $0.4 < r < 0.6$ зв'язок слабкий, а при $r < 0.4$ можна вважати, що зв'язок відсутній.

За розрахованим значенням коефіцієнта кореляції можна скласти два рівняння регресії

$$y_i - \bar{y} = \Delta y_i = r\frac{\sigma_y}{\sigma_x}\Delta x_i, \tag{8.8}$$

$$x_i - \bar{x} = \Delta x_i = r\frac{\sigma_x}{\sigma_y}\Delta y_i. \tag{8.9}$$

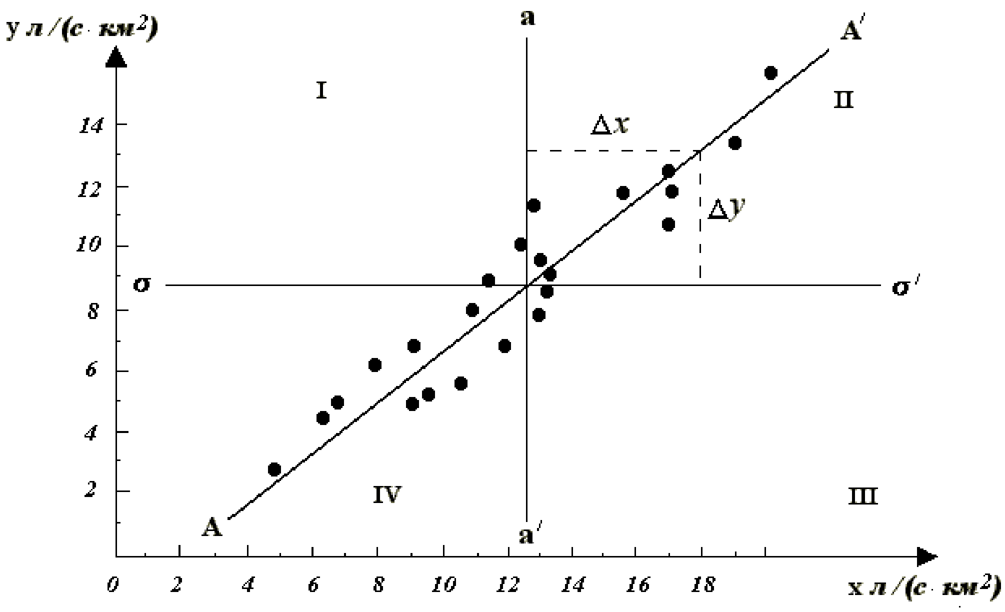


Рисунок 8.1. Приклад графік зв'язку між модулями стоку р.Стир-м.Луцьк *(у)* та р.Горинь-с.Оженін *(х)*

Коефіцієнти $r\frac{\sigma_y}{\sigma_x}$ та $r\frac{\sigma_x}{\sigma_y}$ відображають кут нахилу (через його тангенс і котангенс) прямої регресії до шкали ординат і шкали абсцис. За допомогою цих коефіцієнтів можна провести лінії зв'язку на графіку $y = f(x)$ або $x = f(y)$ (наприклад, рис. 8.1). Рівняння (8.8) і (8.9) характеризуються тим, що визначення y за відомим x можливо тільки за рівнянням (8.8), а x за відомим y - за рівнянням (8.9).

Абсолютні середні квадратичні похибки визначення y та x за рівняннями 8.8-8.9 знаходять за формулами

$$\Delta E_y = \pm\sigma_y\sqrt{1-r^2}, \tag{8.10}$$

$$\Delta E_x = \pm\sigma_x\sqrt{1-r^2}. \tag{8.11}$$

Якщо змінні x та y незалежні між собою, то сума добутку їх відхилень від середнього, що у чисельнику рівняння (8.1), дорівнює нулю, й як слід, дорівнює нулю коефіцієнт кореляції. У тому випадку, коли зв'язок між змінними функціональна та ще й лінійна, коефіцієнт кореляції дорівнює +1 або – 1. За наявності кореляційного зв'язку та в залежності від його тісноти величина коефіцієнта парної кореляції змінюється від –1 до +1.На рисунках

8.2 та 8.3 показані типові залежності змінних випадкових величин та деякі приклади кореляційних зв'язків між X та Y.

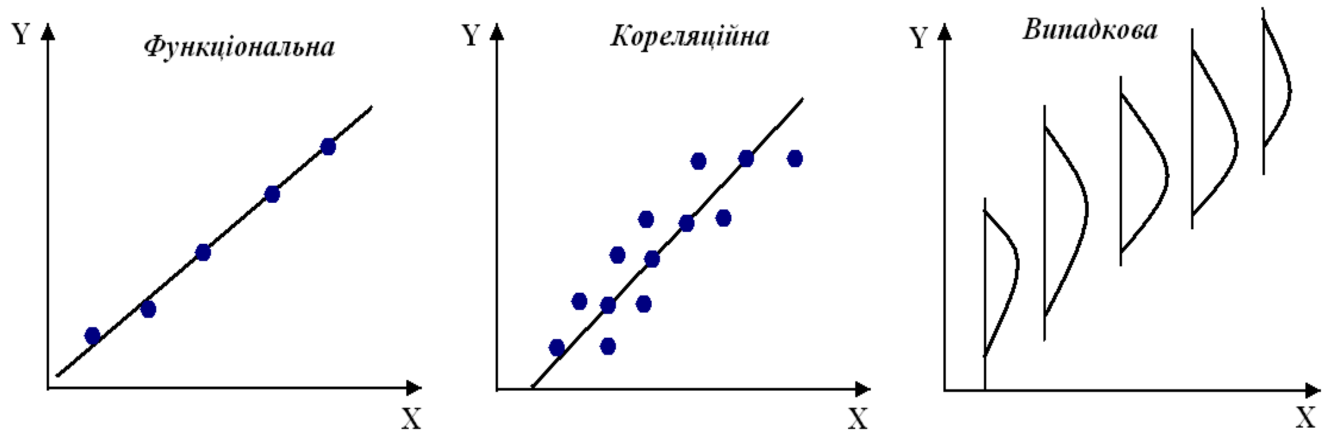


Рисунок 8.2.Типові залежності змінних випадкових величин

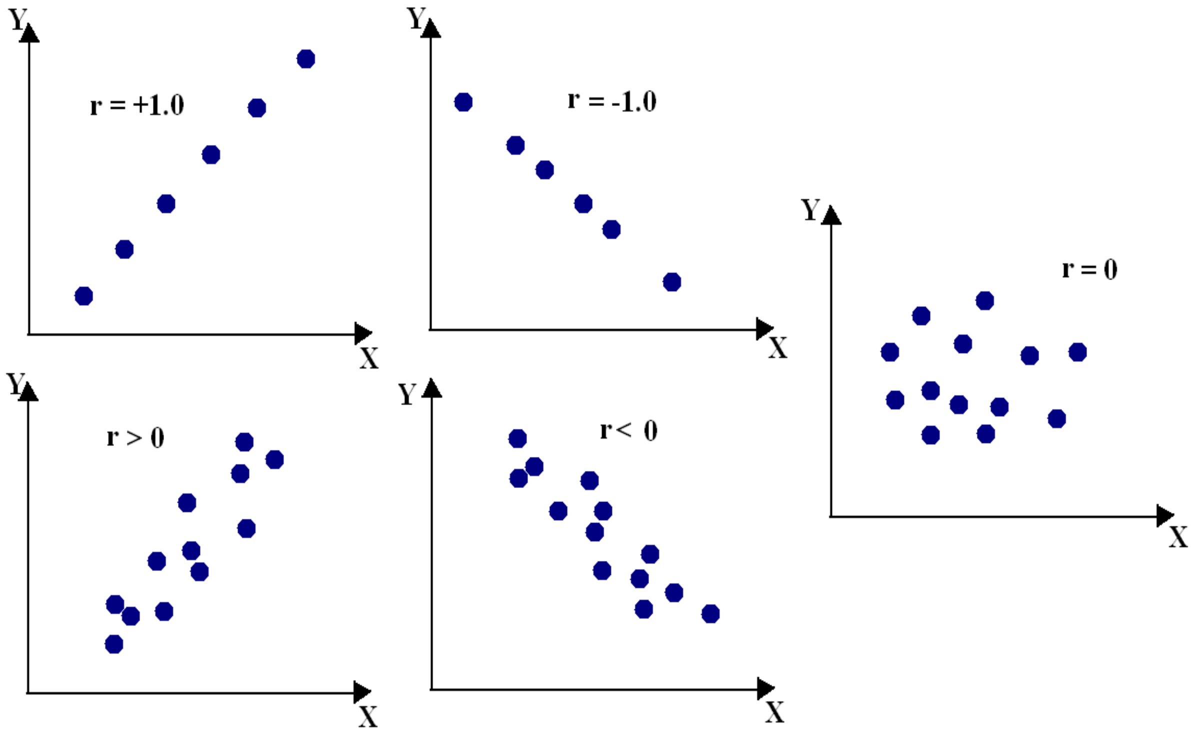


Рисунок 8.3. Деякі приклади кореляційних зв'язків між X та Y

Множинна кореляція полягає у розповсюдженні положень парної кореляції на декілька змінних. У її основі, як й у випадку з двома змінними, лежить спосіб найменших квадратів. За допомогою множинної кореляції проводиться оцінювання впливу кожного фактора на досліджувану величину. Введення великої кількості змінних в рівняння регресії призводить до великих похибок. Тому найбільш допустимим є находження зв'язку між трьома або чотирма змінними. Для трьох змінних рівняння регресії має вигляд

$$z_i - \bar{z} = \Delta z_i = \frac{\sigma_z}{\sigma_x} \cdot \frac{r_{zx} - r_{zy} \cdot r_{xy}}{1 - r^2{}_{xy}} \Delta x_i + \frac{\sigma_z}{\sigma_y} \cdot \frac{r_{zy} - r_{zx} \cdot r_{xy}}{1 - r^2{}_{xy}} \cdot \Delta y_i . \qquad (8.12)$$

У цьому у рівнянні $\sigma_z = \sqrt{\frac{\sum \Delta z_i{}^2}{n}}$; $\sigma_x = \sqrt{\frac{\sum \Delta x_i{}^2}{n}}$; $\sigma_y = \sqrt{\frac{\sum \Delta y_i{}^2}{n}}$ є середні квадратичні відхилення від відповідних середніх арифметичних значень; r_{xy} ; r_{zy} ; r_{zx} - часткові коефіцієнти кореляції між відповідними парами змінних, які розраховуються за формулою (8.1). Загальний (або зведений) коефіцієнт кореляції дорівнює

$$R_r = \sqrt{\frac{r_{zx}^2 + r_{zy}^2 - 2r_{xy} \cdot r_{zy} \cdot r_{zx}}{1 - r_{xy}^2}} . \qquad (8.13)$$

Ймовірна похибка коефіцієнта множинної кореляції $\boldsymbol{R_r}$ дорівнює

$$\varepsilon_r = \pm 0.674 \frac{1 - R_r^2}{\sqrt{n}} . \qquad (8.14)$$

Середня квадратична похибка спостережених величин від розрахованих за рівнянням регресії визначається за формулою:

$$E_z = \sigma_z \sqrt{1 - R_r^2} . \qquad (8.15)$$

У цьому завданні можна, наприклад, встановити наявність кореляційного зв'язку між шаром весняного стоку води (Z), запасами води у снігові (X) і опадами, які випали у період весняного сніготанення (Y). Приклад обчислення коефіцієнта регресії трьох змінних показано в таблиці 8.2.

Таблиця 8.2 .Вихідні дані для розрахунку множинної кореляції між шаром весняного стоку води (Z),запасами води у снігові (X)та опадами у період сніготанення (Y)

р. п. за період

№ пп.	Рік	Шар весняного стоку води, Z (мм)	Запас води в снігові, X (мм)	Опади в період сніготанення, Y (мм)	$\Delta Z = Z_i - \bar{Z}$	$\Delta X = X_i - \bar{X}$	$\Delta Y = Y_i - \bar{Y}$	ΔZ^2	ΔX^2	ΔY^2	$\Delta X \cdot \Delta Z$	$\Delta Z \cdot \Delta Y$	$\Delta X \cdot \Delta Y$

Загальний коефіцієнт множинної кореляції змінюється від 0 до 1.

Розрахунок *коефіцієнта автокореляції* проводиться за тими ж рівняннями та законами, що і коефіцієнта парної кореляції (рівняння 8.1-8.3 та

8.8-8.9). Необхідність розрахунку цього коефіцієнту пов'язана з виявленням внутрішньорядного зв'язку в ряді випадкових величин. В цьому випадку визначити зв'язок між двома послідовно розташованими членами статистичного ряду або зв'язок між членами ряду, зсунутими відносно один одного на декілька членів.

Розрахунок коефіцієнта автокореляції проводиться за формою таблиці 8.3, де за аргумент, наприклад, прийнято фактичний ряд спостережень за витратами води на спаді повені (x), а функцією буде той же ряд, але зсунутий на добу (y), тобто розраховується зв'язок наступної витрати води від попереднього її значення. Таким чином, досліджуваних пар повинно бути $(n-1)$.

Таблиця 8.3. Вихідні дані для виявлення зв'язку між витратами води на спаді повені

№ пп.	Рік	$Q_{(t)}$ (x)	$Q_{(t+1)}$ (y)	Δx	Δy	Δx^2	Δy^2	$\Delta x \cdot \Delta y$

У випадку, якщо $r \geq 0.2$, то можна вважати, що внутрішньорядний зв'язок існує. Тобто, це означає, що кожний наступний член можна оцінити за значенням попереднього. Аналогічно можна встановити наявність зв'язку, наприклад, між $x_{(t)}$ та $x_{(t+2)}$, або $x_{(t)}$ та $x_{(t+5)}$ тощо.

Рекомендована література

Основна:

1. Будкина Л.Г., Ободовский А.Г. Методические указания к выполнению лабораторных работ по спец.курсу «Статистические методы в гидрологии и гидрохимии» - К.: КГУ. -1985
2. Гидрологические и водобалансовые расчеты/ под ред. Н.Г.Галущенко – К.:Вища школа. 1987.
3. Митропольский А.К. Техника статистических вычислений. – М.: Изд-во «Наука», 1971
4. Пособие по определению расчетних гидрологичних характеристик. – Л.:1984
5. Рождественский А.В., Чеботарев А.И. Статистические методы в гидрологии - Л.: Гидрометеоиздат, 1974
6. Самохин А.А., Соловьева Н.Н., Догановский А.М. Практикум по гидрологии. - Л.: Гидрометеоиздат, 1980
7. Статистические методы в гидрологии/под ред. Г.А.Алексеева. - Л.: Гидрометеоиздат, 1970

Додаткова:

8. Пелешенко В.И., Ромась Н.И. Применение вероятностно-статистических методов для анализа гидрохимических данных- К.: КГУ. -1977

9. Рождественский А.В. Оценка точности кривых распределения гидрологических характеристик. - Л. : Гидрометеиздат, 1977.
10. Христофоров А.В. Теория вероятностей и математическая статистика: Учеб.пособие. – М.:Изд-во Моск.ун-та, 1988
11. Христофоров А.В. Теория случайных процессов в гидрологии: Учеб.пособие. – М.:Изд-во Моск.ун-та, 1994

ДОДАТОК А

Таблиця А.1 – Нормовані відхилення ординат розподілу Пірсона III типу від середини (від1.0)

$$\phi(x) = f(C_S, P) = \frac{K_p - 1}{C_V}, \text{ при } C_V = 1.0$$

C_S	Відсотки ймовірності перевищення (забезпеченості)													
	0,01	*0,1*	*1*	*3*	*5*	*10*	*25*	*50*	*75*	*90*	*95*	*97*	*99*	*99,9*
-4,0	0,500	0,500	0,500	0,500	0,500	0,500	0,49	0,41	-0,010	-0,96	-1,90	-2,65	-4,34	-8,17
-3,8	0,527	0,527	0,526	0,526	0,526	0,526	0,52	0,42	-0,032	-1,00	-1,90	-2,65	-4,29	-7,97
-3,6	0,556	0,556	0,556	0,556	0,556	0,555	0,54	0,42	-0,064	-1,03	-1,93	-2,66	-4,24	-7,72
-3,4	0,588	0,588	0,588	0,588	0,587	0,586	0,57	0,41	-0,11	-1,06	-1,94	-2,66	-4,18	-7,54
-3,2	0,625	0,625	0,625	0,625	0,625	0,621	0,59	0,41	-0,15	-1,09	-1,96	-2,66	-4,11	-7,35
-3,0	0,667	0,667	0,666	0,666	0,665	0,661	0,62	0,40	-0,19	-1,13	-1,97	-2,66	-4,05	-7,10
-2,8	0,715	0,715	0,715	0,714	0,711	0,703	0,64	0,39	-0,22	-1,18	-2,00	-2,65	-3,86	-6,86
-2,6	0,770	0,770	0,770	0,766	0,764	0,746	0,66	0,37	-0,25	-1,21	-2,00	-2,63	-3,86	-6,54
-2,4	0,835	0,833	0,830	0,826	0,820	0,792	0,67	0,35	-0,29	-1,25	-2,00	-2,60	-3,78	-6,37
-2,2	0,914	0,910	0,905	0,895	0,882	0,842	0,69	0,33	-0,35	-1,27	-2,02	-2,54	-3,68	-6,14
-2,0	1,01	1,00	0,990	0,970	0,950	0,900	0,71	0,31	-0,39	-1,30	-2,00	-2,51	-3,60	-5,91
-1,8	1,11	1,11	1,09	1,06	1,02	0,940	0,72	0,28	-0,42	-1,32	-1,99	-2,46	-3,50	-5,64
-1,6	1,26	1,24	1,20	1,14	1,10	0,990	0,73	0,25	-0,46	-1,33	-1,97	-2,42	-3,39	-5,37
-1,4	1,41	1,39	1,32	1,23	1,17	1,04	0,73	0,22	-0,49	-1,34	-1,95	-2,37	-3,27	-5,09
-1,2	1,68	1,58	1,45	1,33	1,24	1,08	0,74	0,19	-0,52	-1,34	-1,92	-2,31	-3,15	-4,81
-1,0	1,92	1,79	1,59	1,42	1,32	1,13	0,73	0,16	-0,55	-1,34	-1,88	-2,25	-3,02	-4,53
-0,8	2,23	2,02	1,74	1,52	1,38	1,17	0,73	0,13	-0,58	-1,34	-1,84	-2,18	-2,89	-4,24
-0,6	2,57	2,27	1,88	1,61	1,45	1,20	0,72	0,10	-0,61	-1,33	-1,80	-2,12	-2,75	-3,96
-0,4	2,98	2,54	2,03	1,70	1,52	1,23	0,71	0,07	-0,63	-1,32	-1,75	-2,04	-2,61	-3,66
-0,2	3,37	2,81	2,18	1,79	1,58	1,26	0,69	0,03	-0,65	-1,30	-1,70	-1,96	-2,47	-3,38
0,0	3,72	3,09	2,33	1,88	1,64	1,28	0,67	0,00	-0,67	-1,28	-1,64	-1,88	-2,33	-3,09
0,2	4,16	3,38	2,47	1,96	1,70	1,30	0,65	-0,03	-0,69	-1,26	-1,58	-1,79	-2,11	-2,81
0,4	4,61	3,66	2,61	2,04	1,75	1,32	0,63	-0,07	-0,71	-1,23	-1,52	-1,70	-2,03	-2,54
0,6	5,05	3,96	2,75	2,12	1,80	1,33	0,61	-0,10	-0,72	-1,20	-1,45	-1,61	-1,88	-2,27
0,8	5,50	4,24	2,89	2,18	1,84	1,34	0,58	-0,13	-0,73	-1,17	-1,38	-1,52	-1,74	-2,02

Продовження табл.А.1

C_S	Відсотки ймовірності перевищення (забезпеченості)													
	0,01	*0,1*	*1*	*3*	*5*	*10*	*25*	*50*	*75*	*90*	*95*	*97*	*99*	*99,9*
1,0	5,96	4,53	3,02	2,25	1,88	1,34	0,55	-0,16	-0,73	-1,13	-1,32	-1,42	-1,59	-1,79
1,2	6,41	4,81	3,15	2,31	1,92	1,34	0,52	-0,19	-0,74	-1,08	-1,24	-1,33	-1,45	-1,58
1,4	6,87	5,09	3,27	2,37	1,95	1,34	0,49	-0,22	-0,73	-1,04	-1,17	-1,23	-1,32	-1,39
1,6	7,31	5,37	3,39	2,42	1,97	1,33	0,46	-0,25	-0,73	-0,99	-1,10	-1,14	-1,20	-1,24
1,8	7,76	5,64	3,50	2,46	1,99	1,32	0,42	-0,28	-0,72	-0,94	-1,02	-1,06	-1,09	-1,11
2,0	8,21	5,91	3,60	2,51	2,00	1,30	0,39	-0,31	-0,71	-0,90	-0,95	-0,97	-0,99	-1,00
2,2	8,63	6,14	3,68	2,54	2,02	1,27	0,35	-0,33	-0,69	-0,842	-0,882	-0,895	-0,905	-0,910
2,4	9,00	6,37	3,78	2,60	2,00	1,25	0,29	-0,35	-0,67	-0,792	-0,820	-0,826	-0,830	-0,833
2,6	9,39	6,54	3,86	2,63	2,00	1,21	0,25	-0,37	-0,66	-0,746	-0,764	-0,766	-0,770	-0,770
2,8	9,77	6,86	3,96	2,65	2,00	1,18	0,22	-0,39	-0,64	-0,703	-0,711	-0,714	-0,715	-0,715
3,0	10,16	7,10	4,05	2,66	1,97	1,13	0,19	-0,40	-0,62	-0,661	-0,665	-0,666	-0,666	-0,667
3,2	10,55	7,35	4,11	2,66	1,96	1,09	0,15	-0,41	-0,59	-0,621	-0,625	-0,625	-0,625	-0,625
3,4	10,90	7,54	4,18	2,66	1,94	1,06	0,11	-0,41	-0,57	-0,585	-0,587	-0,588	-0,588	-0,588
3,6	11,30	7,72	4,24	2,66	1,93	1,03	0,064	-0,42	-0,54	-0,555	-0,556	-0,556	-0,556	-0,556
3,8	11,67	7,97	4,29	2,65	1,90	1,00	0,032	-0,42	-0,52	-0,526	-0,526	-0,526	-0,526	-0,527
4,0	12,02	8,17	4,34	2,65	1,90	0,96	0,010	-0,41	-0,49	-0,500	-0,500	-0,500	-0,500	-0,500
4,2	12,40	8,38	4,39	2,64	1,88	0,93	-0,010	-0,41	-0,473	-0,476	-0,476	-0,476	-0,477	-0,477
4,4	12,76	8,60	4,42	2,63	1,86	0,91	-0,032	-0,40	-0,454	-0,455	-0,455	-0,455	-0,455	-0,455
4,6	13,12	8,79	4,46	2,62	1,84	0,87	-0,052	-0,40	-0,454	-0,435	-0,435	-0,435	-0,435	-0,435
4,8	13,51	8,96	4,50	2,60	1,81	0,82	-0,075	-0,39	-0,416	-0,416	-0,416	-0,416	-0,417	-0,417
5,0	13,87	9,12	4,54	2,60	1,78	0,78	-0,099	-0,38	-0,400	-0,400	-0,400	-0,400	-0,400	-0,400
5,2	14,25	9,27	4,59	2,60	1,74	0,73	-0,120	-0,37	-0,385	-0,385	-0,385	-0,385	-0,385	-0,385
5,4	14,60	9,42	4,62	2,60	1,70	0,67	-0,100	-0,37	-0,37	-0,37	-0,37	-0,37	-0,37	-0,37
5,6	14,95	9,59	4,65	2,60	1,67	0,62	-0,120	-0,36	-0,36	-0,36	-0,36	-0,36	-0,36	-0,36
5,8	15,32	9,70	4,70	2,60	1,64	0,57	-0,140	-0,35	-0,35	-0,35	-0,35	-0,35	-0,35	-0,35
6,0	15,67	9,84	4,70	2,60	1,60	0,51	-0,150	-0,34	-0,34	-0,34	-0,34	-0,34	-0,34	-0,34
6,2	16,04	9,35	4,71	2,60	1,56	0,47	-0,150	-0,34	-0,34	-0,34	-0,34	-0,34	-0,34	-0,34
6,4	16,40	10,05	4,71	2,60	1,52	0,42	-0,150	-0,33	-0,33	-0,33	-0,33	-0,33	-0,33	-0,33

Таблиця А.2 – Ординати кривих трипараметричного гамма-розподілу

(P% - ймовірність перевищення (забезпеченість))

P%	C_V							*P%*	C_V						
	0,1	0,2	0,3	0,4	0,5	0,6	0,7		0,1	0,2	0,3	0,4	0,5	0,6	0,7
$C_S = 0.5C_V$									$C_S = 0.5C_V$						
0,001	1,42	1,87	2,29	2,66	2,94	3,08	3,00	***30***	1,05	1,10	1,16	1,21	1,27	1,34	1,42
0,01	1,38	1,76	2,13	2,47	2,74	2,91	2,89	***40***	1,02	1,05	1,07	1,10	1,12	1,15	1,16
0,03	1,35	1,71	2,05	2,37	2,64	2,81	2,83	***50***	0,999	0,997	0,993	0,988	0,980	0,962	0,920
0,05	1,34	1,68	2,01	2,32	2,58	2,76	2,79	***60***	0,974	0,946	0,915	0,881	0,839	0,780	0,690
0,1	1,31	1,63	1,95	2,25	2,5	2,69	2,74	***70***	0,947	0,882	0,834	0,769	0,693	0,596	0,476
0,3	1,28	1,56	1,85	2,12	2,36	2,55	2,64	***75***	0,932	0,862	0,789	0,709	0,615	0,503	0,376
0,5	1,26	1,53	1,79	2,05	2,28	2,48	2,59	***80***	0,915	0,829	0,740	0,643	0,533	0,409	0,282
1,0	1,24	1,48	1,72	1,95	2,17	2,37	2,50	***90***	0,872	0,744	0,615	0,480	0,343	0,215	0,115
3	1,19	1,38	1,58	1,78	1,97	2,16	2,33	***95***	0,837	0,676	0,517	0,362	0,221	0,113	0,047
5	1,17	1,33	1,51	1,68	1,86	2,03	2,22	***97***	0,814	0,633	0,458	0,296	0,160	0,070	0,024
10	1,13	1,26	1,39	1,53	1,67	1,83	2,01	***99***	0,772	0,554	0,354	0,189	0,080	0,025	0006
0	1,08	1,17	1,25	1,35	1,44	1,56	1,70	***99,5***	0,748	0,511	0,302	0,144	0,051	0,013	0,002
25	1,07	1,13	1,20	1,27	1,35	1,45	1,56	***99,7***	0,732	0,482	0,269	0,117	0,037	0,008	0,001
								99,9	0,700	0,428	0,210	0,076	0,019	0,003	0

Продовження табл.А.2

P%	C_V						
	0,1	**0,2**	**0,4**	**0,6**	**0,8**	**1,0**	**1,2**
$C_S = C_V$							
0,01	**1,38**	**1,81**	**2,70**	**3,57**	**4,31**	**4,92**	**5,34**
0,1	**1,32**	**1,67**	**2,40**	**3,13**	**3,82**	**4,42**	**4,92**
1,0	**1,24**	**1,49**	**2,03**	**2,59**	**3,16**	**3,75**	**4,36**
3,0	**1,19**	**1,39**	**1,82**	**2,27**	**2,75**	**3,29**	**3,92**
5,0	**1,17**	**1,34**	**1,70**	**2,10**	**2,53**	**3,02**	**3,63**
10,0	**1,13**	**1,26**	**1,54**	**1,83**	**2,16**	**2,55**	**3,03**
25,0	**1,07**	**1,13**	**1,26**	**1,39**	**1,52**	**1,64**	**1,69**
50	**0,998**	**0,993**	**0,972**	**0,928**	**0,836**	**0,666**	**0,446**
75	**0,932**	**0,861**	**0,708**	**0,528**	**0,321**	**0,144**	**0,046**
90	**0,873**	**0,748**	**0,500**	**0,264**	**0,092**	**0,019**	**0,002**
95	**0,838**	**0,683**	**0,392**	**0,157**	**0,036**	**0,004**	**0**
97	**0,816**	**0,642**	**0,329**	**0,107**	**0,018**	**0,001**	**0**
99	**0,775**	**0,568**	**0,229**	**0,047**	**0,004**	**0**	**0**
99.9	**0,707**	**0,451**	**0,108**	**0,008**	**0**	**0**	**0**
$C_S = 1{,}5C_V$							
0,01	**1,40**	**1,86**	**2,94**	**4,19**	**5,61**	**7,19**	**8,92**
0,1	**1,33**	**1,70**	**2,54**	**3,52**	**4,62**	**5,84**	**7,18**
1,0	**1,24**	**1,51**	**2,09**	**2,76**	**3,49**	**4,30**	**5,21**
3	**1,19**	**1,40**	**1,85**	**2,34**	**2,88**	**3,46**	**4,12**
5	**1,17**	**1,35**	**1,72**	**2,13**	**2,57**	**3,03**	**3,55**
10	**1,13**	**1,26**	**1,54**	**1,82**	**2,11**	**2,41**	**2,71**
25	**1,07**	**1,13**	**1,25**	**1,35**	**1,43**	**1,48**	**1,49**
50	**0,998**	**0,990**	**0,958**	**0,902**	**0,814**	**0,690**	**0,541**
75	**0,931**	**0,860**	**0,708**	**0,545**	**0,377**	**0,223**	**0,111**
90	**0,874**	**0,751**	**0,518**	**0,310**	**0,148**	**0,053**	**0,014**
95	**0,840**	**0,689**	**0,419**	**0,207**	**0,074**	**0,018**	**0,003**
97	**0,819**	**0,651**	**0,363**	**0,155**	**0,045**	**0,008**	**0,001**
99	**0,780**	**0,581**	**0,268**	**0,084**	**0,015**	**0,001**	**0**
99.9	**0,720**	**0,472**	**0,150**	**0,021**	**0**	**0**	**0**

Продовження табл.А.2

P%	C_V										
	0,1	**0,2**	**0,4**	**0,6**	**0,8**	**1,0**	**1,2**	**1,4**	**1,6**	**1,8**	**2,0**
$C_S = 2C_V$											
0,01	**1,42**	**1,92**	**3,20**	**4,85**	**6,85**	**9,21**	**11,8**	**14,7**	**18,2**	**22,2**	**26,6**
0,10	**1,34**	**1,73**	**2,70**	**3,87**	**5,30**	**6,91**	**8,65**	**10,6**	**12,5**	**14,6**	**17,0**
1,0	**1,25**	**1,52**	**2,16**	**2,89**	**3,71**	**4,60**	**5,53**	**6,55**	**7,5**	**8,6**	**9,8**
3,0	**1,20**	**1,41**	**1,87**	**2,39**	**2,94**	**3,51**	**4,12**	**4,71**	**5,2**	**5,8**	**6,5**
5,0	**1,17**	**1,35**	**1,74**	**2,15**	**2,57**	**3,00**	**3,40**	**3,80**	**4,0**	**4,5**	**5,0**
10	**1,13**	**1,26**	**1,54**	**1,80**	**2,06**	**2,30**	**2,50**	**2,64**	**2,7**	**2,6**	**2,6**
25	**1,06**	**1,13**	**1,23**	**1,31**	**1,37**	**1,39**	**1,35**	**1,31**	**1,26**	**1,22**	**1,18**
50	**0,997**	**0,986**	**0,948**	**0,886**	**0,800**	**0,693**	**0,580**	**0,460**	**0,355**	**0,265**	**0,20**
75	**0,931**	**0,858**	**0,708**	**0,556**	**0,416**	**0,288**	**0,193**	**0,106**	**0,060**	**0,040**	**0,025**
90	**0,873**	**0,754**	**0,532**	**0,352**	**0,208**	**0,105**	**0,049**	**0,016**	**0,005**	**0,003**	**0,001**
95	**0,842**	**0,696**	**0,448**	**0,256**	**0,120**	**0,051**	**0,016**	**0,004**	**0,001**	**0**	**0**
97	**0,821**	**0,660**	**0,392**	**0,202**	**0,088**	**0,030**	**0,008**	**0,002**	**0**	**0**	**0**
99	**0,782**	**0,594**	**0,304**	**0,130**	**0,040**	**0,010**	**0,002**	**0**	**0**	**0**	**0**
99,9	**0,719**	**0,492**	**0,192**	**0,052**	**0,008**	**0,001**	**0**	**0**	**0**	**0**	**0**
$C_S = 2.5C_V$											
0,01	**1,44**	**1,98**	**3,49**	**5,54**	**8,10**	**11,1**	**14,6**	**18,4**	**22,5**	**27,0**	**31,9**
0,1	**1,35**	**1,77**	**2,85**	**4,24**	**5,90**	**7,76**	**9,81**	**12,0**	**14,4**	**17,0**	**19,8**
1	**1,25**	**1,54**	**2,21**	**3,00**	**3,87**	**4,78**	**5,73**	**6,71**	**7,70**	**8,71**	**9,74**
3	**1,20**	**1,42**	**1,90**	**2,42**	**2,96**	**3,50**	**4,04**	**4,56**	**5,06**	**5,54**	**6,01**
5	**1,17**	**1,35**	**1,74**	**2,15**	**2,55**	**2,94**	**3,31**	**3,65**	**3,96**	**4,26**	**4,52**
10	**1,13**	**1,26**	**1,53**	**1,78**	**2,01**	**2,22**	**2,39**	**2,53**	**2,64**	**2,73**	**2,79**
25	**1,07**	**1,12**	**1,22**	**1,28**	**1,32**	**1,33**	**1,31**	**1,27**	**1,21**	**1,14**	**1,05**
50	**0,997**	**0,984**	**0,938**	**0,870**	**0,787**	**0,695**	**0,600**	**0,505**	**0,415**	**0,332**	**0,259**
75	**0,931**	**0,858**	**0,712**	**0,571**	**0,443**	**0,332**	**0,238**	**0,164**	**0,107**	**0,066**	**0,039**
90	**0,875**	**0,757**	**0,549**	**0,381**	**0,250**	**0,155**	**0,089**	**0,047**	**0,023**	**0,010**	**0,004**
95	**0,843**	**0,702**	**0,467**	**0,293**	**0,172**	**0,093**	**0,046**	**0,020**	**0,008**	**0,002**	**0,001**
97	**0,823**	**0,667**	**0,420**	**0,247**	**0,134**	**0,065**	**0,028**	**0,011**	**0,003**	**0,001**	**0**
99	**0,784**	**0,606**	**0,341**	**0,175**	**0,080**	**0,032**	**0,011**	**0,003**	**0,001**	**0**	**0**

Продовження табл.А.2

P%	C_V										
	0,1	**0,2**	**0,4**	**0,6**	**0,8**	**1,0**	**1,2**	**1,4**	**1,6**	**1,8**	**2,0**
$C_S = 3C_V$											
0,01	**1,46**	**2,05**	**3,80**	**6,26**	**9,30**	**12,8**	**16,8**	**21,2**	**25,9**	**31,0**	**36,5**
0,1	**1,36**	**1,81**	**3,01**	**4,56**	**6,38**	**8,41**	**10,6**	**13,0**	**15,4**	**18,0**	**20,8**
1	**1,25**	**1,55**	**2,26**	**3,07**	**3,96**	**4,87**	**5,79**	**6,74**	**7,68**	**8,61**	**9,53**
3	**1,20**	**1,42**	**1,91**	**2,43**	**2,95**	**3,47**	**3,98**	**4,44**	**4,89**	**5,31**	**5,70**
5	**1,17**	**1,36**	**1,75**	**2,14**	**2,52**	**2,88**	**3,22**	**3,52**	**3,80**	**4,04**	**4,26**
10	**1,13**	**1,26**	**1,52**	**1,76**	**1,97**	**2,15**	**2,30**	**2,42**	**2,51**	**2,58**	**2,62**
25	**1,07**	**1,12**	**1,21**	**1,26**	**1,28**	**1,29**	**1,27**	**1,23**	**1,18**	**1,12**	**1,05**
50	**0,997**	**0,981**	**0,930**	**0,862**	**0,783**	**0,699**	**0,614**	**0,531**	**0,452**	**0,379**	**0,313**
75	**0,931**	**0,858**	**0,715**	**0,583**	**0,465**	**0,363**	**0,277**	**0,206**	**0,149**	**0,105**	**0,071**
90	**0,876**	**0,761**	**0,563**	**0,406**	**0,284**	**0,193**	**0,126**	**0,078**	**0,047**	**0,026**	**0,014**
95	**0,844**	**0,708**	**0,487**	**0,326**	**0,210**	**0,129**	**0,076**	**0,042**	**0,022**	**0,010**	**0,004**
97	**0,825**	**0,675**	**0,443**	**0,282**	**0,171**	**0,099**	**0,054**	**0,027**	**0,012**	**0,005**	**0,002**
99	**0,786**	**0,618**	**0,369**	**0,213**	**0,116**	**0,058**	**0,027**	**0,011**	**0,004**	**0,001**	**0**
99,9	**0,732**	**0,531**	**0,273**	**0,131**	**0,057**	**0,022**	**0,007**	**0,002**	**0**	**0**	**0**
	$C_S = 3.5C_V$										
0,01	**1,48**	**2,12**	**4,12**	**6,94**	**10,4**	**14,4**	**18,8**	**23,5**	**28,6**	**34,0**	**39,9**
0,1	**1,37**	**1,84**	**3,14**	**4,79**	**6,77**	**8,90**	**11,2**	**13,6**	**16,1**	**18,8**	**21,4**
1	**1,25**	**1,57**	**2,31**	**3,13**	**4,00**	**4,90**	**5,80**	**6,71**	**7,61**	**8,49**	**9,36**
3	**1,20**	**1,43**	**1,93**	**2,43**	**2,94**	**3,43**	**3,90**	**4,34**	**4,75**	**5,14**	**5,50**
5	**1,17**	**1,36**	**1,75**	**2,13**	**2,49**	**2,83**	**3,14**	**3,42**	**3,67**	**3,89**	**4,08**
10	**1,13**	**1,26**	**1,52**	**1,74**	**1,93**	**2,10**	**2,23**	**2,34**	**2,42**	**2,48**	**2,52**
25	**1,07**	**1,12**	**1,20**	**1,24**	**1,26**	**1,26**	**1,24**	**1,21**	**1,16**	**1,11**	**1,05**
50	**0,997**	**0,978**	**0,925**	**0,856**	**0,781**	**0,703**	**0,625**	**0,549**	**0,477**	**0,410**	**0,350**
75	**0,931**	**0,858**	**0,719**	**0,593**	**0,482**	**0,386**	**0,304**	**0,236**	**0,180**	**0,135**	**0,099**
90	**0,877**	**0,764**	**0,576**	**0,427**	**0,311**	**0,221**	**0,154**	**0,104**	**0,069**	**0,044**	**0,027**
95	**0,840**	**0,713**	**0,504**	**0,351**	**0,239**	**0,158**	**0,101**	**0,062**	**0,037**	**0,021**	**0,011**
97	**0,827**	**0,683**	**0,463**	**0,309**	**0,201**	**0,126**	**0,076**	**0,044**	**0,024**	**0,013**	**0,006**
99	**0,788**	**0,629**	**0,396**	**0,244**	**0,145**	**0,082**	**0,044**	**0,022**	**0,011**	**0,005**	**0,002**

Продовження табл.А.2

P%	C_V										
	0,1	**0,2**	**0,4**	**0,6**	**0,8**	**1,0**	**1,2**	**1,4**	**1,6**	**1,8**	**2,0**
$C_S = 4C_V$											
0,01	**1,50**	**2,18**	**4,43**	**7,58**	**11,4**	**15,5**	**20,3**	**25,4**	**30,8**	**36,5**	**42,4**
0,1	**1,38**	**1,88**	**3,29**	**5,07**	**7,08**	**9,26**	**11,6**	**14,0**	**16,6**	**19,2**	**21,9**
1	**1,25**	**1,58**	**2,31**	**3,17**	**4,03**	**4,91**	**5,79**	**6,66**	**7,52**	**8,37**	**9,19**
3	**1,20**	**1,44**	**1,93**	**2,43**	**2,92**	**3,39**	**3,83**	**4,25**	**4,64**	**5,01**	**5,34**
5	**1,17**	**1,36**	**1,75**	**2,12**	**2,46**	**2,78**	**3,07**	**3,34**	**3,57**	**3,78**	**3,96**
10	**1,13**	**1,26**	**1,51**	**1,72**	**1,90**	**2,05**	**2,18**	**2,28**	**2,36**	**2,42**	**2,45**
25	**1,07**	**1,12**	**1,19**	**1,23**	**1,24**	**1,24**	**1,22**	**1,19**	**1,15**	**1,10**	**1,05**
50	**0,997**	**0,976**	**0,920**	**0,853**	**0,781**	**0,707**	**0,634**	**0,562**	**0,495**	**0,433**	**0,375**
75	**0,931**	**0,858**	**0,722**	**0,601**	**0,495**	**0,403**	**0,325**	**0,258**	**0,203**	**0,158**	**0,121**
90	**0,877**	**0,767**	**0,586**	**0,444**	**0,331**	**0,243**	**0,176**	**0,125**	**0,087**	**0,060**	**0,040**
95	**0,846**	**0,719**	**0,519**	**0,372**	**0,261**	**0,180**	**0,121**	**0,080**	**0,051**	**0,032**	**0,019**
97	**0,829**	**0,690**	**0,481**	**0,332**	**0,224**	**0,147**	**0,095**	**0,058**	**0,056**	**0,021**	**0,012**
99	**0,790**	**0,638**	**0,417**	**0,269**	**0,168**	**0,102**	**0,060**	**0,034**	**0,018**	**0,009**	**0,004**
99,9	**0,742**	**0,561**	**0,331**	**0,189**	**0,104**	**0,054**	**0,026**	**0,012**	**0,005**	**0,002**	**0,001**

Таблиця А.3 – - Різниці нормованих відхилень ординат кривої від середини (5% та 95% ймовірності перевищення) та коефіцієнти скісності

C_S	$\phi_{5\%}$-$\phi_{95\%}$	$S=\frac{X_{5\%}+X_{95\%}-2X_{50\%}}{X_{5\%}-X_{95\%}}$	C_S	$\phi_{5\%}$-$\phi_{95\%}$	$S=\frac{X_{5\%}+X_{95\%}-2X_{50\%}}{X_{5\%}-X_{95\%}}$
-4,0	2,40	-0,93	***1,4***	3,12	0,39
-3,8	2,43	-0,91	***1,6***	3,07	0,45
-3,6	2,49	-0,89	***1,8***	3,01	0,51
-3,4	2,53	-0,86	***2,0***	2,95	0,57
-3,2	2,58	-0,83	***2,2***	2,89	0,62
-3,0	2,64	-0,80	***2,4***	2,82	0,67
-2,8	2,71	-0,76	***2,6***	2,76	0,72
-2,6	2,76	-0,71	***2,8***	2,71	0,76
-2,4	2,82	-0,67	***3,0***	2,64	0,80
-2,2	2,90	-0,62	***3,2***	2,59	0,83
-2,0	2,95	-0,57	***3,4***	2,53	0,86
-1,8	3,01	-0,51	***3,6***	2,48	0,89
-1,6	3,07	-0,45	***3,8***	2,43	0,91
-1,4	3,12	-0,39	***4,0***	2,00	0,92
-1,2	3,16	-0,34	***4,2***	2,36	0,94
-1,0	3,20	-0,27	***4,4***	2,32	0,95
-0,8	3,22	-0,22	***4,6***	2,28	0,97
-0,6	3,25	-0,17	***4,8***	2,23	0,98
-0,4	3,27	-0,11	***5,0***	2,18	0,98
-0,2	3,28	-0,08	***5,2***	2,12	0,98
0,0	3,28	0,00	***5,4***	2,07	1,00
0,2	3,28	0,06	***5,6***	2,03	1,00
0,4	3,27	0,11	***5,8***	1,99	1,00
0,6	3,25	0,17	***6,0***	1,94	1,00
0,8	3,22	0,22	***6,2***	1,90	1,00
1,0	3,20	0,28	***6,4***	1,85	1,00
1,2	3,16	0,34			

Примітка. ($\phi_{5\%}$-$\phi_{95\%}$) - різниця нормованих відхилень ординат кривої ймовірностей перевищення від середини (5% та 95% ймовірності перевищення); $S=\frac{X_{5\%}+X_{95\%}-2X_{50\%}}{X_{5\%}-X_{95\%}}$ - коефіцієнт якісності.

ДОДАТОК Б

Таблиця Б.1 – Рівні значимості та відповідні їм відхилення Z при односторонньому (*Q*) та двосторонньому (*2Q*) критеріях у випадку нормального розподілу

Q	2Q	Z	Q	2Q	Z
0,4	0,8	0,25	0,01	0,02	2,33
0,25	0,5	0,67	0,005	0,01	2,58
0,1	0,2	1,28	0,0025	0,005	2,81
0,05	0,1	1,64	0,001	0,002	3,09
0,025	0,05	1,96	0,0005	0,001	3,29

Таблиця Б.2 – Критичні значення ($t_{f,K}$) t –розподілу Стьюдента при різних рівнях значимості f і даної кількості степенів вільності K

K	Рівні значимості для одностороннього обмеження критерію (f)							
	0,1	0,05	0,025	0,01	0,005	0,0025	0,001	0,0005
	Рівні значимості для двостороннього обмеження критерію ($2f$)							
	0,2	0,1	0,05	0,02	0,01	0,005	0,002	0,001
1	3,08	6,31	12,71	31,82	63,66	127,32	318,30	636,61
2	1,89	2,92	4,30	6,96	9,92	14,09	22,33	31,60
3	1,64	2,35	3,18	4,54	5,84	7,45	10,21	12,92
4	1,53	2,13	2,78	3,75	4,60	5,60	7,17	8,61
5	1,48	2,02	2,57	3,36	4,03	4,77	5,89	6,87
6	1,44	1,94	2,45	3,14	3,71	4,32	5,21	5,96
7	1,41	1,89	2,36	3,00	3,50	4,03	4,79	5,41
8	1,40	1,86	2,31	2,90	3,36	3,83	4,50	5,04
9	1,38	1,83	2,26	2,82	3,25	3,69	4,30	4,78
10	1,37	1,81	2,23	2,76	3,17	3,58	4,14	4,59
11	1,36	1,80	2,20	2,72	3,11	3,50	4,02	4,44
12	1,36	1,78	2,18	2,68	3,05	3,43	3,93	4,32
13	1,35	1,77	2,16	2,65	3,01	3,37	3,85	4,22
14	1,34	1,76	2,14	2,62	2,98	3,33	3,79	4,14
15	1,34	1,75	2,13	2,60	2,95	3,29	3,73	4,07
16	1,34	1,75	2,12	2,58	2,92	3,25	3,69	4,02
17	1,33	1,74	2,11	2,57	2,90	3,22	3,65	3,97
18	1,33	1,73	2,10	2,55	2,88	3,20	3,61	3,92
19	1,33	1,73	2,09	2,54	2,86	3,17	3,58	3,88
20	1,33	1,72	2,09	2,53	2,85	3,15	3,55	3,85
25	1,32	1,71	2,06	2,49	2,79	3,08	3,45	3,73
30	1,31	1,70	2,04	2,46	2,75	3,03	3,39	3,65
40	1,30	1,68	2,02	2,42	2,70	2,97	3,31	3,55
50	1,30	1,68	2,01	2,41	2,68	2,94	3,27	3,50

Продовження табл.Б.2

60	1,30	1,67	2,00	2,39	2,66	2,91	3,23	3,46
70	1,30	1,67	2,00	2,38	2,65	2,89	3,21	3,44
80	1,30	1,66	1,99	2,37	2,64	2,88	3,19	3,42
90	1,29	1,66	1,99	2,37	2,63	2,88	3,18	3,40
100	1,29	1,66	1,98	2,36	2,63	2,87	3,17	3,39
120	1,29	1,66	1,98	2,36	2,62	2,86	3,16	3,37
∞	1,28	1,64	1,96	2,33	2,58	2,81	3,09	3,29

Таблиця Б.3 – Критичні значення статистики ***w*** критерію Вилькосона

m	***n***	w_H при рівнях значимості (двосторонніх)			$2M_w$	***m***	***n***	w_H при рівнях значимості (двосторонніх)			$2M_w$
		1%	***5%***	***10%***				***1%***	***5%***	***10%***	
7	7	32	36	39	105		18	62	72	77	216
	8	34	38	41	112		19	64	74	80	224
	9	35	40	43	119		20	66	77	83	232
	10	37	42	45	126		21	68	79	85	240
	11	38	44	47	133		22	70	81	88	248
	12	40	46	49	140		23	71	84	90	256
	13	41	48	52	147		24	73	86	93	264
	14	43	50	54	154		25	75	89	96	272
	15	44	52	56	161	9	9	56	62	66	171
	16	46	54	58	168		10	58	65	69	180
	17	47	56	61	175		11	61	68	72	189
	18	49	58	63	182		12	63	71	75	198
	19	50	60	65	189		13	65	73	78	207
	20	52	62	67	196		14	67	76	81	216
	21	53	64	69	203		15	69	79	84	225
	22	55	66	72	210		16	72	82	87	234
	23	57	68	74	217		17	74	84	90	243
	24	58	70	76	224		18	76	87	93	252
	25	60	72	78	231		19	78	90	96	261
8	8	43	49	51	136		20	81	93	99	270
	9	45	51	54	144		21	83	95	102	279
	10	47	53	56	152		22	85	98	105	288
	11	49	55	59	160		23	88	101	108	297
	12	51	58	62	168		24	90	104	111	306
	13	53	60	64	176		25	92	107	114	315
	14	54	62	67	184	10	10	71	78	82	210
	15	56	65	69	192		11	73	81	86	220
	16	58	67	72	200		12	76	84	89	230
	17	60	70	75	208		13	79	88	92	240

Продовження табл.Б.3

m	n	w_H при рівнях значимості (двосторонніх)			$2M_w$	m	n	w_H при рівнях значимості (двосторонніх)			$2M_w$
		1%	*5%*	*10%*				*1%*	*5%*	*10%*	
	14	81	91	96	250		21	136	151	159	408
	15	84	94	99	260		22	139	155	163	420
	16	86	97	103	270		23	142	159	168	432
	17	89	100	106	280		24	146	163	172	444
	18	92	103	110	290		25	149	167	176	456
	19	94	107	113	300	13	13	125	136	142	351
	20	97	110	117	310		14	129	141	147	364
	21	99	113	120	320		15	133	145	152	377
	22	102	116	123	330		16	136	150	156	390
	23	105	119	127	340		17	140	154	161	403
	24	107	122	130	350		18	144	158	166	416
	25	110	126	134	360		19	148	163	171	429
11	11	87	96	100	253		20	151	167	175	452
	12	90	99	104	264		21	155	171	180	455
	13	93	103	108	275		22	159	176	185	468
	14	96	106	112	286		23	163	180	194	481
	15	99	110	116	297		24	166	185	199	494
	16	102	113	120	308		25	170	189	199	507
	17	105	117	123	319	14	14	147	160	166	406
	18	108	121	127	330		15	151	164	171	420
	19	111	124	131	341		16	155	169	176	434
	20	114	128	135	352		17	159	174	182	448
	21	117	131	139	363		18	163	179	187	462
	22	120	135	143	374		19	168	183	192	476
	23	123	139	147	385		20	172	188	197	490
	24	126	142	151	396		21	176	193	202	504
	25	129	146	155	407		22	180	198	207	518
12	12	105	115	120	300		23	184	203	212	532
	13	109	119	125	312		24	188	207	218	546
	14	112	123	129	324		25	192	212	223	560
	15	115	127	133	336	15	15	171	184	192	465
	16	119	131	138	348		16	175	190	197	480
	17	122	135	142	360		17	180	195	203	495
	18	125	139	146	372		18	184	200	208	510
	19	129	143	150	384		19	189	205	214	525
	20	132	147	155	396		20	193	210	220	540

Продовження табл.Б.3

m	*n*	w_H при рівнях значимості (двосторонніх)			$2M_w$	*m*	*n*	w_H при рівнях значимості (двосторонніх)			$2M_w$
		1%	*5%*	*10%*				*1%*	*5%*	*10%*	
	21	198	216	225	555		24	313	337	350	836
	22	202	221	231	570		25	319	344	357	855
	23	207	226	236	585	20	20	315	337	348	820
	24	211	231	242	600		21	322	344	356	840
	25	216	237	248	615		22	328	351	364	860
16	16	196	211	219	528		23	335	359	371	880
	17	201	217	225	544		24	341	366	379	900
	18	206	222	231	560		25	348	373	387	920
	19	210	228	237	576	21	21	349	373	385	903
	20	215	234	243	592		22	256	381	393	924
	21	220	239	249	608		23	363	388	401	945
	22	225	245	255	624		24	400	396	410	966
	23	230	251	261	640		25	408	404	418	987
	24	235	256	267	656	22	22	386	411	424	990
	25	240	262	273	672		23	393	419	434	1012
17	17	223	240	249	595		24	400	427	441	1034
	18	228	246	255	612		25	408	435	450	1056
	19	234	252	262	629	23	23	424	451	465	1081
	20	239	258	268	646		24	431	459	474	1104
	21	244	264	274	663		25	439	468	483	1127
	22	249	270	281	680	24	24	464	492	507	1176
	23	255	276	287	697		25	472	501	517	1200
	24	260	282	294	714	25	25	505	536	552	1275
	25	265	288	300	731						
18	18	252	270	280	666						
	19	258	277	287	684						
	20	263	283	294	702						
	21	269	290	301	720						
	22	275	296	307	738						
	23	280	303	314	756						
	24	286	309	321	774						
	25	292	316	328	792						
19	19	283	303	313	741						
	20	289	309	320	760						
	21	295	316	328	779						
	22	301	323	335	798						
	23	307	330	342	817						

ДОДАТОК В

Таблиця В.1 – Значення χ^2 відповідно значенням рівнів значимості $P(\chi^2)$ та кількостям степенів вільності l

l	$P(\chi^2)$										
	0,99	0,95	0,90	0,50	0,25	0,10	0,05	0,025	0,01	0,005	0,001
1	0,02	$0,0^24$	0,02	0,46	1,32	2,71	3,84	5,02	6,63	7,88	10,8
2	0,02	0,10	0,21	1,39	2,77	4,61	5,99	7,38	9,21	10,6	13,8
3	0,12	0,35	0,58	2,37	4,11	6,25	7,81	9,35	11,3	12,8	16,3
4	0,30	0,71	1,06	3,36	5,39	7,78	9,49	11,1	13,3	14,9	18,5
5	0,55	1,15	1,61	4,35	6,63	9,24	11,1	12,8	15,1	16,7	20,5
6	0,87	1,64	2,20	5,35	7,84	10,6	12,6	14,4	16,8	18,5	22,5
7	1,24	2,17	2,83	6,35	9,04	12,0	14,1	16,0	18,5	20,3	24,3
8	1,65	2,73	3,49	7,34	10,2	13,4	15,5	17,5	20,1	22,0	26,1
9	2,09	3,33	4,17	8,34	11,4	14,7	16,9	19,0	21,7	23,6	27,9
10	2,56	3,94	4,87	9,34	12,5	16,0	18,3	20,5	23,2	25,2	29,6
11	3,05	4,57	5,58	10,3	13,7	17,3	19,7	21,9	24,7	26,8	31,3
12	3,57	5,23	6,30	11,3	14,8	18,5	21,0	23,3	26,2	28,3	32,9
13	4,11	5,89	7,04	12,3	16,0	19,8	22,4	24,7	27,7	29,8	34,5
14	4,66	6,57	7,79	13,3	17,1	21,1	23,7	26,1	29,1	31,3	36,1
15	5,23	7,26	8,55	14,3	18,2	22,3	25,0	27,5	30,6	32,8	37,7
16	5,81	7,96	9,31	15,3	19,4	23,5	26,3	28,8	32,0	34,3	39,3
17	6,41	8,67	10,1	16,3	20,5	24,8	27,6	30,2	33,4	35,7	40,8
18	7,01	9,39	10,9	17,3	21,6	26,0	28,9	31,5	34,8	37,2	42,3
19	7,63	10,1	11,7	18,3	22,7	27,2	30,1	32,9	36,2	38,6	43,8
20	8,26	10,9	12,4	19,3	23,8	28,4	31,4	34,2	37,6	40,0	45,8
21	8,90	11,6	13,2	20,3	24,9	29,6	32,7	35,5	38,9	41,4	46,8
22	9,54	12,3	14,0	21,3	26,0	30,8	33,9	36,8	40,3	42,8	48,3
23	10,2	13,1	14,8	22,3	27,1	32,0	35,2	38,1	41,6	44,2	49,7
24	10,9	13,8	15,7	23,3	28,2	33,2	36,4	39,4	43,0	45,6	51,2
25	11,5	14,6	16,5	24,3	29,3	34,4	37,7	40,6	44,3	46,9	52,6
26	12,2	15,4	17,3	25,3	30,4	35,6	38,9	41,9	45,6	48,3	54,1
27	12,9	16,2	18,1	26,3	31,5	36,7	40,1	43,2	47,0	49,6	55,5
28	13,6	16,9	18,9	27,3	32,6	37,9	41,3	44,5	48,3	51,0	56,9
29	14,3	17,7	19,8	28,3	33,7	39,1	42,6	45,7	49,6	52,3	58,3
30	15,0	18,5	20,6	29,3	34,8	40,3	43,8	47,0	50,9	53,7	59,7

З М І С Т

Навчальне видання

МЕТОДИЧНІ ВКАЗІВКИ ДО ВИКОНАННЯ ПРАКТИЧНИХ РОБІТ З ДИСЦИПЛІНИ «**МАТЕМАТИЧНІ МЕТОДИ В ГІДРОМЕТЕОРОЛОГІЇ**»

ДЛЯ СТУДЕНТІВ ГЕОГРАФІЧНОГО ФАКУЛЬТЕТУ

НАПРЯМ: 6.040105 ГІДРОМЕТЕОРОЛОГІЯ
СПЕЦІАЛЬНІСТЬ: ГІДРОЛОГІЯ ТА ГІДРОХІМІЯ

Упорядник Лук’янець Ольга Іванівна