

Міністерство освіти і науки, молоді та спорту України  
Харківський національний університет імені В. Н. Каразіна

# **СТАТИСТИЧНІ ТА ХЕМОМЕТРИЧНІ МЕТОДИ В ХІМІЇ**

Навчальний посібник

Харків – 2012

УДК 54:519.2(075.8)  
ББК 24в641я73  
С 76

**Рецензенти:**

**В. М. Зайцев** – член-кореспондент НАН України, д. х. н., професор, завідувач кафедри аналітичної хімії Київського національного університету імені Тараса Шевченка;

**М. В. Косевич** – д. ф.-м. н., провідний науковий співробітник Фізико-технічного інституту низьких температур імені Б. І. Веркіна НАН України;

**М. О. Мchedlov-Петросян** – д. х. н., професор, завідувач кафедри фізичної хімії Харківського національного університету імені В. Н. Каразіна.

*Затверджено до друку рішенням Вченої ради  
Харківського національного університету імені В. Н. Каразіна  
(протокол № 4 від 23.03.2012 р.)*

**Статистичні та хемометричні методи в хімії: навчальний посібник /**  
С 76 А. В. Пантелеймонов, І. В. Христенко, В. В. Іванов та ін. – Х. : ХНУ імені В. Н. Каразіна, 2012. – 40 с.

У посібнику наведено програму та структуру навчальної дисципліни, типові задачі для розв'язання, алгоритми розв'язання завдань, приклади контрольних завдань. Посібник розрахований на аудиторне та самостійне виконання завдань.

Електронна версія посібника розміщена на сайті кафедри хімічного матеріалознавства [www-chemo.univer.kharkov.ua](http://www-chemo.univer.kharkov.ua).

**УДК 54:519.2(075.8)  
ББК 24в641я73**

© Харківський національний університет імені В. Н. Каразіна, 2012

© Пантелеймонов А. В., Христенко І. В., Іванов В. В., Холін Ю. В., 2012

© Дончик І. М., макет обкладинки, 2012

## ЗМІСТ

ВСТУП.....	4
Програма навчальної дисципліни.....	5
Структура навчальної дисципліни.....	7
Розподіл балів, які отримують студенти.....	7
Шкала оцінювання.....	7
<b>I. Проста обробка експериментальних даних.</b>	
Апроксимація залежностей. Робота з функціями.	
Вивчення властивостей деяких неперервних розподілів випадкових величин.....	8
<i>Контрольна робота № 1</i> .....	15
<b>II. Деякі статистичні методи перевірки гіпотез.</b>	
Перевірка нормальності розподілу результатів хімічного аналізу за допомогою критерію $\chi^2$ .....	16
<i>Контрольна робота № 2</i> .....	21
<b>III. Перевірка гіпотез про функції розподілу ймовірності в задачах якісного аналізу.....</b>	22
<b>IV. Сингулярний розклад матриць у регресійному аналізі.....</b>	25
<i>Контрольна робота № 3</i> .....	28
Екзаменаційний білет письмового екзамену.....	29
Додаток.....	31
Рекомендована література.....	38

## ВСТУП

Хімія – експериментальна наука, і для хімічних досліджень типовим є виконання повторних вимірювань. Коректна обробка масиву отриманих даних потребує певних знань щодо статистичної природи результатів вимірювань. Отже, курс статистики є необхідним етапом загальної системи хімічної освіти. Такий загальний курс – «Статистичні та хемометричні методи в хімії» – читається і на хімічному факультеті Харківського національного університету імені В. Н. Каразіна. Він складається з курсу лекцій та практичних занять, які звичайно виконуються за допомогою комп'ютерів. Цей навчальний посібник є збірником завдань, що ілюструють основні статистичні концепції, що є важливими в сучасних хімічних дослідженнях.

Слід зазначити, що на сьогодні теоретична статистика пропонує широкий набір методів дослідження чисельних масивів, починаючи від «елементарної статистичної обробки даних» і закінчуючи складними хемометричними підходами на кшталт «DATA MINING». У навчальному посібнику ми орієнтувалися, перш за все, на стандартні статистичні проблеми, що виникають в аналітичних та фізико-хімічних дослідженнях. Складні питання, пов'язані з «вилученням знань» із хімічних даних, розглядаються в спеціальних курсах, що читаються на кафедрі хімічного матеріалознавства.

Посібник розрахований на студентів хімічних факультетів, які у навчальних курсах аналітичної, фізичної, колоїдної, теоретичної або органічної хімії використовують засоби обробки даних. У посібнику наведені алгоритми розв'язку основних задач, що виникають у студентів хімічних факультетів при виконанні лабораторних робіт нормативних курсів.

На сьогодні існує величезна кількість різноманітних статистичних пакетів та програм, які дозволяють провести усі необхідні розрахунки. Для розв'язання задач, розглянутих у посібнику, можна використовувати як універсальні загальнонаукові статистичні пакети (Statistica, SPSS та інші), так і більш компактні програми, значною мірою орієнтовані на графічну обробку даних. Останні, на наш погляд, найбільш пристосовані для цілей навчального курсу статистики. Серед таких програм слід згадати, перш за все, ORIGIN, CurvExpert, SciDAVis та QtiPlot. Ці програми значною мірою є аналогічними, з близькою системою функцій та команд. Можливості програм дозволяють не лише будувати графіки і діаграми, але й знаходити апроксимуючі функції (в тому числі і нелінійні), а також програмувати досить складні математичні вирази. Крім того, SciDAVis та QtiPlot безкоштовно розповсюджуються в мережі Інтернет. Значна частина завдань, наведених у посібнику, може бути виконана і за допомогою «електронних таблиць» – EXCEL, OpenOffice.

Звичайно, що в сучасних дослідженнях використовуються комп'ютерні методи аналізу статистичних даних. Але для більш глибокого розуміння предмета важливою є також «ручна» обробка статистичних даних. Тому в посібнику наведено певну кількість і таких завдань.

У посібнику наведені також програма та структура навчальної дисципліни, система оцінювання знань, приклади контрольних завдань.

## Програма навчальної дисципліни

### Модуль 1. Лекції

№ теми	Зміст теми
1.	Експериментальні дані. Фактографічна та бібліографічна хімічна інформація. «Дані»: визначення, типи. Шкали: порядкова; відношень; інтервальна. Значення хемометрії та інформатики для хімії (хімічний аналіз, параметрична ідентифікація моделей, QSAR, молекулярна інформатика, автоматизація обробки даних експерименту тощо). Джерела хімічної інформації, бази даних і пакети прикладних програм (Science Citation Index, CSD тощо). Хімічні періодичні видання. Друковані та електронні версії. Імпакт-фактори журналів. Пошук інформації в мережі Інтернет. Статистика, хімічна інформатика, хемометрія.
2.	Представлення та статистична обробка даних. Первинне представлення даних. Дескриптивне представлення даних. Гістограми. Результат вимірювань як випадкова величина. Генеральна сукупність та вибірка. Вибіркові оцінки. Моменти. Середнє. Стандартне відхилення. Дисперсія. Коваріаційні матриці. Коефіцієнти кореляції. Перетворення даних (масштабування, автомасштабне перетворення). Задачі обробки первинних експериментальних даних: дослідження однорідності вибірки, визначення функції розподілу, кореляцій між змінними, класифікація, факторний аналіз. Структурна і параметрична ідентифікація моделей, перевірка адекватності. Статистичні розподіли випадкової величини. Дискретні й неперервні випадкові величини. Біноміальний розподіл. Розподіли неперервних величин: рівномірний, Гауса, Лапласа, Пуассона, $\chi^2$ . Центральна гранична теорема. Метод максимуму правдоподібності. Функція правдоподібності. Правдоподібні оцінки параметрів генеральної сукупності при нормальному та Лапласівському розподілах похибок.
3.	Перевірка статистичних гіпотез. Задача перевірки статистичних гіпотез. Схема перевірки гіпотези. Помилки I та II родів. Потужність критеріїв. Перевірка гіпотез про функції розподілу. Критерій $\chi^2$ , графічні способи перевірки гіпотез про функції розподілу.
4.	Основи кореляційного та регресійного аналізу. Кореляційний аналіз. Приклади кореляцій в хімії, значення кореляцій. Принцип лінійності вільних енергій як основа багатьох хімічних кореляцій. Теоретичні засади методу найменших квадратів (МНК) та статистичні властивості оцінок МНК. Розрахункова схема МНК. Вибір найкращого набору регресорів: методи всіх регресій, покрокової регресії, вилучення регресорів. Приклади використання МНК у хімічних задачах. Лінійний та нелінійний МНК як приклад некоректної задачі (теоретичний аналіз та приклади), мультиколінеарність. Її формальні та неформальні причини. Способи подолання мультиколінеарності: $\alpha$ регуляризація за Тихоновим, застосування ортогональних поліномів, сингулярний розклад. Типові приклади математично некоректних задач в хімії. Вплив викидів на

	оцінки МНК. Уявлення про робастні оцінки.
5.	Класифікація та кластерний аналіз. Види класифікацій та їх значення для аналізу даних. Типи ознак. Міри схожості об'єктів. Класифікація без навчання. Ієрархічна класифікація, дендрограми. Найпростіші алгоритми ієрархічного кластерного аналізу сукупності об'єктів. Проблема стійкості класифікації. Факторний аналіз. Характеристика моделей з латентними змінними. Кореляційна та коваріаційна матриці – об'єкт факторного аналізу. Формулювання задачі факторного аналізу. Матриця навантажень, вектори характеристик і загальностей. Основна факторна теорема. Експлораторний та конфірматорний факторний аналіз. Алгоритми факторного аналізу. Факторний аналіз хроматографічних даних.
<b>Модуль 2. Лабораторні заняття</b>	
6.	Елементарні засоби апроксимації експериментальних залежностей. Робота з програмними засобами.
7.	Вивчення властивостей деяких неперервних розподілів випадкових величин.
8.	Перевірка нормальності розподілу випадкових величин за критерієм $\chi^2$ .
9.	Апроксимація концентраційних частот виявлення неспадними функціями.
10.	Сингулярний розклад.
11.	Підсумкове заняття.

## Структура навчальної дисципліни

№ теми	Кількість годин					
	Усього	У тому числі				
		лекції	практичні	лабораторні	індивідуальні	самостійна робота
1.	9	3				6
2.	12	4				8
3.	12	4				8
4.	12	4				8
5.	9	3				6
6.	12			4		8
7.	24			8		16
8.	24			8		16
9.	24			8		16
10.	18			6		12
11.	6			2		4
Усього годин	162	36		36		108

## Розподіл балів, які отримують студенти

Поточне тестування та самостійна робота	Модуль 1	Модуль 2					
	Теми 1-5	T6	T7	T8	T9	T10	T11
		контрольна робота					
		10	25	25			
Підсумковий семестровий контроль (екзамен)	40						
Сума	100						

## Шкала оцінювання

Сума балів за всі види навчальної діяльності протягом семестру	Оцінка ECTS	Оцінка за національною шкалою
90 – 100	A	відмінно
80 – 89	B	добре
70 – 79	C	
60 – 69	D	задовільно
50 – 59	E	
1 – 49	FX	незадовільно

**I. Проста обробка експериментальних даних.  
Апроксимація залежностей. Робота з функціями.  
Вивчення властивостей деяких неперервних  
розподілів випадкових величин**

**Вказівки до розв'язання задач**

Диференціювання та інтегрування функцій проводять за допомогою меню «Analysis».

Встановлення значень змінних з певним кроком та розрахунок значень функцій проводяться за допомогою меню «Column → Set column values».

Для апроксимації графіків функцій поліноміальними моделями використовується меню «Analysis → Fit polynomial».

Розрахунок підгінних параметрів графіків функцій нелінійним методом найменших квадратів (МНК) виконується за допомогою меню «Analysis».

**Завдання для аудиторного та самостійного розв'язання**

1. Наведіть у графічному вигляді залежність оптичної густини розчину бетаїнового індикатору  $A$  від довжини хвилі  $\lambda$  (таблиця I.1). Побудуйте графіки трьох типів: точковий, лінійний, лінійний з точками. Для лінійного графіка змініть тип з'єднувальної лінії з суцільної на пунктирну. Згладьте лінійний з точками графік сплайн-функцією. На графіку підпишіть вісі.

2. Наведіть у графічному вигляді залежність оптичної густини двох розчинів індикатору метилового оранжевого  $A_1$  та  $A_2$  від довжини хвилі  $\lambda$  (таблиця I.2). Оберіть залежність (лінійна, поліноміальна, експоненціальна тощо), що адекватно апроксимує експериментальні дані для кожного з розчинів. Знайдіть максимуми поглинання кожного з розчинів.

Таблиця I.1

$\lambda$ , нм	$A$
316	1.053
330	1.956
350	2.396
380	0.993
400	0.444
430	0.167
450	0.215
470	0.332
500	0.546
520	0.655
550	0.711
570	0.66
600	0.474
630	0.265
650	0.154
700	0.03



Таблиця І.2

$\lambda$ , нм	$A_1$	$A_2$
350	0.004	0.103
375	0.006	0.396
400	0.067	0.835
425	0.257	1.144
450	0.736	1.372
475	1.549	1.338
500	1.961	0.906
525	1.93	0.402
550	1.148	0.061
575	0.138	0.05

3. Побудуйте залежність логарифмів констант протонізації двох слабких основ  $\lg K_1$  та  $\lg K_2$  від іонної сили розчинів  $I$  (таблиця І.3). Оберіть залежність (лінійна, поліноміальна, експоненціальна тощо), що адекватно апроксимує експериментальні дані для кожного з розчинів. Знайдіть значення термодинамічної константи протонізації кожної з основ.

4. Побудуйте залежність рН розчину оцтової кислоти від об'єму доданого титранту  $V$  (NaOH) при потенціометричному титруванні (таблиця І.4). Знайдіть точку еквівалентності.

Таблиця І.3

$I$ , ммоль/л	$\lg K_1$	$\lg K_2$
0.5	7.26	7.28
0.10	7.09	7.11
0.25	6.79	6.76
0.30	6.77	6.68
0.51	6.42	6.46
0.77	6.30	6.34
0.82	6.26	6.32
1.03	6.35	6.31

Таблиця І.4

$V$ (NaOH), мл	рН
0	3
1	3.7
2	4.1
3	4.3
4	4.7
5	4.9
6	5.2
7	5.5
7.5	5.9
8	10
8.5	11
9	11.4
10	11.5

5. Фрагмент ПМР спектру олеїнової кислоти ( $C_{17}H_{33}COOH$ ) надано в таблиці І.5. Відомо, що один атом Н має інтенсивність 16 відносних одиниць. Визначте, яка функціональна група дає вказаний фрагмент спектру.

Таблиця І.5

$\delta$ , м. д.	$I$ , відн. од.
0.725	0.838
0.750	4.255
0.775	16.827
0.800	51.831
0.825	124.337
0.850	232.292
0.875	337.983
0.900	382.985
0.925	337.983
0.950	232.292
0.975	124.337
1.000	51.831
1.025	16.827
1.050	4.255
1.075	0.838

6. Побудуйте спектри флуоресценції комплексів кальцеїну ( $C_{13}H_{26}N_2O_{13}$ ) з іонами нікелю в лужному ( $A_1$ ) та кислому ( $A_2$ ) середовищах (таблиця І.6). Побудуйте спектри нормалізованої флуоресценції.

Таблиця І.6

$\lambda$ , нм	$A_1$	$A_2$
470	112.3	103.3
480	100.1	126.4
490	192.6	250.5
500	382.4	449.5
510	577.4	638.7
520	653.8	738.1
530	675.9	792.2
540	684.3	824.5
550	656.3	805.4
560	565.3	711.1
570	459.7	604.2
580	373.3	508
590	297.2	411.5
600	245.5	337.3
610	192.8	268.7
620	160	221
630	135.9	186.9
640	114.8	156.5
650	101.9	136.3
660	91.45	119.5
670	107.9	149
680	85.76	106
690	82.39	100.5
698	79.94	95.55

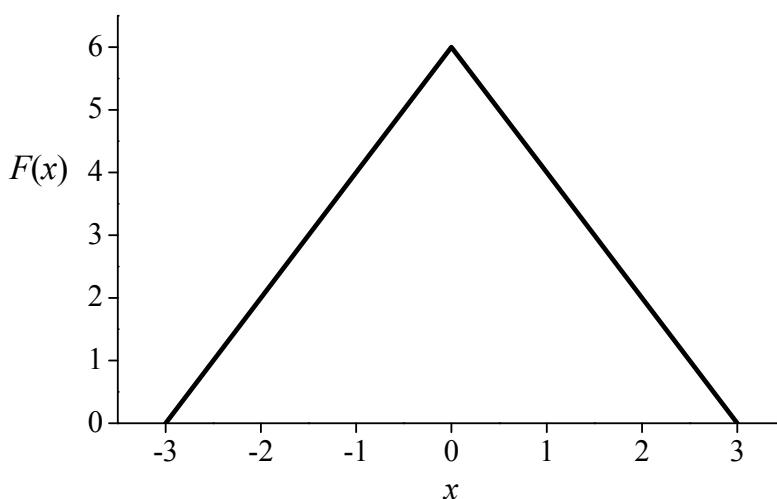
6. Змінна приймає значення  $x = [-0.45; -0.30; \dots; 6.90]$ . Розрахуйте значення функції  $y(x) = \sin\left(\frac{\pi x}{1.5}\right) \cdot e^{-0.5x}$ . Апроксимуйте наведену функцію поліномом в інтервалі  $[0; 2]$  так, щоб похибка апроксимації була менше 0.1. Знайдіть площу під графіком функції у вказаному інтервалі і порівняйте отримане значення з результатом аналітичного розрахунку інтегралу  $\int_0^2 y(x) dx$ .

7. Плоска спіраль описується параметричною кривою  $\begin{cases} x = s \cdot \sin(s) \\ y = s \cdot \cos(s) \end{cases}$ , де  $s = [0; 0.1; \dots; 6\pi]$ . Побудуйте графік плоскої спіралі.

8. Побудуйте графік  $y = \begin{cases} x + 4, x \in [-4, -2) \\ 2, x \in [-2, 2] \\ -x + 4, x \in (2, 4] \end{cases}$

Знайдіть площу під графіком функції. Порівняйте отримане значення з розрахунком за формулою трапеції.

9. Побудуйте графік, що представлено на рис. 1. Знайдіть площу під графіком функції в інтервалах  $[-3, 3]$  та  $[-1, 1]$ . Порівняйте отримані значення з результатами аналітичних розрахунків.



**Рисунок 1**

10. Дані з таблиці I.7 можна апроксимувати функцією  $y(x) = \alpha + e^{-\beta x}$ . За допомогою нелінійного методу найменших квадратів знайдіть параметри даної функції.

Таблиця І.7

$x$	$y$
0.35	0.60
0.45	0.50
0.55	0.43
0.65	0.38
0.75	0.33
0.85	0.30
0.95	0.27
1.05	0.25
1.15	0.24
1.25	0.22
1.35	0.21
1.45	0.21
1.55	0.20
1.65	0.20
1.75	0.19

11. У таблиці І.8 наведено залежність енергії зв'язку ( $E$ , ат. од.) двохатомної молекули від відхилень довжини зв'язку від рівноважного значення ( $r$ , ат. од.). Енергію двохатомної молекули можна описати функцією Морзе:

$$E(r) = \alpha(1 - \exp(-\beta r))^2,$$

узагальненою функцією Морзе  $E(r) = \alpha(1 - \exp(-\beta r))^2 + \gamma(1 - \exp(-\delta r))^3$

або функцією Рідберга  $E(r) = \alpha(1 + \beta_1 r + \beta_2 r^2 + \beta_3 r^3) \exp(-\beta_1 r)$ .

Знайдіть параметри вказаних функцій. Яка з цих функцій найбільш адекватно описує дисоціацію двохатомної молекули? Дані з таблиці запишіть у 2 стовпчиках –  $r$  та  $E$ , і, таким чином, отримаєте 30 рядків даних.

Таблиця І.8

$r$	$E$	$r$	$E$
-0.5375	0.19702	0.4625	0.03858
-0.4375	0.11031	0.6625	0.06448
-0.3375	0.05601	0.8625	0.09012
-0.2875	0.03768	1.0625	0.11402
-0.2375	0.02388	1.2625	0.13546
-0.1875	0.01385	1.4625	0.15411
-0.1375	0.00695	1.6625	0.16986
-0.0875	0.00263	1.8625	0.18275
-0.0375	$4.4978 \cdot 10^{-4}$	2.0625	0.19299
-0.0125	$4.7634 \cdot 10^{-5}$	2.2625	0.2009
0	0	2.4625	0.20686
0.0125	$4.8132 \cdot 10^{-5}$	2.7625	0.21297
0.0625	0.00111	3.2625	0.2183
0.1625	0.00666	4.2625	0.22148
0.2625	0.01547	5.2625	0.22204

12. Функція сигмоїда  $f(x) = \frac{1}{1 + \exp(-2 \cdot x)}$  має важливе значення для побудови штучних нейронних мереж. Для  $x = [-3.0; -2.7; \dots; 3.0]$  розрахуйте значення сигмоїди, знайдіть її похідну в кожній точці, апроксимуйте графік похідної функцією густини ймовірності Гауса та знайдіть площу під графіком апроксимуючої функції на всьому інтервалі значень  $x$ .

13. Для  $x = [1.0; 1.5; \dots; 6.0]$  побудуйте на одному графіку густини ймовірності експоненціального розподілу  $f(x) = \frac{1}{b} \exp\left(-\frac{x}{b}\right)$  з параметрами  $b = 1; 2; 4$ . Знайдіть площі під кривими.

14. Для випадкової змінної  $x = [-3.0; -2.4; \dots; 3.0]$ :

а) знайдіть середнє значення, стандартне відхилення та стандартне відхилення середнього;

б) побудуйте на одному графіку густини ймовірності розподілу Гауса  $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  зі вказаними параметрами (середнє ( $\mu$ ); дисперсія ( $\sigma$ )): (0; 1), (0; 0.5), (0; 2), (1; 1). Знайдіть площі під кривими;

в) для густини ймовірності розподілу Гауса з параметрами (0; 1) знайдіть інтеграли  $\int_{-1}^1 f(x)dx$ ,  $\int_{-2}^2 f(x)dx$ ,  $\int_{-3}^3 f(x)dx$ ,  $\int_0^3 f(x)dx$ . Який фізичний зміст цих величин?

г) побудуйте на одному графіку функції густини ймовірності логістичного розподілу  $f(x) = \frac{\exp\left(\frac{x-a}{k}\right)}{k \cdot \left(1 + \exp\left(\frac{x-a}{k}\right)\right)^2}$  з параметрами ( $a; k$ ): (0; 1), (0; 0.5), (0; 2). Порівняйте їх з кривими густини ймовірності розподілу Гауса.

15. У розчині присутні дві забарвлені речовини. В таблиці І.9 наведено залежність світлопоглинання розчину від довжини хвилі.

а) Побудуйте точковий графік.

б) Апроксимуйте кожний з піків окремим поліномом.

в) Для кожного з піків розрахуйте площу під кривою.

г) Розрахуйте значення  $\hat{A}$  для сумарної кривої (за двома поліномами).

д) Апроксимуйте експериментальні дані функціями

$$\hat{A} = \alpha_1 \exp(-\beta_1(\lambda - \bar{\lambda}_1)^2) + \alpha_2 \exp(-\beta_2(\lambda - \bar{\lambda}_2)^2) + \gamma \text{ та}$$

$$\hat{A} = \alpha_1 \exp(-\beta_1|\lambda - \bar{\lambda}_1|) + \alpha_2 \exp(-\beta_2|\lambda - \bar{\lambda}_2|) + \gamma, \text{ знайдіть значення підгінних}$$

параметрів для кожної функції ( $\bar{\lambda}_1$  та  $\bar{\lambda}_2$  – максимуми поглинання відповідних речовин).

е) Для кожного з отриманих наборів  $\hat{A}$  (три набори) знайдіть значення критерію  $U = \sum_{i=1}^N (A_i - \hat{A}_i)^2$  ( $i$  – номер експериментальної точки,  $N$  – число експериментальних точок).

Таблиця І.9

$\lambda$	$A$
410	1.053
414	1.956
424	2.396
434	0.993
445	0.444
456	0.167
468	0.215
478	0.332
489	0.546
502	0.655
512	0.711
522	0.66
534	0.474
546	0.265
556	0.154

**Контрольна робота № 1**  
**Тема «Проста обробка експериментальних даних»**

1. При потенціометричному титруванні отримали такі дані:

$V$ , мл	1.5	1.8	1.9	1.95	1.98	2.00	2.02	2.05	2.10
pH	2.64	3.05	3.36	3.64	4.05	6.98	9.95	10.53	10.65

Точка еквівалентності спостерігається при об'ємі титранту \_\_\_\_\_.

2. Ступінь полінома, необхідний для апроксимації функції

$f(x) = \frac{\exp(x) + \sqrt{x}}{x^2}$  для  $x = 0.2, 0.4, \dots, 2.2$  (за умови, що стандартне відхилення апроксимації задовольняє подвійній нерівності  $0.5 < \sigma < 2$ ), дорівнює \_\_\_\_\_.

3. Для випадкової величини  $x$ , що приймає значення в інтервалі 6, ..., 8 (крок 0.1), знайдіть середнє значення (\_\_\_\_\_), дисперсію (\_\_\_\_\_ ) і стандартне відхилення середнього (\_\_\_\_\_ ) (в дужках впишіть значення).

Площа під кривою Гауса (використайте для побудови знайдені параметри) на всьому інтервалі зміни випадкової величини складає \_\_\_\_\_, а в інтервалі [6.5; 7.5] – \_\_\_\_\_.

4. У дослідженні області ненадійності тестової реакції отримали такі дані для залежності ймовірності виявлення компонента  $P$  від його концентрації  $c$ , мг/л:

$c$ , мг/л	1.9	2	2.1	2.3	2.4	2.5	2.6
$P$	0.2	0.53	0.51	0.63	0.47	0.79	0.91

Припускаючи, що ймовірність виявлення компонента описується функцією експоненціального розподілу, оцініть значення параметрів цієї функції і запишіть їх нижче:

Знайдіть значення критерію  $U = \sum_{i=1}^N (P - \hat{P})^2$ , де  $i$  – номер експериментальної точки,  $N$  – загальна кількість експериментальних точок,  $P$  – експериментальне значення ймовірності виявлення,  $\hat{P}$  – ймовірність виявлення, розрахована з використанням одержаних оцінок параметрів.

$U =$  \_\_\_\_\_.

Необхідні формули:

Густина розподілу ймовірностей Гауса:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Функція експоненціального розподілу:

$$P = 1 - \exp\left(-\frac{c-a}{b}\right).$$

## II. Деякі статистичні методи перевірки гіпотез. Перевірка нормальності розподілу результатів хімічного аналізу за допомогою критерію $\chi^2$

### Вказівки до розв'язання задач

Статистика  $\chi^2$  – міра узгодженості між певним теоретичним розподілом і емпіричним розподілом результатів.

Значення  $\chi_{\text{експ}}^2 = \sum_{i=1}^K \frac{(v_i - v_i^{\text{теор}})^2}{v_i^{\text{теор}}}$ , де  $i$  – номер біну,  $K$  – кількість бінів,  $v$  –

частоти потрапляння вимірів у даний бін, необхідно порівняти з критичним значенням  $\chi_{f, \alpha}^2$ , яке розраховується для  $f$  ступенів свободи та рівня значущості  $\alpha$  (зазвичай, 5%). Гіпотеза про те, що результати вимірювань не суперечать визначеному закону розподілу, приймається, якщо  $\chi_{\text{експ}}^2 < \chi_{f, \alpha}^2$ . Таблиця квантилей розподілу  $\chi^2$  наведена у додатку.

Для перевірки відповідності між певним теоретичним розподілом та емпіричним розподілом результатів можна використовувати критерій Романовського:  $Q = \frac{|\chi^2 - k|}{\sqrt{2k}}$ , де  $k = n_b - 2$ ,  $n_b$  – кількість бінів. Якщо  $Q \leq 3$ , то теоретичний розподіл відповідає емпіричним даним, у протилежному випадку (при  $Q > 3$ ) обраний теоретичний розподіл не відповідає емпіричному.

Певними показниками форми емпіричного розподілу є значення коефі-

цієнтів асиметрії  $\tilde{A} = \frac{\sum_{i=1}^K v_i (x_i - \bar{x})^3}{N\sigma^3}$  та ексцесу  $\gamma_2 = \frac{\sum_{i=1}^K v_i (x_i - \bar{x})^4}{N\sigma^4} - 3$ , де  $x_i$  – координати центру біну,  $\bar{x}$  – середнє значення для вибірки,  $N$  – загальна кількість значень у вибірці. Дисперсії цих величин визначаються за формулами:  $\sigma^2(\tilde{A}) = \frac{6(N-1)}{(N+1)(N-3)}$  та  $\sigma^2(\gamma_2) = \frac{24(N-2)(N-3)}{(N+1)^2(N+3)(N+5)}$ . Розподіл результатів можна

вважати нормальним, якщо виконується нерівність  $|\tilde{A}| \leq 3\sigma(\tilde{A})$  та  $|\gamma_2| \leq 5\sigma(\gamma_2)$ .

Оцінку такого типу використовують у випадку вибірки з  $N < 20$ .

Задача розв'язується як у графічних пакетах (ORIGIN, SciDAVis), так і за допомогою електронних таблиць (Open Office, EXCEL).

З наведеної вище формули для розрахунку статистики  $\chi^2$  видно, що для отримання критерію необхідно розрахувати теоретичні значення частот потрапляння вимірів до даного біну. Алгоритм їх розрахунку такий.



ORIGIN, SciDAVis	EXCEL, Open Office	
1. Розрахуйте значення функції густини розподілу Гауса в кожній точці центру біну (зі знайденими раніше параметрами).	1. Значення функції густини розподілу Гауса в кожній точці центру біну розраховуються за допомогою вбудованої функції нормального розподілу диференційного вигляду (зі знайденими раніше параметрами).	1. Розрахуйте значення функції розподілу Гауса в кожній точці кінця біну (зі знайденими раніше параметрами) за допомогою вбудованої функції нормального розподілу інтегрального вигляду.
2. Знайдіть добуток густини розподілу Гауса та ширини біну. Отримані дані – ймовірність потрапляння вимірів у даний бін.	2. Знайдіть ймовірність потрапляння вимірів у даний бін як різницю між $(i+1)$ -ю та $(i)$ -ю ймовірностями.	
3. Розрахуйте добуток ймовірності потрапляння вимірів у даний бін та загальної кількості вимірювань. Отримані значення і є теоретичними частотами потрапляння вимірів до даного біну.		

### Завдання для аудиторного та самостійного розв'язання

1. У таблицях II.1 – II.2 наведено приклади даних, отриманих у міжлабораторному експерименті з визначення речовини. Дані таблиць та інші набори даних розміщені у файлах \*.dat на сайті [www-chemo.univer.kharkov.ua](http://www-chemo.univer.kharkov.ua) (символ \* відповідає номеру таблиці). Для кожного з наборів даних:

а) знайдіть середнє, стандартне відхилення, стандартне відхилення середнього;

б) побудуйте гістограму частот;

в) знайдіть вибіркові значення коефіцієнтів асиметрії та ексцесу;

г) згрупуйте дані так, щоб у кожному біні було не менше 5 вимірювань;

д) розрахуйте значення статистики  $\chi^2$ , порівняйте його з критичним при вказаних викладачем рівнях значущості;

е) зробіть висновок щодо нормальності розподілу результатів аналізу;

ж) розрахуйте вибіркові значення коефіцієнтів асиметрії та ексцесу за допомогою електронних таблиць і порівняйте їх зі знайденими в п. в);

з) побудуйте графіки експериментальних та теоретичних частот.

Таблиця II.1

2.634	2.602	2.749	2.495	2.517	2.363	2.510	2.412	2.531	2.552
2.705	2.314	2.536	2.534	2.265	2.427	2.647	2.361	2.795	2.824
2.445	2.521	2.619	2.582	2.588	2.852	2.433	2.614	2.862	2.311
2.816	2.413	2.453	2.451	2.680	2.635	2.644	2.670	2.629	2.639
2.337	2.471	2.767	2.248	2.723	2.518	2.666	2.376	2.588	2.562
2.771	2.639	2.719	2.813	2.342	2.588	2.750	2.534	2.674	2.467
2.708	2.434	2.546	2.584	2.367	2.465	2.649	2.688	2.731	2.524
2.403	2.371	2.582	2.542	2.638	2.342	2.393	2.537	2.334	2.715
2.561	2.457	2.695	2.588	2.587	2.527	2.547	2.646	2.599	2.441

Таблиця II.1. Продовження

2.586	2.734	2.263	2.364	2.603	2.433	2.385	2.845	2.559	2.324
2.526	2.734	2.687	2.500	2.675	2.230	2.423	2.602	2.495	2.424
2.379	2.669	2.585	2.715	2.582	2.647	2.629	2.492	2.963	2.518
2.443	2.578	2.745	2.761	2.475	2.580	2.589	2.579	2.433	2.685
2.718	2.688	2.549	2.619	2.602	2.760	2.556	2.415	2.546	2.682
2.335	2.283	2.593	2.530	2.373	2.452	2.475	2.511	2.606	2.536
2.500	2.580	2.504	2.606	2.619	2.351	2.431	2.392	2.353	2.400
2.259	2.678	2.374	2.640	2.463	2.543	2.693	2.468	2.819	2.418
2.650	2.527	2.892	2.612	2.690	2.377	2.495	2.675	2.413	2.626
2.599	2.874	2.638	2.541	2.413	2.671	2.717	2.686	2.742	2.604
2.786	2.771	2.514	2.236	2.524	2.806	2.555	2.605	2.438	2.648

Таблиця II.2

12.94	12.99	12.95	13.02
13.07	12.31	13.03	13.11
13.08	12.83	13.07	12.98
13.77	13.9	13.27	13.4
12.74	12.62	13.02	13.13
12.9	12.89	13.15	13.02
12.38	13.27	13.01	12.86
13.13	12.82	12.89	12.77
12.62	12.86	12.62	12.92
12.9	12.95	13.11	12.97
13.0	12.87	12.92	13.13
12.88	13.06	13.02	12.91
13.0	12.94	13.01	12.9
12.82	12.95	12.77	12.52
12.02	12.02	12.9	13.14
12.64	12.9	12.9	12.77
12.9	12.77	13.11	12.97
13.07	12.95	12.7	12.74
12.89	12.52	12.59	12.54
12.81	12.62	12.74	12.62
12.94	12.9	13.01	12.99
13.0	12.97	12.9	13.01
12.82	12.68	12.98	12.95
12.94	13.01	13.27	12.02
13.52	12.52	13.27	12.9
13.53	13.4	13.27	12.77
13.14	13.15	12.94	12.81

Вказівка. Кожне число в наведених таблицях відповідає результату окремого вимірювання. Дані таблиць можуть бути імпортовані з відповідного dat-файлу.

2. Розрахуйте довірчі інтервали для визначення оксиду заліза в залізній руді та нікелю в сталі (експериментальні дані наведені в табл. II.3).

Таблиця II.3

% Fe <sub>2</sub> O <sub>3</sub>	% Ni
38.82	4.85
39.01	4.89
38.72	4.82
38.85	

Вказівка. Розрахунок довірчого інтервалу зазвичай виконують за формулою  $\bar{x} \pm \Delta\bar{x} = \bar{x} \pm \frac{t(P, f) \cdot s}{\sqrt{N}}$ , де  $\bar{x}$  – середнє значення вимірювань,  $s$  – стандартне відхилення,  $N$  – обсяг експериментальної вибірки,  $t(P, f)$  – значення  $t$ -критерію для довірчої ймовірності  $P$  при  $f$  ступенях вільності. Таблиця квантилей  $t$ -розподілу наведена у додатку.

3. Зробіть висновок щодо відтворюваності двох фотометричних методів визначення фосфору. Відомо, що відносне стандартне відхилення першого методу  $s_1 = 4.2\%$ , а другого –  $s_2 = 2.3\%$ . Число вимірювань в обох методах дорівнює 10.

Вказівка. Задачі такого плану розв'язують за допомогою обчислення  $F$ -критерію  $F = \frac{s_1^2}{s_2^2}$ , де  $s_1$  і  $s_2$  – відповідно більше й менше стандартні відхилення, які необхідно порівняти між собою. Висновок про однорідність дисперсій може бути зроблено, якщо  $F < F(P, f_1, f_2)$  для довірчої ймовірності  $P$  при  $f$  ступенях вільності. Таблиця квантилей  $F$ -розподілу наведена у додатку.

4. Дві незалежні групи при визначенні вмісту селену в пробі отримали такі дані (табл. II.4).

Таблиця II.4

Група 1	Група 2
9.31	9.51
9.42	9.43
9.35	9.61
9.37	9.63

Перевірте, чи є статистично значуща різниця між середніми значеннями вмісту селену у пробі для результатів, отриманих двома незалежними групами.

Вказівка. Для перевірки значущості різниці між двома середніми значеннями необхідно, по-перше, перевірити відсутність статистично значущої різниці між їх стандартними відхиленнями (див. п. 3). Якщо  $F < F(P, f_1, f_2)$ ,

розраховують  $t$ -критерій  $t = \frac{|\bar{X}_1 - \bar{X}_2|}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ , де  $\bar{X}_i$  – середнє значення кожної вибірки,

$s_i$  – стандартні відхилення вибірок,  $n_i$  – їх обсяг. У разі, якщо  $t < t(P, f)$ , роблять висновок про відсутність статистично значущої різниці між математичними очікуваннями вибірок. Таблиця квантилей  $t$ -розподілу наведена у додатку.

5. При визначенні вмісту вуглецю в сталі отримали такі дані (табл. II.5). Перевірте, чи не є значення 1.90% таким, яке можна інтерпретувати як грубий промах.

Таблиця II.5

1.90
2.08
2.09
2.10
2.12

Вказівка. Для ідентифікації грубого промаху розраховують співвідношення  $Q = \frac{|x_1 - x_2|}{R}$ , де  $x_1$  – вимірювання, яке є кандидатом на грубий промах,  $x_2$  – сусіднє вимірювання,  $R = x_{\max} - x_{\min}$  – розмах варіювання вибірки. Якщо  $Q < Q(P, n)$ , то припущення щодо інтерпретації  $x_1$  як грубого промаху є статистично необґрунтованими. Числові значення для  $Q(P, n)$  наведені в додатку.

## Контрольна робота № 2

### Тема «Перевірка нормальності результатів хімічного аналізу»

У файлі \*.dat, де \* – номер варіанта, містяться дані про міжлабораторне визначення заліза у питній воді (концентрація заліза, мг/л).

1. Середнє значення складає \_\_\_\_\_, дисперсія \_\_\_\_\_, стандартне відхилення \_\_\_\_\_, стандартне відхилення середнього \_\_\_\_\_.

2. Знайдіть вибіркові значення коефіцієнтів асиметрії (\_\_\_\_\_) та ексцесу (\_\_\_\_\_).

3. При кроці побудови гістограми 0.03 згрупуйте дані так, щоби в кожному біні було не менше 5 вимірювань. Значення статистики  $\chi^2$  при перевірці нормальності розподілу результатів аналізу складає \_\_\_\_\_, критичне значення статистики  $\chi^2$  при рівні значущості 1% і \_\_\_\_ ступенях вільності дорівнює \_\_\_\_\_.

4. Гіпотезу про нормальний розподіл результатів аналізу слід \_\_\_\_\_, тому що \_\_\_\_\_.

5. Для наявних даних укажіть інтервал значень випадкової величини (в численному вигляді), в якому ймовірність виявити одиничний вимір складає 68% (\_\_\_\_\_), 95% (\_\_\_\_\_), 99.7% (\_\_\_\_\_).

Необхідні формули:

$$\text{Густина розподілу Гауса } f(x) = \frac{1}{\sigma(2\pi)^{1/2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right].$$

$$\text{Коефіцієнт асиметрії } \tilde{A} = \frac{\sum_{i=1}^K v_i (x_i - \bar{x})^3}{N\sigma^3}.$$

$$\text{Коефіцієнт ексцесу } \gamma_2 = \frac{\sum_{i=1}^K v_i (x_i - \bar{x})^4}{N\sigma^4} - 3.$$

### III. Перевірка гіпотез про функції розподілу ймовірності в задачах якісного аналізу

#### Вказівки до розв'язання задач

Задачі цього розділу – елемент сучасного якісного аналізу. Для розробки нових методик тест-визначення аналітів необхідно оцінити метрологічні характеристики таких методик, однією з найважливіших метрологічних характеристик є межа виявлення (концентрація аналіту, вище якої ймовірність його виявлення перевищує заданий рівень, зазвичай 95 %).

Ймовірність виявлення компонента  $P(c)$  при концентрації  $c$  розраховується як відношення кількості позитивних відповідей спостерігачів про присутність компонента ( $n_i$ ) до загальної кількості спостерігачів ( $N_i$ ).

Залежність  $P(c)$  – неспадна функція, для апроксимації якої можна використовувати, наприклад, статистичні функції розподілу випадкових величин.

Для розрахунку значення межі виявлення необхідно оцінити параметри апроксимуючих функцій, для чого можна використовувати або нелінійний МНК, або так звані графічні методи оцінки, суть яких складається в лінеаризації залежностей  $P(c)$ . Лінеаризація функції експоненціального розподілу доволі проста, нормальний і логнормальний розподіли лінеаризуються за допомогою вбудованих функцій.

Висновок про те, яка саме функція максимально точно відтворює експериментальні дані, робиться на основі розрахунку коефіцієнтів кореляції або статистики  $\chi^2$ .

#### Завдання для аудиторного та самостійного розв'язання

1. Для даних з таблиць III.1 – III.10 для трьох функцій розподілу ймовірностей (експоненціальної  $P(c) = 1 - \exp\left(\frac{-(c-a)}{b}\right)$ ,

$$\text{нормальної } P(c) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^c \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx$$

$$\text{та логнормальної } P(c) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^c \exp\left[-\frac{(\ln(x/\mu))^2}{2\sigma^2}\right] dx):$$

а) знайдіть параметри кожної з функцій;

б) зробіть висновок про те, яку з вищенаведених функцій доцільно використовувати для апроксимації експериментальних залежностей  $P(c)$ ;

в) знайдіть значення статистики  $\chi^2$  за умови апроксимації експериментальних даних трьома вищезгаданими функціями розподілу, зробіть висновок щодо адекватності моделі (для вказаних викладачем рівнів значущості);

г) для кожної системи оцініть значення межі виявлення на вказаному викладачем рівні ймовірності (використайте для розрахунків всі три функції).

2. За допомогою нелінійного методу найменших квадратів апроксимуйте вказані викладачем експериментальні дані функцією розподілу Вейбула

$P(c) = 1 - \exp\left(\frac{-(c-a)}{b}\right)^k$  та функцією логістичного розподілу

$P(c) = \frac{1}{1 + \exp\left(-\frac{c-k}{t}\right)}$ . Знайдіть їхні параметри, значення статистики  $\chi^2$ , зробіть

висновок про адекватність апроксимації та оцініть значення межі виявлення.

Таблиця III.1

$c_i$ , МГ/Л	$N_i$	$n_i$
$4.3 \cdot 10^{-3}$	300	275
$4.2 \cdot 10^{-3}$	280	260
$4.1 \cdot 10^{-3}$	300	240
$4 \cdot 10^{-3}$	280	181
$3.9 \cdot 10^{-3}$	300	132
$3.8 \cdot 10^{-3}$	280	78
$3.7 \cdot 10^{-3}$	300	34

Таблиця III.2

$c_i$ , МГ/Л	$N_i$	$n_i$
3.7	30	2
3.8	45	11
3.9	45	27
4.0	30	26
4.1	45	38
4.2	30	26
4.3	30	26

Таблиця III.3

$c_i$ , МКМОЛЬ/Л	$N_i$	$n_i$
2	316	30
2.25	300	82
2.5	316	163
2.75	300	205
3	300	259
3.25	300	271
3.5	316	308
3.75	300	296

Таблиця III.4

$c_i$ , МГ/Л	$N_i$	$n_i$
55	228	32
57.5	228	68
60	228	108
62.5	226	147
65	228	188
67.5	221	200
70	134	127
72.5	68	65

Таблиця III.5

$c_i$ , МГ/Л	$N_i$	$n_i$
4.4	50	7
4.5	50	11
4.6	50	20
4.7	50	29
4.8	50	44

Таблиця III.6

$c_i$ , МГ/Л	$N_i$	$n_i$
0.02	168	37
0.04	168	67
0.06	168	95
0.1	168	132
0.14	168	161
0.2	156	154

Таблиця ІІІ.7

$c_i$ , мг/л	$N_i$	$n_i$
8.1	263	57
8.2	263	91
8.3	263	110
8.4	263	120
8.5	263	156
8.6	263	176
8.7	263	184

Таблиця ІІІ.8

$c_i$ , мг/л	$N_i$	$n_i$
30	73	9
40	73	17
60	92	50
80	92	67
100	92	78
150	92	89

Таблиця ІІІ.9

$c_i$ , мг/л	$N_i$	$n_i$
1.90	100	20
2.00	200	106
2.10	100	51
2.20	150	84
2.30	100	63
2.40	200	94
2.50	100	79
2.60	200	182

Таблиця ІІІ.10

$c_i$ , мкмоль/л	$N_i$	$n_i$
9.0	98	2
9.5	42	1
10.0	88	5
11.7	99	8
13.0	98	21
15.0	42	15
20.0	58	57
25.0	42	41



## IV. Сингулярний розклад матриць у регресійному аналізі

### Вказівки до розв'язання задач

Якщо певну задачу представити у матричному вигляді  $Y = XA$ , то мета регресійного аналізу – знайти вектор  $A$ , для чого необхідно розрахувати  $X^{-1}$ ,  $X$  – квадратна матриця (переконайтеся, що знаєте навіщо).

У деяких хімічних задачах, сформульованих у матричній формі (наприклад, методи розрахунку адитивних схем), може виявитись неможливість знаходження оберненої матриці внаслідок того, що вихідна матриця близька до виродженої. В таких ситуаціях використовують методи знаходження псевдообернених матриць, які можуть бути реалізовані за допомогою сингулярного розкладу матриці (Singular value decomposition, SVD).

Виконати сингулярний розклад матриці  $X$  – значить представити її у вигляді добутку трьох матриць:  $X = U\Sigma V^T$ , де  $U$  та  $V$  – ортонормовані матриці,  $\Sigma$  – матриця діагональної структури, яка містить у собі сингулярні числа.

Кінцевою метою є розрахунок вектора  $A = V(\Sigma^T\Sigma)^{-1}\Sigma^T U^T Y$ .

Сингулярний розклад матриці виконується в програмі MATCAD або SCILAB за допомогою функції `svd`.

### Завдання для аудиторного та самостійного розв'язання

1. Для матриць  $X_1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ ,  $X_2 = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ ,  $X_3 = \begin{pmatrix} 0 & 2 & 1 \\ 4 & 2 & 1 \end{pmatrix}$ :

- виконайте сингулярний розклад;
- перевірте ортонормованість отриманих матриць;
- відновіть матриці  $X_1$ ,  $X_2$  та  $X_3$  з результатів сингулярного розкладу.

2. Для матриць  $X_4 = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ ,  $X_5 = \begin{pmatrix} 3 & 1 & 2 & 0 \\ 3 & 12 & 1 & 1 \\ 2 & 1 & 31 & 3 \\ 4 & 2 & 0 & 45 \end{pmatrix}$ :

- знайдіть обернені матриці;
- доведіть, що знайшли саме обернені до даних матриці.

3. Для матриць  $X_6 = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{pmatrix}$ ,  $X_7 = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 4 \\ 8 & 2 & 0 \end{pmatrix}$ ,  $X_8 = \begin{pmatrix} 1 & 5 \\ 2 & 0.1 \\ 3 & 1 \\ 2 & 10 \\ 10 & 1 \end{pmatrix}$ :

- знайдіть їх ранг;
- спробуйте знайти обернені.

#### 4. Основи факторного аналізу: згортання спектрофотометричної інформації.

Знайдіть число поглинаючих частинок у водному розчині етилендіамін-тетраацетату міді (вимірювання проводились у 21 розчині при 9 довжинах хвиль, результати наведені в таблиці IV.1).

Вказівка. Закон Бугера–Ламберта–Бера у матричному вигляді записується як  $A = C \times E$ , розмірності  $A - 21 \times 9$ ,  $C - 21 \times S$ ,  $E - S \times 9$ , де  $S$  – число поглинаючих частинок.  $\text{Rank}(A) \leq \min\{\text{Rank}(C), \text{Rank}(E)\}$ , значить, ранг  $A$  не більше числа поглинаючих частинок. Таким чином, треба виконати сингулярний розклад  $A$  та знайти ранг як кількість сингулярних чисел, для яких  $q_i / q_1 > 0.01$ .

Таблиця IV.1

		Зміна довжин хвиль								
Зміна розчинів		0.273	0.46	0.285	0.265	0.119	0.03	0.4076	0.27389	0.24325
		0.256	0.46	0.294	0.271	0.124	0.029	0.4093	0.27744	0.25032
		0.247	0.459	0.296	0.275	0.127	0.03	0.4085	0.27925	0.25255
		0.235	0.461	0.315	0.293	0.147	0.039	0.4182	0.29585	0.27081
		0.224	0.461	0.32	0.298	0.155	0.039	0.419	0.29862	0.27531
		0.202	0.46	0.333	0.312	0.172	0.047	0.4221	0.30958	0.28854
		0.187	0.46	0.346	0.326	0.192	0.056	0.427	0.32175	0.30212
		0.162	0.46	0.365	0.349	0.207	0.069	0.4338	0.34036	0.32153
		0.14	0.455	0.38	0.364	0.23	0.08	0.4364	0.3534	0.33718
		0.112	0.45	0.402	0.387	0.26	0.096	0.4421	0.37354	0.36004
		0.087	0.432	0.4	0.39	0.271	0.107	0.4292	0.37441	0.36175
		0.07	0.414	0.4	0.392	0.288	0.119	0.4188	0.3768	0.36524
		0.045	0.376	0.39	0.386	0.31	0.142	0.393	0.37405	0.36322
		0.034	0.327	0.369	0.37	0.329	0.17	0.3591	0.36588	0.35315
		0.02	0.285	0.343	0.345	0.327	0.181	0.324	0.34602	0.33405
		0.014	0.25	0.32	0.327	0.332	0.195	0.2955	0.33262	0.31884
		0.015	0.225	0.302	0.312	0.336	0.202	0.2752	0.32123	0.30604
		0.012	0.179	0.275	0.285	0.337	0.225	0.24	0.30506	0.28855
		0.007	0.142	0.25	0.265	0.335	0.235	0.2097	0.28891	0.27055
		0.008	0.135	0.246	0.264	0.341	0.25	0.205	0.29168	0.27093
	0.01	0.129	0.244	0.256	0.345	0.244	0.2012	0.28546	0.26757	

#### 5. Застосування SVD для боротьби з мультиколінеарністю в задачах лінійного МНК.

У таблиці IV.2 містяться дані про залежність теплоти схватування цементу ( $Y$ ) від складу клінкеру,  $X_i$  – мольні долі компонентів клінкеру. Розрахуйте теоретичні значення теплоти схватування цементу (виходячи з лінійної залежності цієї величини від складу клінкеру), розрахуйте коваріаційну матрицю параметрів регресії, їхню корельованість та перевірте адекватність моделі за критерієм  $\chi^2$ .

Таблиця IV.2

$X_1$	$X_2$	$X_3$	$X_4$	$Y$
6	7	61	26	78.5
1	15	55	29	74.3
11	8	25	56	104.3
11	8	50	31	87.6
7	6	35	52	95.9
11	9	25	55	109.2
3	17	9	71	102.7
1	22	46	31	72.5
2	18	26	54	93.1
21	4	28	47	115.9
1	23	36	40	83.8
11	9	14	66	113.3
10	8	14	68	109.4

6. Розв'язання задачі лінійного МНК із застосуванням SVD.

У таблиці IV.3 наведено дані про розчинність  $[\text{Co}(\text{thio})_3]\text{Br}_3$  в залежності від концентрації бромід-іонів. Оцініть константи рівноваги утворення комплексів.

Вказівка. Припускається, що розчинення обумовлене набором реакцій  $\text{MA}_3\downarrow + i \text{A}^+ = \text{MA}^i_{(3+i)}$ ,  $i = -3, -2, -1, 0, 1$  із невідомими константами рівноваги  $K_i$ . Тоді розчинність  $S = K_1 \times [\text{A}]^{-3} + K_2 [\text{A}]^{-2} + K_3 \times [\text{A}]^{-1} + K_4 + K_5 \times [\text{A}]$  або у матричному вигляді  $S = \text{A} \times \text{K}$ , де  $S$  – вектор,  $\text{A}$  – матриця, стовпці якої сформовані з величин  $[\text{A}]^{-3}$ ,  $[\text{A}]^{-2}$ ,  $[\text{A}]^{-1}$ , 1 і  $[\text{A}]$ .

Таблиця IV.3

$[\text{Br}^-]$ , моль/л	$S$ , моль/л
0.004685	0.003576
0.004966	0.003327
0.006768	0.002477
0.010234	0.002094
0.027615	0.003128
0.06378	0.006535
0.010234	0.002094
0.13643	0.013716
0.227313	0.022775
0.363661	0.036394
0.681832	6.82E-02

### Контрольна робота № 3

#### Тема «Графічна перевірка гіпотез у візуальному тестовому аналізі»

Для наведених у таблиці даних перевірте відповідність залежності частот виявлення компонента від його концентрації функціям нормального, логнормального й експоненціального розподілів.

$c_i$	$n_i$	$N_i$
1.9	5	25
2.0	23	43
2.1	22	43
2.2	24	43
2.3	27	43
2.4	20	43
2.5	34	43
2.6	39	43

Параметри функцій розподілу:

Експоненціального		Нормального		Логнормального	
a =	$r^2 =$	$\mu =$	$r^2 =$	$\mu =$	$r^2 =$
b =		$\sigma =$		$\sigma =$	

Стандартне відхилення середнього  $s_{\bar{x}} =$  \_\_\_\_\_.

Функція розподілу	Значення межі виявлення (концентрація на рівні ймовірності виявлення компонента $P = 0.95$ )
Нормального	
Логнормального	
Експоненціального	

Залежність  $P(c)$  описується функцією \_\_\_\_\_ розподілу, тому що \_\_\_\_\_

Значення статистики  $\chi^2$  для перевірки відповідності експериментальних частот виявлення компонента функції експоненціального розподілу складає \_\_\_\_\_. Критичне значення статистики  $\chi^2$  при числі ступенів вільності \_\_\_\_\_ і рівні значущості 10% дорівнює \_\_\_\_\_. Відповідно, гіпотезу слід \_\_\_\_\_

Необхідні формули:

Функція нормального розподілу: 
$$P(c) = \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^c \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx.$$

Функція логнормального розподілу: 
$$P(c) = \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^c \exp\left[-\frac{(\ln(x/\mu))^2}{2\sigma^2}\right] dx.$$

Функція експоненціального розподілу: 
$$P(c) = 1 - \exp\left(\frac{-(c-a)}{b}\right).$$

## Екзаменаційний білет письмового екзамену

Відповіді наводьте з правильною кількістю значущих цифр!

Функція нормального розподілу $P(u) = \frac{1}{\sigma(2\pi)^{1/2}} \int_{-\infty}^u \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx$	Функція біноміального розподілу $p(X=i) = \frac{n!}{i!(n-i)!} p^i (1-p)^{(n-i)}$
Перевірка гіпотез про функцію розподілу $\chi_{\text{exp}}^2 = \sum_{i=1}^K \frac{(v_i - v_i^{\text{теор}})^2}{v_i^{\text{теор}}}$	Перевірка гіпотез про адекватність регресійної моделі $\chi_{\text{експ}}^2 = \sum_{i=1}^N \xi_i^2$

1. **14 балів.** У паралельних вимірюваннях знаходили масу монети.

№	маса, г	№	маса, г
1	3.025	6	3.111
2	3.024	7	3.022
3	3.028	8	3.029
4	3.027	9	3.021
5	3.028	10	3.022

1.1. Знайдіть середнє значення, медіану, стандартне відхилення, відносне стандартне відхилення та стандартне відхилення середнього значення. **(5 балів)**

1.2. Приймаючи, що результати вимірювання маси розподілені за законом Гауса, а середнє значення та вибіркова дисперсія є добрими наближеннями до параметрів генеральної сукупності – математичного сподівання  $\mu$  та дисперсії  $\sigma^2$ , розрахуйте частки результатів вимірювань, що дають значення маси в інтервалах: а) 3.022-3.025, б)  $>3.106$ . **(5 балів)**

1.3. З урахуванням відповіді на питання 1.2.б, поясніть, яку з величин (середнє значення чи медіану) доцільно використовувати як оцінку математичного сподівання  $\mu$ ? **(4 бали)**

2. **9 балів.** Молярну масу ідеального газу ( $M$ ) можна визначити за рівнянням Клапейрона–Менделєєва:

$$M = \frac{mRT}{PV}, \text{ де } m - \text{маса, г; } R - \text{універсальна газова стала } 0.082056$$

л·атм/(моль·К),  $P$  – тиск, атм,  $V$  – об'єм, л. У досліді одержали такі дані:  $m = 0.118$  г ( $s_m = 0.001$  г);  $T = 298.2$  К ( $s_T = 0.05$  К);  $P = 0.724$  атм ( $s_P = 0.005$  атм);  $V = 0.250$  л ( $s_V = 0.005$  л); всі виміряні величини є незалежними.

2.1. Розрахуйте молярну масу газу, її стандартне відхилення та відносне стандартне відхилення. **(6 балів)**

2.2. Яка з виміряних величин дає найбільший внесок у похибку (стандартне відхилення) молярної маси газу? Відповідь обґрунтуйте. **(3 бали)**

3. **10 балів.** При вимірюванні масової частки нітрогену в добриві одержали такі результати.

Масова частка, %	Кількість дослідів	Масова частка, %	Кількість дослідів
менше 12.00	1	12.07-12.08	4
12.00-12.02	3	12.08-12.09	3
12.02- 12.03	3	12.09-12.10	4
12.03- 12.04	5	12.10-12.11	2
12.04- 12.05	5	12.11-12.12	2
12.05-12.06	6	12.12-12.13	1

3.1. Побудуйте гістограму частот результатів аналізу. (3 бали)

3.2. За критерієм  $\chi^2$  перевірте гіпотезу про узгодження експериментальних даних з нормальним розподілом. (7 балів)

4. **10 балів.** В цьому завданні ви повинні оцінити адекватність регресійної моделі за локальними та глобальним критерієм адекватності. В кінетичних дослідженнях за залежністю швидкості реакції ( $V$ , моль·л<sup>-1</sup>·с<sup>-1</sup>) від концентрації реагентів методом найменших квадратів знайшли кінетичне рівняння. В таблиці наведено швидкості реакції, виміряні та розраховані за знайденим регресійним рівнянням (містить два підгоночні параметри). Відносне стандартне відхилення виміряних швидкостей реакції становить 2.0%.

№ досліду	$V_{\text{експеримент}}$	$V_{\text{розрахунок}}$	№ досліду	$V_{\text{експеримент}}$	$V_{\text{розрахунок}}$
1	3.03	2.98	5	4.84	4.70
2	4.04	4.11	6	5.12	5.14
3	4.56	4.60	7	5.36	5.39
4	4.73	4.69	8	5.43	5.37

4.1. Розрахуйте статистичні ваги  $w_k$  та локальні критерії адекватності – зважені залишки  $\xi_k = w_k^{1/2} \cdot (V_{\text{розрахунок}} - V_{\text{експеримент}})$ . (4 бали)

4.2. Побудуйте графік залежності  $\xi_k$  від  $V_{\text{розрахунок}}$  та зробіть висновок щодо адекватності регресійної моделі. (3 бали)

4.3. Перевірте гіпотезу про адекватність регресійної моделі на основі порівняння  $\chi_{\text{експ}}^2$  з відповідним критичним значенням (для довірчої ймовірності 5%). (3 бали)

5. **7 балів.** Карбон має два стабільні нукліди, <sup>12</sup>C та <sup>13</sup>C. Мольні частки цих нуклідів становлять, відповідно, 98.89% і 1.11%.

5.1. Визначте середнє значення та стандартне відхилення кількості атомів <sup>13</sup>C в молекулі холестерину C<sub>27</sub>H<sub>44</sub>O. (2 бали)

5.2. Якою є ймовірність зустріти в зразку холестерину молекулу, що не містить жодного атома <sup>13</sup>C? (5 балів)

## Додаток

Таблиця Д1

Основні функції, вбудовані в розглянуті розрахункові пакети, які використовуються для розв'язання задач посібника.

Опис функції	Запис функції
Значення $\pi$	PI()
Обчислення $e^x$	exp(x)
Знаходження логарифма	
- десяткового	log10(x)
- натурального	ln(x)
Обчислення $\sqrt{x}$	sqrt(x)
Обчислення $x^y$	x^y
Обчислення модуля $ a $	abs(a)
Функція нормального розподілу	NORMDIST
Функція, оборотна до функції нормального розподілу	NORMINV
Функція для розрахунку критичного значення статистики $\chi^2$	CHIINV
Функція для розрахунку t-критерію	TINV
Функція регресійного аналізу	LINEST
Значення коефіцієнта асиметрії	KURT
Значення коефіцієнта ексцесу	SKEW
Сингулярний розклад матриці X	[U,A,V]=SVD(X)

Таблиця Д2. Значення функції стандартного нормального розподілу

$u$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000



Квантілі розподілу  $\chi^2$ 

$f; P$	0.5	0.6	0.7	0.8	0.9	0.95	0.975	0.99
1	0.4549	0.7083	1.0742	1.6424	2.7055	3.8415	5.0239	6.6349
2	1.3863	1.8326	2.4079	3.2189	4.6052	5.9915	7.3778	9.2103
3	2.366	2.9462	3.6649	4.6416	6.2514	7.8147	9.3484	11.3449
4	3.3567	4.0446	4.8784	5.9886	7.7794	9.4877	11.1433	13.2767
5	4.3515	5.1319	6.0644	7.2893	9.2364	11.0705	12.8325	15.0863
6	5.3481	6.2108	7.2311	8.5581	10.6446	12.5916	14.4494	16.8119
7	6.3458	7.2832	8.3834	9.8032	12.017	14.0671	16.0128	18.4753
8	7.3441	8.3505	9.5245	11.0301	13.3616	15.5073	17.5345	20.0902
9	8.3428	9.4136	10.6564	12.2421	14.6837	16.919	19.0228	21.666
10	9.3418	10.4732	11.7807	13.442	15.9872	18.307	20.4832	23.2093
11	10.341	11.5298	12.8987	14.6314	17.275	19.6751	21.92	24.725
12	11.3403	12.5838	14.0111	15.812	18.5493	21.0261	23.3367	26.217
13	12.3398	13.6356	15.1187	16.9848	19.8119	22.362	24.7356	27.6882
14	13.3393	14.6853	16.2221	18.1508	21.0641	23.6848	26.1189	29.1412
15	14.3389	15.7332	17.3217	19.3107	22.3071	24.9958	27.4884	30.5779
16	15.3385	16.7795	18.4179	20.4651	23.5418	26.2962	28.8454	31.9999
17	16.3382	17.8244	19.511	21.6146	24.769	27.5871	30.191	33.4087
18	17.3379	18.8679	20.6014	22.7595	25.9894	28.8693	31.5264	34.8053
19	18.3377	19.9102	21.6891	23.9004	27.2036	30.1435	32.8523	36.1909
20	19.3374	20.9514	22.7745	25.0375	28.412	31.4104	34.1696	37.5662

Квантілі  $t$ -розподілу

$f; P$	0.5	0.6	0.7	0.8	0.9	0.95	0.98
1	1	1.3764	1.9626	3.0777	6.3138	12.7062	31.8205
2	0.8165	1.0607	1.3862	1.8856	2.92	4.3027	6.9646
3	0.7649	0.9785	1.2498	1.6377	2.3534	3.1824	4.5407
4	0.7407	0.941	1.1896	1.5332	2.1318	2.7764	3.7469
5	0.7267	0.9195	1.1558	1.4759	2.015	2.5706	3.3649
6	0.7176	0.9057	1.1342	1.4398	1.9432	2.4469	3.1427
7	0.7111	0.896	1.1192	1.4149	1.8946	2.3646	2.998
8	0.7064	0.8889	1.1081	1.3968	1.8595	2.306	2.8965
9	0.7027	0.8834	1.0997	1.383	1.8331	2.2622	2.8214
10	0.6998	0.8791	1.0931	1.3722	1.8125	2.2281	2.7638
11	0.6974	0.8755	1.0877	1.3634	1.7959	2.201	2.7181
12	0.6955	0.8726	1.0832	1.3562	1.7823	2.1788	2.681
13	0.6938	0.8702	1.0795	1.3502	1.7709	2.1604	2.6503
14	0.6924	0.8681	1.0763	1.345	1.7613	2.1448	2.6245
15	0.6912	0.8662	1.0735	1.3406	1.7531	2.1314	2.6025
16	0.6901	0.8647	1.0711	1.3368	1.7459	2.1199	2.5835
17	0.6892	0.8633	1.069	1.3334	1.7396	2.1098	2.5669
18	0.6884	0.862	1.0672	1.3304	1.7341	2.1009	2.5524
19	0.6876	0.861	1.0655	1.3277	1.7291	2.093	2.5395
20	0.687	0.86	1.064	1.3253	1.7247	2.086	2.528
21	0.6864	0.8591	1.0627	1.3232	1.7207	2.0796	2.5176
22	0.6858	0.8583	1.0614	1.3212	1.7171	2.0739	2.5083
23	0.6853	0.8575	1.0603	1.3195	1.7139	2.0687	2.4999
24	0.6848	0.8569	1.0593	1.3178	1.7109	2.0639	2.4922
25	0.6844	0.8562	1.0584	1.3163	1.7081	2.0595	2.4851
26	0.684	0.8557	1.0575	1.315	1.7056	2.0555	2.4786
27	0.6837	0.8551	1.0567	1.3137	1.7033	2.0518	2.4727
28	0.6834	0.8546	1.056	1.3125	1.7011	2.0484	2.4671
29	0.683	0.8542	1.0553	1.3114	1.6991	2.0452	2.462
30	0.6828	0.8538	1.0547	1.3104	1.6973	2.0423	2.4573
31	0.6825	0.8534	1.0541	1.3095	1.6955	2.0395	2.4528
32	0.6822	0.853	1.0535	1.3086	1.6939	2.0369	2.4487
33	0.682	0.8526	1.053	1.3077	1.6924	2.0345	2.4448
34	0.6818	0.8523	1.0525	1.307	1.6909	2.0322	2.4411
35	0.6816	0.852	1.052	1.3062	1.6896	2.0301	2.4377
36	0.6814	0.8517	1.0516	1.3055	1.6883	2.0281	2.4345

Таблиця Д4  
Продовження

$f; P$	0.5	0.6	0.7	0.8	0.9	0.95	0.98
37	0.6812	0.8514	1.0512	1.3049	1.6871	2.0262	2.4314
38	0.681	0.8512	1.0508	1.3042	1.686	2.0244	2.4286
39	0.6808	0.8509	1.0504	1.3036	1.6849	2.0227	2.4258
40	0.6807	0.8507	1.05	1.3031	1.6839	2.0211	2.4233
41	0.6805	0.8505	1.0497	1.3025	1.6829	2.0195	2.4208
42	0.6804	0.8503	1.0494	1.302	1.682	2.0181	2.4185
43	0.6802	0.8501	1.0491	1.3016	1.6811	2.0167	2.4163
44	0.6801	0.8499	1.0488	1.3011	1.6802	2.0154	2.4141
45	0.68	0.8497	1.0485	1.3006	1.6794	2.0141	2.4121
46	0.6799	0.8495	1.0483	1.3002	1.6787	2.0129	2.4102
47	0.6797	0.8493	1.048	1.2998	1.6779	2.0117	2.4083
48	0.6796	0.8492	1.0478	1.2994	1.6772	2.0106	2.4066
49	0.6795	0.849	1.0475	1.2991	1.6766	2.0096	2.4049
50	0.6794	0.8489	1.0473	1.2987	1.6759	2.0086	2.4033

Квантілі  $F$ -розподілу

$f; f$	1	2	3	4	5	6	7	8	9	10
2	18.51	19	19.16	19.25	19.3	19.33	19.36	19.37	19.38	19.39
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63
8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39	3.34
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13
10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97
11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9	2.86
12	4.75	3.88	3.49	3.26	3.11	3	2.92	2.85	2.8	2.76
13	4.67	3.8	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67
14	4.6	3.74	3.34	3.11	2.96	2.85	2.77	2.7	2.65	2.6
15	4.54	3.68	3.29	3.06	2.9	2.79	2.7	2.64	2.59	2.55
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.2	2.96	2.81	2.7	2.62	2.55	2.5	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.9	2.74	2.63	2.55	2.48	2.43	2.38
20	4.35	3.49	3.1	2.87	2.71	2.6	2.52	2.45	2.4	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.3	3.44	3.05	2.82	2.66	2.55	2.47	2.4	2.35	2.3
23	4.28	3.42	3.03	2.8	2.64	2.53	2.45	2.38	2.32	2.28
24	4.26	3.4	3.01	2.78	2.62	2.51	2.43	2.36	2.3	2.26
25	4.24	3.38	2.99	2.76	2.6	2.49	2.41	2.34	2.28	2.24
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22

Числові значення  $Q$ -критерію

$n$	$P$		
	0.9	0.95	0.99
3	0.89	0.94	0.99
4	0.68	0.77	0.89
5	0.56	0.64	0.76
6	0.48	0.56	0.70
7	0.43	0.51	0.64
8	0.40	0.48	0.58

## Рекомендована література

### Базова

1. Худсон Д. Статистика для физиков / Д. Худсон. – М. : Мир, 1970. – 295 с.
2. Тейлор Дж. Введение в теорию ошибок / Дж. Тейлор. – М. : Мир, 1985. – 272 с.
3. Уилкс С. Математическая статистика / С. Уилкс. – М. : Наука, 1967. – 632 с.
4. Налимов В. В. Применение математической статистики при анализе вещества / В. В. Налимов. – М. : Гос. изд-во физ.-мат. лит-ры, 1960. – 430 с.

### Допоміжна

1. Доерфель К. Статистика в аналитической химии / К. Доерфель. – М. : Мир, 1969. – 248 с.
2. Демиденко Е. З. Линейная и нелинейная регрессия / Е. З. Демиденко. – М. : Финансы и статистика, 1981.
3. Родионова О. Е. Хемометрика: достижения и перспективы / О. Е. Родионова, А. Л. Померанцев // Успехи химии. – 2006. – Т. 75, № 4. – С. 302–321.
4. Большев Л. Н. Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – М. : Наука, 1983. – 416 с.
5. Ван дер Варден Б. Л. Математическая статистика / Б. Л. Ван дер Варден. – М. : Изд-во иностр. л-ры, 1960. – 434 с.

**Для нотаток**

Навчальне видання

**Пантелеймонов** Антон Віталійович  
**Христенко** Інна Василівна  
**Іванов** Володимир Венедиктович  
**Холін** Юрій Валентинович

**Статистичні та хемометричні  
методи в хімії**

Навчальний посібник

Коректор *О. В. Гавриленко*  
Комп'ютерна верстка *В. В. Савінкова*  
Макет обкладинки *І. М. Дончик*

Формат 60×84/16. Ум. друк. арк. 2,48. Наклад 100 пр. Зам. № 74/12.

Видавець і виготовлювач  
Харківський національний університет імені В. Н. Каразіна,  
61022, м. Харків, майд. Свободи, 4  
Свідоцтво суб'єкта видавничої справи ДК № 3367 від 13.01.2009

Видавництво ХНУ імені В. Н. Каразіна  
Тел. 705-24-32