

Н.О. Лисенко

**Методичні вказівки щодо виконання лабораторних робіт
 з курсу «Статистичний аналіз та моделювання вимірів в
 автоматизованих системах та комп'ютерно-інтегрованих
 технологіях»**

$$W^*(x | |x|) = \frac{\alpha}{n\sqrt{2\pi}} \sum_{k=1}^n \exp\left[-\frac{\alpha^2}{2}(x - x_k)^2\right]$$

$$A = \frac{n}{n - \sum_{k=1}^n (1 - \Phi(\alpha x_k))}$$

$$\alpha_{1a}^* = \frac{1.592 + 0.018n}{\sqrt{D^*}}$$

$$A = \frac{n}{n - \sum_{k=1}^n \Phi(\alpha(a - x_k)) - \sum_{k=1}^n (1 - \Phi(\alpha(b - x_k)))}$$

$$\alpha_{2a}^* = \frac{3.93 + 0.071n}{\sqrt{D^*}}$$

м. Дніпро
 2022 р.

Міністерство освіти і науки України
Дніпровський національний університет імені Олеся Гончара
Фізико-технічний факультет

**Методичні вказівки щодо виконання лабораторних робіт
з курсу «Статистичний аналіз та моделювання вимірів в
автоматизованих системах та комп'ютерно-інтегрованих
технологіях»**

Укладач: к.т.н., доц. Лисенко Н.О.

Дніпро
2022 р.

Введення

Методичний посібник розроблений з метою допомогти студентам оволодіти навичками обробки вимірювань та отримати досвід роботи з експериментальними даними. Зміст та назви наведених в посібнику лабораторних робіт в повному обсязі відповідають змісту дисципліни “Статистичний аналіз та моделювання вимірів в автоматизованих системах та комп’ютерно-інтегрованих технологіях” для студентів, що навчаються за спеціальністю 151 «Автоматизація та комп’ютерно-інтегровані технології».

За результатами виконання кожної лабораторної роботи потрібно сформулювати звіт, в якому необхідно навести методику, результати дослідження, таблиці, графіки та обов’язково аналіз і висновки. Звіт по лабораторній роботі є індивідуальним та рукописним.

Для виконання лабораторних робіт рекомендовано використовувати математичне середовище MathCad. Передбачається, що студенти вже ознайомлені з основними правилами роботи в цьому програмному пакеті із курсу “Основи теорії ймовірності і математичної статистики”. В кожному розділі методичного посібника коротко приведений набір функцій MathCad, які, на погляд автора, будуть у нагоді для успішної реалізації даної лабораторної роботи.

До методичного посібника додається диск з електронними даними. А саме, диск містить всі необхідні вибірки випадкових величин в форматі .prn, які студент отримує в якості емпіричних даних для виконання наведених в цьому посібнику лабораторних робіт.

Лабораторна робота №1

«Оцінка параметрів амплітудно – частотних характеристик лінійних фільтрів»

1. Теоретична частина

Розглядаються лінійні фільтри, спроектовані як послідовність ланок першого й другого порядку з передавальними функціями

$$H_1(p) = \frac{b}{p+a}, \quad H_2(p) = \frac{b}{p^2 + a_1 p + a_2}.$$

Квадрати їх амплітудно-частотних характеристик (АЧХ) дорівнюють:

$$|K_1(j\omega)|^2 = \frac{b^2}{\omega^2 + a^2}, \quad |K_2(j\omega)|^2 = \frac{b^2}{(a_2 - \omega^2)^2 + a_1^2 \omega^2}.$$

Припустимо, що вимірюються величини, зворотні квадратам амплітудно-частотних характеристик

$$y_k = |K_1(j\omega_k)|^{-2}.$$

Формули для оцінювання параметрів a і b можна отримати за допомогою методу найменших квадратів. За цим методом необхідно мінімізувати функцію

$$\xi_k^2(a, b) = \frac{1}{n} \sum_{k=1}^n \left(y_k - \frac{\omega_k^2 + a^2}{b^2} \right)^2.$$

Для того, щоб знайти значення невідомих параметрів a і b при яких функція $\xi_k^2(a, b)$ дістається мінімуму необхідно обчислити часткові похідні по a і b та вирішити систему рівнянь

$$\begin{cases} \frac{\partial \xi_k^2(a, b)}{\partial a} = 0 \\ \frac{\partial \xi_k^2(a, b)}{\partial b} = 0 \end{cases},$$

Отже система рівнянь для оцінювання параметрів a і b за методом найменших квадратів запишеться у вигляді:

$$\begin{cases} b^2 \sum_{k=1}^n y_k - na^2 = \sum_{k=1}^n \omega_k^2 \\ b^2 \sum_{k=1}^n y_k^2 - a^2 \sum_{k=1}^n y_k = \sum_{k=1}^n \omega_k^2 y_k \end{cases}.$$

Передавальну функцію фільтра другого порядку представимо через коефіцієнт загасання δ і резонансну частоту ω_0 :

$$H(p) = \frac{b}{p^2 + 2\delta\omega_0 p + \omega_0^2}.$$

Тоді квадрат АЧХ запишеться у вигляді

$$|K_2(j\omega)|^2 = \frac{b^2}{(\omega_0^2 - \omega^2)^2 + 4\delta^2 \omega_0^2 \omega^2}.$$

Величини, що вимірюються, дорівнюють квадратам амплітудно-частотних характеристик

$$x_k = |K_2(j\omega_k)|^2.$$

Припустимо, що амплітудно-частотна характеристика має максимум ($\delta^2 \leq 0,5$) і можна за графіком оцінити x_m й ω_m . Тоді з умови максимуму витікає формула

$$\omega_m^2 = \omega_0^2 (1 - 2\delta^2). \quad (1.1)$$

Максимум АЧХ у точці ω_m дорівнює

$$K_{\max}^2 = \frac{b^2}{\omega_0^4 - \omega_m^4}. \quad (1.2)$$

Замінивши K_{\max}^2 на оцінку x_m для вимірювань АЧХ одержимо формулу

$$x_k = \frac{x_m (\omega_0^4 - \omega_m^4)}{(\omega_0^2 - \omega_k^2)^2 + 2\omega_k^2 (\omega_0^2 - \omega_m^2)}. \quad (1.3)$$

Якщо ввести наступне позначення

$$A = \omega_0^4 - \omega_m^4,$$

то формулу (1.3) можна трансформувати до вигляду

$$x_k = \frac{Ax_m}{A + (\omega_m^2 - \omega_k^2)^2}.$$

Середній квадрат помилки визначається як

$$\Delta^2(A) = \frac{1}{n} \sum_{k=1}^n \left(x_k - \frac{Ax_m}{A + (\omega_m^2 - \omega_k^2)^2} \right)^2. \quad (1.4)$$

Якщо обчислити похідну $\frac{d\Delta^2(A)}{dA}$ та дорівняти її нулю, можна отримати значення A^* при якому середній квадрат помилки (1.4) буде мінімальним

$$A^* = \frac{\sum_{k=1}^n x_k (\omega_m^2 - \omega_k^2)^4}{\sum_{k=1}^n (x_m - x_k) (\omega_m^2 - \omega_k^2)^2}.$$

Оцінивши A , можна визначити оцінку резонансної частоти ω_0^* за формулою

$$\omega_0^* = (A^* + \omega_m^4)^{\frac{1}{4}}.$$

Із формул (1.1) та (1.2) отримуються оцінки δ^* і b^* , відповідно,

$$\delta^* = \sqrt{0.5 \left[1 - \left(\frac{\omega_m}{\omega_0^*} \right)^2 \right]}, \quad b^* = \sqrt{x_m \left[(\omega_0^*)^4 - \omega_m^4 \right]}.$$

2. Завдання на дослідження

Частина 1. Оцінювання параметрів лінійного фільтру першого порядку.

Дано: вибірка випадкових величин $|y| = y_1, y_2, \dots, y_n$, які зворотні квадрату амплітудно – частотної характеристики фільтру першого порядку та вибірка відповідних $|y|$ частот $|\omega|$.

Визначити: параметри a і b амплітудно – частотної характеристики фільтру першого порядку.

Частина 2. Оцінювання параметрів лінійного фільтру другого порядку.

Дано: вибірка випадкових величин $|x| = x_1, x_2, \dots, x_n$, які дорівнюють квадрату амплітудно – частотної характеристики фільтру другого порядку та вибірка відповідних $|x|$ частот $|\omega|$.

Визначити: оцінити параметри ω_0 , b , δ амплітудно – частотної характеристики фільтру другого порядку.

3. Порядок виконання роботи

Частина 1. Оцінювання параметрів лінійного фільтру першого порядку.

- 1) Отримати вибірки випадкових величин $|y|$ та частот $|\omega|$ (на диску у викладача) відповідно до заданого варіанту.
- 2) Визначити довжину вибірок.
- 3) Вивести формули для визначення параметрів амплітудно – частотної характеристики лінійного фільтру першого порядку за методом найменших квадратів.
- 4) Обчислити значення невідомих параметрів a і b .
- 5) Побудувати графік квадрату амплітудно – частотної характеристики фільтру першого порядку $|K_1(j\omega)|^2$.

Частина 2. Оцінювання параметрів лінійного фільтру другого порядку.

- 1) Отримати вибірки випадкових величин $|x|$ та частот $|\omega|$ (на диску у викладача) відповідно до заданого варіанту.
- 2) Визначити довжину вибірок.
- 3) Вивести формули для визначення параметрів амплітудно – частотної характеристики лінійного фільтру другого порядку.
- 4) Отримати оцінки невідомих параметрів ω_0 , δ , b .
- 5) Побудувати графік квадрату амплітудно – частотної характеристики фільтру другого порядку $|K_2(j\omega)|^2$.

4. Застосування математичного середовища MathCad для розв'язання задач даного типу

Вибірki випадкових величин на диску наведені в форматі .prn. Для того, щоб побачити дані наочно та для їх подальшого використання необхідно

звернутися до заданого в залежності від варіанту файлу за допомогою функції MathCad: READPRN(“ ”). Наприклад, на диску D в папці “data” знаходиться потрібний вам файл B1_1.prn. Якщо записати в MathCad

$$z := \text{READPRN}("d:\text{data}\text{B1_1.prn}"),$$

то зміст файлу B1_1.prn запишеться в змінну z .

Для виконання лабораторної роботи №1 знадобиться знання деяких вбудованих в MathCad функцій. А саме:

- функція $\text{length}(z)$ – функція, що повертає кількість елементів у векторі z , іншими словами, це функція, яка визначає довжину вибірки z ;
- функція $\text{max}(x)$ – функція, що знаходить максимальне значення у вибірці x . Наприклад, якщо записати

$$x_{\text{max}} = \text{max}(x),$$

то в змінну x_{max} запишеться максимальне значення вибірки x .

- функція $\text{min}(x)$ – функція, що знаходить мінімальне значення у вибірці x .
- функція $\text{match}(d, x)$ – функція, що знаходить у векторі x задане значення d , та повертає індекс його позиції. Наприклад, щоб визначити позицію максимального значення x_{max} у векторі x , необхідно записати

	0
0	0.533
1	0.667
2	2.434
3	1.887
4	1.364
5	1.783
6	1.435
7	0.792
8	1.372
9	1.509

$x_{\text{max}} = \text{max}(x)$
 $x_{\text{max}} = 2.434$
 $k := \text{match}(x_{\text{max}}, x)$
 $k = (2)$

$x =$

	0
0	0.533
1	0.667
2	2.434
3	1.887
4	1.364
5	1.783
6	1.435
7	0.792
8	1.372
9	1.509

Функція $\text{match}(d, x)$ знаходить всі значення d , у випадку коли заданих значень у векторі декілька:

	0
0	0.47
1	0.32
2	0.31
3	0.36
4	0.58
5	0.45
6	0.23
7	0.58
8	0.41
9	0.58

$x_{\text{max}} = \text{max}(x)$
 $x_{\text{max}} = 0.58$
 $k := \text{match}(x_{\text{max}}, x)$
 $k = \begin{pmatrix} 4 \\ 7 \\ 9 \end{pmatrix}$

$x =$

	0
0	0.47
1	0.32
2	0.31
3	0.36
4	0.58
5	0.45
6	0.23
7	0.58
8	0.41
9	0.58

Лабораторна робота №2

«Оцінка параметрів законів розподілу ймовірностей і аналіз результатів обчислювальних експериментів»

1. Теоретична частина

При оцінюванні параметрів законів розподілу прагнуть одержати незміщені, спроможні та ефективні оцінки. Але не завжди така вимога виконується. Багато залежить від методу, який використовується для оцінювання невідомого параметру. На практиці широке застосування набули такі методи як мінімуму математичного очікування квадрату помилки, метод максимальної правдоподібності та метод моментів. Як свідчить досвід, найбільш ефективним є метод максимальної правдоподібності, але найпростішим – метод моментів. Нижче стисло приведено основний зміст кожного з перелічених методів оцінювання.

Метод мінімуму математичного очікування квадрату помилки.

Передбачається, що задана вибірка вимірювань $|x| = x_1, x_2, \dots, x_n$ з відомим законом розподілу ймовірностей $W(x/\alpha)$, але невідомим параметром α . Закон розподілу параметра $W(\alpha)$ є відомим. Тоді оптимальна точкова оцінка невідомого параметра закону розподілу α за критерієм мінімуму математичного очікування квадрату помилки $M[(\alpha - \alpha_n^*)^2]$ має вигляд

$$\alpha_n^* = \frac{\int_0^{\infty} \alpha W(\alpha) W\left(\frac{|x|}{\alpha}\right) d\alpha}{\int_0^{\infty} W(\alpha) W\left(\frac{|x|}{\alpha}\right) d\alpha}, \quad (2.1)$$

де $W(|x|/\alpha) = \prod_{i=1}^n W(x_i/\alpha)$ – функція правдоподібності.

Метод максимальної правдоподібності.

Запишемо логарифм функції правдоподібності

$$L\left(\frac{|x|}{\alpha}\right) = \sum_{i=1}^n \ln\left(W\left(\frac{x_i}{\alpha}\right)\right) = \ln\left(W\left(\frac{|x|}{\alpha}\right)\right).$$

Тоді алгоритм оцінювання параметра α за методом максимальної правдоподібності виглядає як рішення рівняння

$$\frac{dL\left(\frac{|x|}{\alpha}\right)}{d\alpha} = 0. \quad (2.2)$$

У більшості практичних випадків це рівняння є складним для розв'язування. Розв'язати рівняння можна чисельними методами.

Метод моментів.

Цей метод базується на властивості рівності теоретичних та статистичних початкових або центральних моментів. Можлива їх комбінація.

Наприклад, параметр закону розподілу Релея b невідомий. Але для розподілу Релея справедливі такі аналітичні визначення моментів:

$$M = \sqrt{\frac{\pi b}{2}}, \quad D = b \left(\frac{4 - \pi}{2} \right),$$

де M - математичне очікування (перший початковий момент), D - дисперсія (другий центральний момент). Оцінку параметра b визначають із рівнянь

$$\sqrt{\frac{\pi b}{2}} = \bar{M} \quad \text{або} \quad b \left(\frac{4 - \pi}{2} \right) = \bar{D},$$

де $\bar{M} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ і $\bar{D} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$ - вибіркове середнє та вибіркова дисперсія відповідно.

Наведені вище методи відносяться до методів точкового оцінювання, коли оцінка параметра отримується у вигляді одного значення. Але існує також інтервальне оцінювання – це визначення двох значень α_1^* і α_2^* у межах яких із заданою ймовірністю P буде знаходитися невідомий параметр α , тобто оцінка отримується у вигляді:

$$\alpha_1^* \leq \alpha \leq \alpha_2^*.$$

Застосування інтервального оцінювання продемонструємо на прикладі нормального закону розподілу. Розглянемо три варіанти умов:

1) Оцінюється математичне очікування, дисперсія відома й дорівнює σ^2 , тоді оцінка має вигляд

$$a_n^* - \frac{\sigma}{\sqrt{n}} \Psi \left(\frac{1+P}{2} \right) \leq a \leq a_n^* + \frac{\sigma}{\sqrt{n}} \Psi \left(\frac{1+P}{2} \right), \quad (2.3)$$

де $\Psi(\beta)$ - функція, зворотна інтегралу ймовірностей Гауса, n - розмір вибірки вимірювань, $a_n^* = \bar{x}$ - вибіркове середнє.

2) Оцінюється математичне очікування, дисперсія вимірювань невідома, тоді інтервальна оцінка буде у вигляді

$$a_n^* - \sqrt{\frac{\bar{D}}{n}} \Psi_c \left(\frac{1+P}{2} \right) \leq a \leq a_n^* + \sqrt{\frac{\bar{D}}{n}} \Psi_c \left(\frac{1+P}{2} \right), \quad (2.4)$$

де \bar{D} - вибіркова дисперсія, $\Psi_c(\beta)$ - функція, зворотна функції розподілу Стюдента.

3) Оцінюється дисперсія, математичне очікування відоме й дорівнює a . Інтервальне оцінювання дисперсії нормального розподілу має вигляд

$$\frac{n\bar{D}}{z_2} \leq D \leq \frac{n\bar{D}}{z_1},$$

де $z_1 = \Psi_x\left(\frac{1-P}{2}\right)$, $z_2 = \Psi_x\left(\frac{1+P}{2}\right)$, $\Psi_x(\beta)$ - функція, зворотна функції розподілу хі-квадрат. При достатній довжині вибірки вимірювань ($n > 30$) значення z_1 і z_2 можна обчислювати за наближеними формулами

$$z_1 = n \left(1 - \frac{2}{9n} + \Psi\left(\frac{1-P}{2}\right) \sqrt{\frac{2}{9n}} \right)^3,$$

$$z_2 = n \left(1 - \frac{2}{9n} + \Psi\left(\frac{1+P}{2}\right) \sqrt{\frac{2}{9n}} \right)^3,$$

де $\Psi(\beta)$ - функція, зворотна інтегралу ймовірностей Гауса.

2. Завдання на дослідження

Частина 1. Точкове оцінювання.

Дано: вибірка випадкових величин і вид її закону розподілу. Закон розподілу невідомого параметра $W(\alpha)$ вважати рівномірним.

Визначити: точкові оцінки невідомого параметра закону розподілу α за допомогою методів мінімуму математичного очікування квадрату помилки, максимальної правдоподібності та методу моментів.

Частина 2. Інтервальне оцінювання.

Дано: три вибірки випадкових величин з нормальним законом розподілу ймовірностей з однаковими параметрами, але з різною довжиною ($n = 25, 75, 125$). Задана також ймовірність знаходження невідомого параметра в інтервалі P .

Визначити: інтервальну оцінку математичного очікування при умові, що дисперсія відома, інтервальну оцінку математичного очікування при умові, що дисперсія невідома, й інтервальні оцінки дисперсії за точними та наближеними формулами. Зробити аналіз результатів дослідження.

3. Порядок виконання роботи

Частина 1. Точкове оцінювання.

- 1) Відповідно до варіанту отримати вибірку випадкових величин (див. на диску у викладача) і вид її закону розподілу (див. табл. 5.1 в додатку).
- 2) Визначити довжину вибірки, задати закон розподілу невідомого параметра $W(\alpha)$.
- 3) Записати функцію правдоподібності $W(|x|/\alpha)$.
- 4) Обчислити оцінку невідомого параметра за (2.1) методом мінімуму математичного очікування квадрату помилки, .
- 5) Записати логарифм функції правдоподібності та, шляхом вирішення рівняння (2.2), отримати оцінку невідомого параметра закону розподілу ймовірностей за методом максимальної правдоподібності.

6) Обчислити вибіркове середнє та вибіркєву дисперсїю. На основї отриманих оцїнок першого початкового та другого центрального моментів обчислити оцїнку невідомого параметра закону розподїлу.

7) Результати представити у виглядї таблицї 2.1.

Таблиця 2.1

n	Оцїнка параметра за методом $\min M \left[(\alpha - \alpha_n^*)^2 \right]$	Оцїнка параметра за методом $\max W(x /\alpha)$	Оцїнка за методом моментів	
			на основї першого початкового моменту	на основї другого центрального моменту

8) Зробити аналіз отриманих оцїнок і порівняти між собою використанї методи оцїнювання.

Частина 2. Інтервальне оцїнювання.

1) Самостїйно згенерувати три вибірки випадкових величин з нормальним законом розподїлу ймовїрностей, якї відрїзняються одна від одної довжиною ($n = 25, 75, 125$). Відповїдно до таблицї варїантів (див. таблицю 5.2) задати параметри нормального закону a і σ^2 та ймовїрність P .

2) Провести інтервальне оцїнювання математичного очїкування, якщо дисперсїя відома. Отримати границї інтервалу a_1^* та a_2^* згїдно з формулою (2.3).

3) Провести інтервальне оцїнювання математичного очїкування, якщо дисперсїя невідома. Отримати границї інтервалу a_1^* та a_2^* згїдно з формулою (2.4).

4) Оцїнити дисперсїю нормального закону розподїлу за точними та наближеними формулами, отримати граничні значення інтервалу D_1^* та D_2^* для обох випадків.

5) Результати дослідження представити у виглядї таблицї 2.2.

Таблиця 2.2

$P = \dots$	Оцїнка мат. очїкування, якщо дисперсїя відома			Оцїнка мат. очїкування, якщо дисперсїя невідома			Оцїнка дисперсїї за точними формулами			Оцїнка дисперсїї за наближеними формулами		
	a_1^*	a	a_2^*	a_1^*	a	a_2^*	D_1^*	D	D_2^*	D_1^*	D	D_2^*
n												
25												
75												
125												

б) Зробити аналіз даних, отриманих в ході виконання лабораторної роботи . Сформулювати висновки.

4. Застосування математичного середовища MathCad для розв'язання задач даного типу

Для генерування вибірок випадкових величин MathCad містить широкий спектр спеціальних функцій. Наприклад, щоб створити вибірку випадкових величин з рівномірним законом розподілу на інтервалі від a до b використовується функція $runif(n,a,b)$, де n - довжина вибірки. Якщо згенерувати вибірку рівномірних випадкових величин на інтервалі від 0 до 1, то шляхом її функціонального перетворення можна отримати вибірку вимірювань з будь-яким законом розподілу. Наприклад, щоб сформулювати вибірку випадкових величин з розподілом Релея потрібно зробити наступне перетворення:

$n := 100$ - довжина вибірки

$r := runif(n,0,1)$ - вибірка з рівномірним розподілом на інтервалі від 0 до 1

$b := 2$ - параметр закону розподілу Релея

$x := \sqrt{2 \cdot b \cdot \ln\left(\frac{1}{r}\right)}$ - вибірка з релеєвським розподілом і параметром b

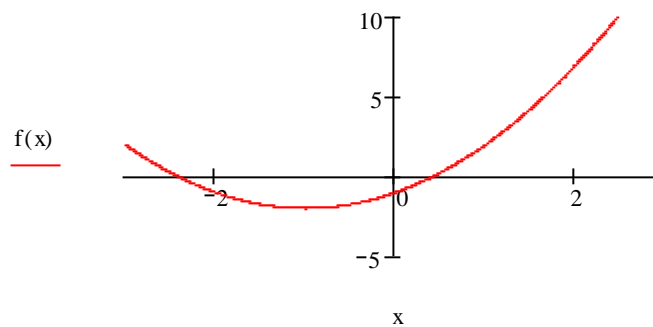
Цей спосіб універсальний, але він не самий легкий. Для деяких, найбільш розповсюджених законів, в середовищі MathCad вбудовано функції генерування випадкових величин з заданим законом розподілу. Так вибірку вимірювань з нормальним законом розподілу можна отримати завдяки використанню функції $norm(n,a,\sigma)$, де a і σ - параметри нормального закону розподілу, а саме математичне очікування та середнє квадратичне відхилення, відповідно.

Автоматично середовище MathCad нумерує масиви даних з нуля, але користувач може це змінити за допомогою функції ORIGIN. У кнопці меню Math знайдіть Options/Built-In Variables/Array Origin та вкажіть число з якого потрібно починати нумерацію масиву. Очевидно, що це може бути будь-яке позитивне чи негативне ціле число. Після цієї установки з вказаного вами числа буде починатися нумерація будь-якого масиву в даному робочому документі. Також можливо змінити нумерацію локально, вказавши прямо в робочому полі значення функції ORIGIN, наприклад, ORIGIN:=5. Тоді з п'яти буде починатися нумерація масивів, які в документі розташовані нижче цього запису. Наприклад:

$d := \text{rnorm}(10, 0, 1)$		0
	0	-0.684
	1	-1.615
	2	1.411
	3	0.972
d =	4	-1.156
	5	0.144
	6	1.697
	7	-0.32
	8	-2.031
	9	-1.411
ORIGIN:= 5		5
	5	-0.684
	6	-1.615
	7	1.411
	8	0.972
d =	9	-1.156
	10	0.144
	11	1.697
	12	-0.32
	13	-2.031
	14	-1.411

Функція $\text{root}(f(x), x, a, b)$ повертає значення змінної x , яке знаходиться в інтервалі від a до b , в якому функція $f(x)$ дорівнює нулю. Нижче наведено приклад застосування функції $\text{root}(f(x), x, a, b)$ для знаходження коренів рівняння:

$$f(x) := x^2 + 2 \cdot x - 1$$



$$k1 := \text{root}(f(x), x, 0, 2)$$

$$k2 := \text{root}(f(x), x, -3, -2)$$

$$k1 = 0.414$$

$$k2 = -2.414$$

Для виконання лабораторної роботи №2 також знадобиться знання вбудованих зворотних функцій:

- $qnorm(p, 0, 1)$ - функція, зворотна інтегралу ймовірностей Гауса;
- $qt(p, n-1)$ - функція, зворотна функції розподілу Стюдента з $n-1$ ступенями свободи;
- $qchisq(p, n-1)$ - функція, зворотна функції розподілу хі-квадрат, де $n-1$ - ступені свободи.

Лабораторна робота №3

«Перевірка гіпотез про закони й функції розподілу ймовірностей»

1. Теоретична частина

Перевірка гіпотези про закон розподілу ймовірності.

Експериментальні дані у вигляді вибірок вимірювань випадкових величин містять інформацію про їх статистичні закономірності. Із аналізу фізичної природи випадковості вимірювань або із їх гістограм можна отримати інформацію про вид закону розподілу ймовірностей. Якщо вибірки вимірювань довгі ($n \geq 100$), то по виду огинаючої гістограми можна висунути гіпотезу про закон розподілу ймовірності випадкових величин. Для перевірки гіпотези проводять порівняння гіпотетичного закону з гістограмою за критерієм хі-квадрат.

Припустимо, в результаті контролю отримана вибірка експериментальних даних $|x| = |x_1, x_2, \dots, x_n|$ з невідомим законом розподілу. Необхідно висунути та перевірити гіпотезу про закон розподілу ймовірності випадкових величин. Щоб визначити гістограмні числа потрібно обчислити мінімальне x_{\min} та максимальне x_{\max} значення вибірки вимірювань, задати кількість інтервалів L . Тоді інтервал дискретизації обчислюється за формулою

$$\Delta x = \frac{x_{\max} - x_{\min}}{L}.$$

Гістограмні числа k_1, k_2, \dots, k_L - це кількість вимірювань досліджуваної вибірки, які влучили в інтервали дискретизації $x_{0i} \pm \Delta x/2$, где $x_{0i} = x_{01} + (i-1)\Delta x$, $x_{01} = x_{\min} + \Delta x/2$. Якщо використати функцію Хевісайда (одиничного стрибка) для визначення гістограмних чисел, то алгоритм їх визначення буде мати вигляд

$$k_i = \sum_{k=1}^n [\text{sgn}(x_k - x_{0i} + \Delta x/2) - \text{sgn}(x_k - x_{0i} - \Delta x/2)],$$

де $\text{sgn}(x)$ - функція одиничного стрибка приймає лише два значення: $\text{sgn}(x) = 0$, якщо $x < 0$, та $\text{sgn}(x) = 1$, якщо $x \geq 0$

Якщо закон розподілу вимірювань $W(x)$, то ймовірності влучення вимірювань в i -тий інтервал гістограми p_i визначаються за формулою

$$p_i = \int_{x_{0i} - \frac{\Delta x}{2}}^{x_{0i} + \frac{\Delta x}{2}} W(x) dx \approx W(x_{0i}) \Delta x, \quad i = 1, 2, 3, \dots, L,$$

де x_{0i} - центр i -го інтервалу гістограми.

Очевидно, що відносні гістограмні числа $p_i^* = \frac{k_i}{n}$ - є оцінками ймовірностей p_i . Якщо ймовірності p_i визначені за гіпотетичним законом розподілу $W(x)$, то зважена сума квадратів різниці p_i и p_i^* може бути показником близькості $W(x)$ та невідомого закону досліджуваної вибірки вимірювань

$$z = \sum_{i=1}^L \left(\frac{p_i^* - p_i}{\sqrt{D_i}} \right)^2 = \sum_{i=1}^L \frac{(k_i - np_i)^2}{np_i(1-p_i)}.$$

Рішення про вірність висунутої гіпотези слід приймати шляхом порівняння обчисленого значення показника z з деяким граничним значенням z_0 : якщо $z \leq z_0$, то слід вважати, що гіпотеза вірна, якщо $z > z_0$, то гіпотезу слід відхилити та запропонувати іншу. Поріг порівняння z_0 визначають за правилом Неймана – Пірсона із умови: ймовірність помилки 1-го роду не повинна перевищувати деяке граничне значення P_0 (наприклад, $P(H/B) \leq 0.05$). Ймовірність $P = 1 - P(H/B)$ характеризує надійність правила Неймана – Пірсона. Для показника z поріг z_0 визначається із рівняння

$$\frac{1}{2^{\frac{L}{2}} \Gamma(L/2)} \int_0^{z_0} z^{\frac{L-2}{2}} \exp(-z/2) dz = P.$$

Якщо задано P і L , то поріг z_0 знаходять за таблицями математичної статистики. Приблизне значення z_0 можна оцінити за формулою

$$z_0 = L \left(1 - \frac{2}{9L} + \Psi(P) \sqrt{\frac{2}{9L}} \right)^3, \quad (3.1)$$

где $\Psi(x)$ - функція, зворотна інтегралу ймовірності Гауса.

На практиці дуже часто параметри гіпотетичних законів розподілу невідомі та оцінюються за вибіркою вимірювань. Параметри гіпотетичних законів замінюються їх оцінками, які самі є випадковими величинами. В теорії ймовірностей доведено, що при даній заміні показник z при великих n має розподіл хі-квадрат з $(L-r)$ ступенями свободи, де r - число параметрів, що замінені оцінками. Таким чином, якщо використовуються оцінки параметрів законів розподілу, то при визначенні граничного значення за формулою (3.1) необхідно L замінити на $(L-r)$.

Якщо висувається декілька гіпотез про закон розподілу вимірювань, то для вибору однієї з них рекомендовано робити наступним чином:

- 1) для кожної гіпотези обчислюються показники близькості z_1, z_2, \dots, z_m ;
- 2) серед них залишають лише ті, які задовольняють нерівності $z_i \leq z_0$, $i = 1, 2, \dots, m$;
- 3) із залишених обирають ту гіпотезу, для якої z_i мінімально.

Перевірка гіпотези про функцію розподілу ймовірності вимірювань.

При невеликій кількості вимірювань значно погіршується інформативність гістограми і тому зростає кількість помилкових рішень при перевірці гіпотези про закон розподілу випадкових величин. В даному випадку кращі результати можна одержати, якщо перевіряти гіпотезу про функцію розподілу.

В якості показника близькості емпіричної та гіпотетичної функцій розподілу ймовірностей використовується математичне очікування інтегрального квадрату різниці цих функцій

$$V = n \int_{-\infty}^{\infty} \left(F_n^* \left(\frac{x}{|x|} \right) - F(x) \right)^2 W(x) dx. \quad (3.2)$$

Цей критерій називається критерієм $n\omega$ - квадрат, або критерієм Крамера – Мізеса – Смирнова. Якщо гіпотеза про функцію розподілу є вірною, то показник V є випадковою величиною з математичним очікуванням

$$M[V] = \frac{1}{6}$$

та дисперсією

$$D[V] = \frac{(4n-3)}{180n}.$$

Ймовірності того, що $V \leq V_0$, визначені та наведені в таблиці 3.1.

Таблиця 3.1

V_0	0,184	0,241	0,367	0,461	0,549	0,62	0,743	0,864	1,168
P	0,7	0,8	0,9	0,95	0,97	0,98	0,99	0,995	0,999

Вираз (3.2) - це математичне очікування інтегрального квадрату різниці емпіричної та гіпотетичної (теоретичної) функцій розподілу. Запишемо формулу (3.2) у вигляді

$$V = n \int_{-\infty}^{\infty} F^2(x) W(x) dx - 2n \int_{-\infty}^{\infty} F(x) F_n^* \left(\frac{x}{|x|} \right) W(x) dx + n \int_{-\infty}^{\infty} \left(F_n^* \left(\frac{x}{|x|} \right) \right)^2 W(x) dx.$$

Після обчислення цих інтегралів одержимо формулу для визначення показника близькості за критерієм $n\omega$ -квадрат

$$V = \sum_{k=1}^n \left(F(\xi_k) - \frac{2k-1}{2n} \right)^2 + \frac{1}{12n},$$

де $|\xi| = |\xi_1, \xi_2, \dots, \xi_n|$ - впорядкована за зростанням вибірка вимірювань $|x| = |x_1, x_2, \dots, x_n|$.

Перевірка правильності гіпотези формулюється в такий спосіб: якщо $V \leq V_0$, то гіпотеза вважається вірною з ймовірністю P .

2. Завдання на дослідження

Частина 1. Перевірка гіпотези про закон розподілу ймовірності (за критерієм χ^2 -квадрат).

Дано: вибірка вимірювань з невідомим законом розподілу ймовірності та довжиною $n \geq 100$; задана ймовірність $P = 0,97$.

Визначити: закон розподілу заданої вибірки шляхом перевірки різноманітних гіпотез за критерієм χ^2 -квадрат.

Частина 2. Перевірка гіпотези про функцію розподілу ймовірностей (за критерієм $n\omega$ -квадрат).

Дано: коротка вибірка вимірювань ($20 \leq n \leq 50$), що сформована із довгої вибірки (див. частину 1 даної лабораторної роботи) шляхом її розрідження.

Визначити: функцію розподілу ймовірності короткої вибірки за критерієм $n\omega$ -квадрат.

3. Порядок виконання роботи

Частина 1. Перевірка гіпотези про закон розподілу ймовірності (за критерієм χ^2 -квадрат).

- 1) Завантажити вибірку випадкових величин (див. на диску у викладача) в MathCad та визначити її довжину. За вибіркою оцінити вибіркове середнє та вибіркочову дисперсію.
- 2) Упорядкувати вибірку за зростанням.
- 3) Визначити мінімальне та максимальне значення вибірки. Задати кількість інтервалів гістограми.
- 4) Обчислити інтервал дискретизації та центри інтервалів гістограми.
- 5) Визначити гістограмні числа та побудувати графік гістограми.
- 6) Проаналізувати вид отриманої гістограми та висунути гіпотезу про закон розподілу ймовірностей заданих випадкових величин.
- 7) Оцінити параметри гіпотетичного закону.
- 8) Обчислити ймовірності влучення вимірювань в кожний інтервал гістограми.
- 9) Обчислити показник близькості χ^2 -квадрат та граничне значення.
- 10) Записати вирішальне правило та прийняти або відхилити дану гіпотезу про закон розподілу.
- 11) Перевірити щонайменше три ймовірні гіпотези.
- 12) Показники близькості та граничні значення, які отримані в результаті дослідження, записати в таблицю. Також занести до звіту гістограму та графіки гіпотетичних законів розподілу.
- 13) Проаналізувати та зробити висновок.

Частина 2. Перевірка гіпотези про функцію розподілу ймовірності (за критерієм $n\omega$ -квадрат).

- 1) Сформуувати вибірку вимірювань довжиною $20 \leq n \leq 50$ шляхом розрідження довгої вибірки із першої частини даної лабораторної роботи.
- 2) Обчислити вибіркочове середнє та вибіркочову дисперсію.
- 3) Упорядкувати вибірку за зростанням.
- 4) Побудувати емпіричну функцію розподілу випадкових величин та представити її графік.
- 5) Висунути гіпотезу про функцію розподілу ймовірності.
- 6) Оцінити параметри гіпотетичної функції й обчислити показник близькості V за критерієм $n\omega$ -квадрат.
- 7) Сформуувати вирішальне правило, тобто порівняти отриманий показник з граничним значенням, та прийняти рішення про вид функції розподілу випадкових величин.

- 8) Перевірити щонайменше три гіпотези про функцію розподілу ймовірності.
- 9) Записати отримані показники та граничні значення. представити графіки емпіричної та гіпотетичних функцій розподілу ймовірностей.
- 10) Проаналізувати отримані результати та зробити висновки.

4. Застосування математичного середовища MathCad для розв'язання задач даного типу

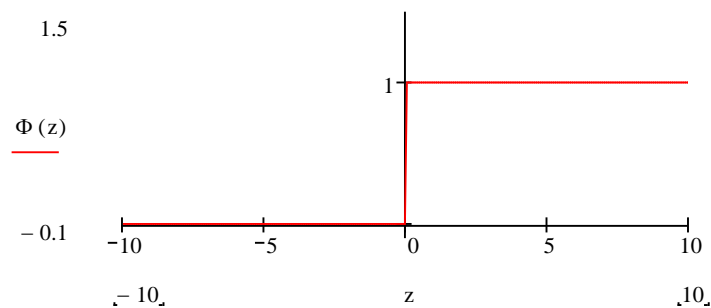
Для роботи з масивами даних в середовищі MathCad вбудовано багато функцій. При виконанні третьої лабораторної роботи буде корисним знання наступних.

Функція $sort(x)$ упорядковує за зростанням вибірку x .

Функція одиничного стрибка $sgn(z)$, яка називається також функцією Хевісайда

$$sgn(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

вбудована в MathCad з іменем *heaviside step* та позначається $\Phi(z)$:



Для формування вирішальних правил та перевірки виконання нерівностей зручним є використання функції $if(cond, x, y)$, де $cond$ - це умова (нерівність) виконання якої ви перевіряєте, x - значення функції або змінної, яке вона набуває при виконанні умови, y - значення функції або змінної при не виконанні заданої умови.

Для перевірки гіпотез про функції розподілу ймовірностей будуть у нагоді наступні функції:

- $pnorm(x, \mu, \sigma)$ - функція розподілу нормальних випадкових величин з математичним очікуванням μ , та середнє квадратичним відхиленням σ .
- $pexp(x, \lambda)$ - функція розподілу експоненціальних випадкових величин з параметром λ .
- $pgamma(x, \alpha)$ - функція гамма-розподілених випадкових величин з параметрами $\beta=1$ та α .
- і т.д., за аналогічним принципом запису вбудовано багато функцій розподілу ймовірностей.

Лабораторна робота №4

«Побудова й аналіз моделей законів розподілу ймовірностей»

1. Теоретична частина

Формування вирішальних правил розпізнавання об'єктів неруйнівного контролю пов'язане з необхідністю оцінювання законів розподілу ймовірностей інформативних параметрів за експериментальними даними. Передбачається, що відомі вибірки вимірювань параметрів, що характеризують нормальний та дефектний стани об'єктів контролю. Необхідно по цих вибірках вимірювань сформулювати адекватні математичні моделі їх законів розподілу. Розглянемо два методи побудови моделей законів розподілу ймовірностей.

Моделі законів розподілу ймовірностей на основі рядів ортонормованих поліномів.

Оцінка закону розподілу ймовірності на основі рядів ортонормованих поліномів має вигляд

$$W^*(x) = A \sum_{i=0}^N b_i^* \sqrt{q(x)} Q_i(x), \quad (4.1)$$

де $Q_i(x)$ - ортонормований поліном i -го порядку, $q(x)$ - вагова функція, коефіцієнти b_i^* визначаються за формулою

$$b_i^* = \frac{1}{n} \sum_{k=1}^n \sqrt{q(x_k)} Q_i(x_k),$$

нормуючий множник A обчислюється із умови рівності одиниці площі під кривою закону розподілу $W^*(x)$:

$$A = \left(\sum_{i=0}^N b_i^* q_i \right)^{-1}, \quad q_i = \int_{-\infty}^{\infty} \sqrt{q(x)} Q_i(x) dx.$$

Запишемо вираз для інтегрального квадрата помилки оцінювання закону розподілу ймовірностей на основі даного методу

$$\delta^2 = \int_{-\infty}^{\infty} \left(W(x) - A \sum_{i=0}^N b_i^* \sqrt{q(x)} Q_i(x) \right)^2 dx. \quad (4.2)$$

Згідно з (4.2) інтегральний квадрат помилки є випадковою величиною й залежить від виду ортонормованого полінома, розмірів вибірки вимірювань, числа членів ряду.

Зі структури ряду (4.1) виявляється, що якщо вид функції $\sqrt{q(x)}$ буде схожий на закон розподілу $W(x)$, то можна чекати, що число членів ряду N , достатнє для його ефективної оцінки, може бути невеликим.

Щоб обґрунтовано вибрати число членів ряду N пропонується використовувати правило «двох нулів». Це правило ґрунтується на перевірці нерівності:

$$-3\sqrt{D^*[b_i^*]} \leq (b_i^* - b_i) \leq 3\sqrt{D^*[b_i^*]},$$

де вибіркова дисперсія оцінки обчислюється за формулою:

$$D^*[b_i^*] = \frac{1}{n(n-1)} \sum_{k=1}^n (\sqrt{q(x_k)} Q_i(x_k) - b_i^*)^2.$$

Якщо для оцінки коефіцієнта b_i^* виконується нерівність

$$-3\sqrt{D^*[b_i^*]} \leq b_i^* \leq 3\sqrt{D^*[b_i^*]}, \quad (4.3)$$

то його математичне очікування дорівнює нулю. Кожна оцінка b_i^* повинна перевірятися «на нуль» відповідно до правила (4.3). В зв'язку з тим, що структура деяких рядів припускає наявність нульових коефіцієнтів, ряд варто обривати, якщо виявлено два нулі підряд. Таким чином, якщо коефіцієнти b_i^* й b_{i+1}^* задовольняють нерівності (4.3), то число членів ряду варто вибрати $i-1$ (тобто $N=i-1$).

Розглянемо деякі розповсюджені види ортонормованих поліномів, які використовуються для формування моделей законів розподілу ймовірностей.

Поліноми Ерміта $H_i(x)$ з ваговою функцією

$$h(x) = e^{-x^2}, \quad -\infty \leq x \leq \infty.$$

Очевидно, що їх можна застосовувати для оцінки симетричних законів розподілу типу Гауса з нульовим математичним очікуванням і одиничною дисперсією. Ортонормовані поліноми Ерміта описуються рекурентною формулою

$$\sqrt{(i+1)2^{i+1}\sqrt{\pi}} H_{i+1}(x) = 2x\sqrt{i2^i\sqrt{\pi}} H_i(x) - 2i\sqrt{(i-1)2^{i-1}\sqrt{\pi}} H_{i-1}(x).$$

Перші три поліноми дорівнюють: $H_0(x) = \frac{1}{\sqrt{\pi}}$, $H_1(x) = \frac{\sqrt{2}x}{\sqrt{\pi}}$, $H_2(x) = \frac{2x^2 - 1}{\sqrt{2\sqrt{\pi}}}$.

При оцінці закону розподілу вимірювань y_1, y_2, \dots, y_n , у котрого огинаюча гістограма нагадує симетричну гаусову криву, необхідно визначити середнє значення \bar{y} й вибіркову дисперсію D_y^* й нормувати вихідну вибірку в такий спосіб

$$x_k = \frac{y_k - \bar{y}}{\sqrt{D_y^*}}.$$

При даному нормуванні складові вимірювання вибірки x_1, x_2, \dots, x_n будуть мати близьке до нуля математичне очікування й майже одиничну дисперсію. Оцінивши закон розподілу випадкових величин x_1, x_2, \dots, x_n рядом (4.1), одержимо

$$W_x^*(x) = A \sum_{i=0}^N b_i^* \exp\left(-\frac{x^2}{2}\right) H_i(x).$$

Таким чином, модель закону розподілу вихідної вибірки y_1, y_2, \dots, y_n запишеться у вигляді

$$W^*(y|y) = W_x^*\left(\frac{y - \bar{y}}{\sqrt{D_y^*}}\right) \frac{1}{\sqrt{D_y^*}}. \quad (4.4)$$

Поліноми Лагерра $L^{(0)}_i(x)$ з ваговою функцією

$$l(x) = e^{-x}, \quad 0 \leq x \leq \infty$$

застосовуються при оцінці вимірювань з асиметричними законами розподілу ймовірностей, схожих на експоненціальний або на півнормальний закони. Ортонормовані поліноми Лагерра визначаються за рекурентною формулою

$$(i+1)L^{(0)}_{i+1}(x) = (2i+1-x)L^{(0)}_i(x) - i^2(i-1)L^{(0)}_{i-1}(x).$$

Перші три поліноми дорівнюють

$$L^{(0)}_0(x) = 1, \quad L^{(0)}_1(x) = 1-x, \quad L^{(0)}_2(x) = \frac{1}{2}(x^2 - 4x + 2).$$

При оцінці закону розподілу вибірки y_1, y_2, \dots, y_n ортонормованими рядами Лагерра $L^{(0)}_i(x)$ необхідно визначити середнє значення \bar{y} й нормувати вихідну вибірку

$$x_k = \frac{y_k}{\bar{y}}.$$

Оцінивши закон розподілу випадкових величин x_1, x_2, \dots, x_n рядом (4.1)

$$W_x^*(x) = A \sum_{i=0}^N b_i^* \exp\left(-\frac{x}{2}\right) L^{(0)}_i(x),$$

запишемо оцінюваний закон розподілу вихідної вибірки y_1, y_2, \dots, y_n

$$W^*(y) = W_x^*\left(\frac{y}{\bar{y}}\right) \frac{1}{\bar{y}}. \quad (4.5)$$

Поліноми Лагерра $L^{(1)}_i(x)$ з ваговою функцією

$$l(x) = xe^{-x}, \quad 0 \leq x \leq \infty$$

застосовуються для оцінки законів розподілу зі слабкою асиметрією типу розподілу Релея або гамма-розподілу. Для поліномів Лагерра $L^{(1)}_i(x)$ справедлива рекурентна формула

$$\sqrt{(i+1)(i+2)}L^{(1)}_{i+1}(x) = (x-2i-2)L^{(1)}_i(x) - \sqrt{i(i+1)}L^{(1)}_{i-1}(x).$$

Перші три поліноми дорівнюють

$$L^{(1)}_0(x) = 1, \quad L^{(1)}_1(x) = \frac{x-2}{\sqrt{2}}, \quad L^{(1)}_2(x) = \frac{3-3x+0,5x^2}{\sqrt{3}}.$$

Вагова функція $l(x) = xe^{-x}$ має максимум при $x_m = 1$. Якщо задано вибірку вимірювань y_1, y_2, \dots, y_n , то за гістограмою можна оцінити найбільш ймовірне значення y_m й перетворити вибірку в нормовану в такий спосіб

$$x_k = \frac{y_k}{y_m}.$$

Оцінивши закон розподілу випадкових величин x_1, x_2, \dots, x_n рядом (4.1)

$$W_x^*(x) = A \sum_{i=0}^N b_i^* \sqrt{x} \exp\left(-\frac{x}{2}\right) L^{(1)}_i(x),$$

запишемо модель закону розподілу вихідної вибірки y_1, y_2, \dots, y_n

$$W^*(y) = W_x^*\left(\frac{y}{y_m}\right) \frac{1}{y_m}. \quad (4.6)$$

Поліноми Лежандра $P_i(x)$ з одиничною ваговою функцією

$$p(x) = 1, \quad -1 \leq x \leq 1$$

використовуються для оцінки законів розподілу, близьких до рівномірного в інтервалі ± 1 . Ортонормовані поліноми Лежандра $P_i(x)$ описуються рекурентною формулою

$$\frac{1}{\sqrt{2i+3}} P_{i+1}(x) = x \frac{\sqrt{2i+1}}{(i+1)} P_i(x) - \frac{i}{(i+1)\sqrt{2i-1}} P_{i-1}(x).$$

Перші три поліноми дорівнюють

$$P_0(x) = \frac{1}{\sqrt{2}}, P_1(x) = \frac{\sqrt{3}x}{\sqrt{2}}, P_2(x) = \frac{\sqrt{5}(3x^2-1)}{2\sqrt{2}}.$$

Вихідну вибірку y_1, y_2, \dots, y_k необхідно перетворювати в такий спосіб

$$x_k = \frac{2y_k - (y_{\max} + y_{\min})}{y_{\max} - y_{\min}},$$

де y_{\max} - максимальне, а y_{\min} - мінімальне значення вибірки $|y|$.

Модель закону розподілу запишеться у вигляді

$$W^*(y) = W_x^* \left(\frac{2y - (y_{\max} + y_{\min})}{y_{\max} - y_{\min}} \right) \frac{2}{(y_{\max} - y_{\min})}, \quad (4.7)$$

де $W_x^*(x) = A \sum_{i=0}^N b_i^* P_i(x)$.

Таким чином, щоб формувати моделі законів розподілу, необхідно по вибірках вимірювань будувати гістограми й вибирати систему ортонормованих поліномів відповідно до виду гістограм. В умовах обмеженості вхідних даних визначити вид закону розподілу за гістограмою важко. Отже, при використанні коротких вибірок ($n < 100$) буде велика ймовірність помилитися у виборі вагової функції й це позначиться на точності оцінок.

Моделі законів розподілу ймовірностей на основі рядів згладжених дельта-функцій

Оцінка закону розподілу ймовірності на основі дельта-функцій, які апроксимовані Гаусовою кривою, має вигляд

$$W^*(x/|x|) = \frac{\alpha}{n\sqrt{2\pi}} \sum_{k=1}^n \exp \left[-\frac{\alpha^2}{2} (x - x_k)^2 \right], \quad (4.8)$$

де α - параметр згладжування, n - довжина вибірки вимірювань, x_k - k -те вимірювання експериментальних даних $|x|$.

Труднощі використання цієї моделі пов'язані з вибором параметра α . В [1] запропоновано формули для оцінки показників згладжування вимірювань зі слабкою й сильною асиметрією

$$\alpha_{1a}^* = \frac{1.592 + 0.018n}{\sqrt{D^*}}, \quad (4.9)$$

$$\alpha_{2a}^* = \frac{3.93 + 0.07 \ln n}{\sqrt{D^*}}. \quad (4.10)$$

Таким чином, щоб на практиці обрати параметр згладжування, необхідно обчислити вибірковий коефіцієнт асиметрії k_{as}^* експериментальних даних. Якщо

$k_{as}^* < 1,7$, то для обчислення параметра згладжування використовують формулу (4.9), а якщо $k_{as}^* \geq 1,7$, то формулу (4.10).

У разі, коли відома апріорна інформація про характер вимірювань, можна скористатися наступною моделлю закону розподілу ймовірності на основі рядів згладжених дельта-функцій

$$W_A^*(x/|x) = AW^*(x/|x) = \frac{A\alpha}{n\sqrt{2\pi}} \sum_{k=1}^n \exp\left[-\frac{\alpha^2}{2}(x-x_k)^2\right]. \quad (4.11)$$

Ця модель доповнена нормуючими множниками, які враховують таку інформацію:

- якщо відомо, що випадкові величини лише позитивні ($x \geq 0$), то нормуючий множник має вигляд

$$A = \frac{n}{n - \sum_{k=1}^n (1 - \Phi(\alpha x_k))},$$

де $\Phi(z)$ - інтеграл ймовірності Гауса;

- якщо відомо, що випадкові величини обмежені інтервалом $[a, b]$ ($a \leq x \leq b$), то нормуючий множник обчислюється за формулою

$$A = \frac{n}{n - \sum_{k=1}^n \Phi(\alpha(a - x_k)) - \sum_{k=1}^n (1 - \Phi(\alpha(b - x_k)))}$$

Показником ефективності оцінювання законів розподілу на основі згладжених дельта-функцій може виступати, як і в першому випадку, інтегральний квадрат помилки:

$$\delta^2 = \int_{-\infty}^{\infty} (W(x) - W^*(x/|x))^2 dx \quad (4.12)$$

2. Завдання на дослідження

Частина 1. Формування моделей законів розподілу ймовірностей на основі рядів ортонормованих поліномів.

Дано: вибірка вимірювань, довжина вибірки $n = 100$.

Визначити: модель закону розподілу ймовірності вибірки вимірювань на основі рядів ортонормованих поліномів.

Частина 2. Формування моделей законів розподілу ймовірностей на основі рядів згладжених дельта-функцій.

Дано: вибірка вимірювань (дивись частину 1 цієї лабораторної роботи)

Визначити: модель закону розподілу ймовірності на основі рядів згладжених дельта-функцій.

3. Порядок виконання роботи

Частина 1.

- 1) Завантажити вибірку вимірювань відповідно до заданого викладачем варіанту (див. файли від викладача).
- 2) Оцінити статистичні характеристики заданої вибірки (математичне очікування, дисперсію та коефіцієнт асиметрії).
- 3) Знайти максимальне та мінімальне значення вибірки вимірювань, задати кількість інтервалів гістограми та обчислити границі інтервалів.
- 4) Визначити гістограмні числа та побудувати гістограму.
- 5) Проаналізувати вид отриманої гістограми та обрати вид вагової функції та ортонормованих поліномів для побудування моделі закону розподілу.
- 6) Провести нормування початкових даних. Обчислити вибіркове середнє та вибіркочну дисперсію нормованих даних й впевнитися, що нормування проведено коректно.
- 7) Обчислили довжину ряду ортонормованих поліномів N на основі правила “двох нулів”.
- 8) Сформувати модель закону розподілу на основі нормованих даних.
- 9) Перейти до моделі закону розподілу для заданої вибірки.
- 10) Побудувати графік закону розподілу за отриманою моделлю.

Частина 2.

- 11) Продовжити роботу з вибіркою із першої частини лабораторної роботи.
- 12) Оцінити параметр згладжування α за формулами (4.9) або (4.10) в залежності від значення вибіркового коефіцієнту.
- 13) Якщо характер вимірювань заданої вибірки володіє такими властивостями, що $x \geq 0$ або $a \leq x \leq b$, то для побудови моделі на основі згладжених дельта-функцій можна скористатися формулою (4.11). Якщо апріорної інформації про характер вибірки отримати не можна, то модель закону розподілу необхідно формувати за формулою (4.8)
- 14) Побудувати графік закону розподілу ймовірності за побудованою моделлю.

Після побудови двох моделей законів розподілу ймовірностей за кожним із методів, отримати у викладача вид та параметри теоретичного закону розподілу й продовжити виконання лабораторної роботи.

- 15) Обчислити помилки оцінювання законів розподілу на основі побудованих моделей за формулами (4.2) та (4.12). Порівняти між собою отримані значення помилок оцінювання законів розподілу.
- 16) Зробити висновки щодо ефективності моделей на основі рядів ортонормованих поліномів та згладжених дельта-функцій.

4. Застосування математичного середовища MathCad для розв'язання задач даного типу

Щоб отримати ортонормовані поліноми, можна скористатися вбудованими в MathCad ортогональними поліномами та провести їх нормування. Таким чином,

- ортонормовані поліноми Ерміта запишуться у вигляді:

$$Q(i, v) := \frac{1}{\sqrt{2^i \cdot i! \cdot \sqrt{\pi}}} \cdot \text{Her}(i, v),$$

- ортонормовані поліноми Лагерра нульового ступеню $L^{(0)}_i(x)$:

$$Q(i, v) := \text{if}(v \geq 0, \text{Lag}(i, v), 0),$$

- ортонормовані поліноми Лагерра першого ступеню $L^{(1)}_i(x)$:

$$Q(i, v) := \frac{(-1)^i \cdot \sqrt{i!}}{\sqrt{G(i+2)}} \cdot \frac{(i+1) \cdot \text{Lag}(i, v) - (i+1) \cdot \text{Lag}(i+1, v)}{v},$$

де $G(v)$ - гамма-функція, яка обчислюється за формулою

$$G(v) := \int_0^{\infty} e^{-t} \cdot t^{v-1} dt,$$

- ортонормовані поліноми Лежандра:

$$Q(i, v) := \text{if}\left(-1 \leq v \leq 1, \frac{\sqrt{2 \cdot i + 1}}{\sqrt{2}} \cdot \text{Leg}(i, v), 0\right).$$

Для реалізації правила “двох нулів” в MathCad зручно скористатися циклом *for*:

$$\begin{aligned} N &:= \text{for } i \in N1..0 \\ N &\leftarrow i \text{ if } R_i = 1 \wedge R_{i+1} = 0 \wedge R_{i+2} = 0, \end{aligned}$$

де $N1$ - початкове значення числа членів ряду ортонормованих поліномів, обране заздалегідь велике, R - це вектор, що містить результат перевірки виконання нерівності (4.3):

$$R_i := \text{if}\left(-3 \cdot \sqrt{Db_i} \leq b_i \leq 3 \cdot \sqrt{Db_i}, 0, 1\right).$$

Вимоги до оформлення звітів за лабораторними роботами

Для захисту лабораторної роботи до усної відповіді студента повинен додаватися рукописний звіт. В звіті необхідно вказати:

- тему лабораторної роботи;
- її мету;
- детальний опис ходу роботи (вхідні дані, послідовності дій, використані формули, припущення та алгоритми);
- результати дослідження (таблиці, графіки);
- висновки.

Звіт повинен бути придатним до перевірки його даних та давати такі самі результати при перевірці його третім лицем.

Таблиці варіантів

Таблиця 5.1

Номер варіанту	Вид закону розподілу	Відомі параметри закону розподілу	Невідомий параметр, який необхідно оцінити
1	Експоненціальний	-	λ
2	Релея	-	b
3	Нормальний	$\mu = 0,5$	σ
4	Гамма-розподіл	$\beta = 1$	α
5	Лапласа	$a = 0$	λ
6	Вейбулла	$\beta = 1$	α
7	Логістичний	$\sigma = 1$	a
8	Релея	-	b
9	Нормальний	$\mu = 4$	σ
10	Експоненціальний	-	λ
11	Вейбулла	$\alpha = 1$	β
12	Лапласа	$a = 0,5$	λ
13	Релея	-	b
14	Експоненціальний	-	λ
15	Гамма-розподіл	$\beta = 1$	α
16	Релея	-	b
17	Нормальний	$\sigma = 1$	μ
18	Вейбулла	$\beta = 1$	α
19	Лапласа	$a = 2,1$	λ
20	Логістичний	$a = 0$	σ
21	Експоненціальний	-	λ
22	Гамма-розподіл	$\beta = 1$	α
23	Релея	-	b
24	Нормальний	$\mu = -3$	σ
25	Вейбулла	$\beta = 1$	α

Таблиця 5.2

Номер варіанту	a	σ^2	P
1	0	0,7	0,95
2	1	0,5	0,97
3	2	1	0,9
4	3	1,1	0,97
5	2,5	1,5	0,99
6	1,5	0,1	0,95
7	-1	0,3	0,93
8	-2	1,2	0,91
9	-3	2	0,99
10	0	1,9	0,98
11	-2,5	0,4	0,95
12	-1,5	1,7	0,96

Номер варіанту	a	σ^2	P
13	0,1	1,85	0,93
14	0,9	0,6	0,94
15	25	0,75	0,91
16	10	0,8	0,92
17	0,3	0,65	0,95
18	16	1,4	0,9
19	0,5	2,01	0,99
20	-0,5	1,99	0,97
21	7	1,3	0,92
22	4	0,3	0,95
23	1,9	0,55	0,98
24	8	1	0,99
25	15	1,4	0,93

Список рекомендованої літератури

1. Лысенко Н.А., Малайчук В.П. Восстановление законов распределения вероятностей рядами сглаженных дельта-функций // Актуальні проблеми автоматизації та інформаційних технологій. – Том 7. – Дніпропетровськ: ДНУ. - 2003. – С. 61-74.
2. Малайчук В.П., Мозговой А.В. Математическая дефектоскопия. – Днепропетровск: Системные технологии, 2005. – 180 с.
3. Бабак В.П., Білецький А.Я., Приставка О.П., Приставка П.О. Основи теорії ймовірностей та математичної статистики: Навчальний посібник. – К.: КВІЦ, 2003. – 432 с.
4. Малайчук В.П., Петренко А.Н., Лысенко Н.А. Восстановление законов распределения вероятностей рядами ортонормированных полиномов // Системні технології. Регіональний міжвузівський збірник наукових праць – Випуск 3 (38) – Дніпропетровськ: “Системні технології”. – 2005. – С. 73-88.
5. Малайчук В.П., Петренко О.М., Рожковський В.Ф. Основи теорії ймовірності і математичної статистики: Навчальний посібник – Дніпропетровськ: РВВ ДНУ, 2001. – 164 с.
6. Лысенко Н.А., Малайчук В.П., Мозговой А.В. Восстановление законов распределения вероятностей в задачах обработки измерений неразрушающего контроля // Материалы одиннадцатой международной конференции «Современные методы и средства неразрушающего контроля и технической диагностики» (г. Ялта). – К.: УИЦ “Наука. Техника. Технология”. – 2003. – С.126-128.
7. Малайчук В.П. Основы теории обработки сигналов в технических системах управления и контроля. – Днепропетровск: Изд-во ДГУ, 1990. – 116 с.

Зміст

<i>Введення</i>	3
<i>Лабораторна робота №1 «Оцінка параметрів амплітудно-частотних характеристик лінійних фільтрів»</i>	4
1. Теоретична частина.	4
2. Завдання на дослідження.	5
3. Порядок виконання роботи.	6
4. Застосування математичного середовища MathCad для розв’язання задач даного типу.	6
<i>Лабораторна робота №2 «Оцінка параметрів законів розподілу ймовірностей і аналіз результатів обчислювальних експериментів»</i>	8
1. Теоретична частина.	8
2. Завдання на дослідження.	10
3. Порядок виконання роботи.	10
4. Застосування математичного середовища MathCad для розв’язання задач даного типу.	12
<i>Лабораторна робота №3 «Перевірка гіпотез про закони й функції розподілу ймовірностей»</i>	14
1. Теоретична частина.	14
2. Завдання на дослідження.	16
3. Порядок виконання роботи.	17
4. Застосування математичного середовища MathCad для розв’язання задач даного типу.	18
<i>Лабораторна робота №4 «Побудова й аналіз моделей законів розподілу ймовірностей»</i>	19
1. Теоретична частина.	19
2. Завдання на дослідження.	23
3. Порядок виконання роботи.	24
4. Застосування математичного середовища MathCad для розв’язання задач даного типу.	25
<i>Вимоги до оформлення звітів за лабораторними роботами</i>	25
<i>Таблиці варіантів</i>	26
<i>Список рекомендованої літератури</i>	27