

**МІНІСТЕРСТВО ВНУТРІШНІХ СПРАВ УКРАЇНИ
КРЕМЕНЧУЦЬКИЙ ЛЬОТНИЙ КОЛЕДЖ
ХАРКІВСЬКОГО НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ
ВНУТРІШНІХ СПРАВ**

Циклова комісія природничих дисциплін

ТЕКСТ ЛЕКЦІЇ

навчальної дисципліни «Вища математика»
обов'язкових компонент
освітньо-професійної програми
першого (бакалаврського) рівня вищої освіти
Аеронавігація

**за темою № 18: Обробка та відображення експериментальних
даних. Аналіз результатів експерименту**

Харків 2022

ЗАТВЕРДЖЕНО

Науково-методичною радою
Харківського національного
університету внутрішніх справ
Протокол від 30.08. 2022 №8

СХВАЛЕНО

Методичною радою Кременчуцького
льотного коледжу Харківського
національного університету
внутрішніх справ
Протокол від 22.08. 2022 №1

ПОГОДЖЕНО

Секцією Науково-методичної ради
ХНУВС з технічних дисциплін
Протокол від 29.08. 2022 №8

Розглянуто на засіданні циклової комісії природничих дисциплін,
протокол від 10.08.2021 № 1

Розробник:

Доцент циклової комісії природничих дисциплін, к.ф.–м.н. Семенов В.О.

Рецензенти:

1. Доцент кафедри автомобілі та трактори Кременчуцького національного університету імені Михайла Остроградського, кандидат технічних наук Черниш А.А.
2. Професор циклової комісії аеронавігації Кременчуцького льотного коледжу Харківського національного університету внутрішніх справ, кандидат технічних наук, старший науковий співробітник Тягній В.Г.

ЛЕКЦІЯ 30

План лекції

1. Первинна обробка результатів спостережень.
2. Вторинна обробка результатів спостережень.

Рекомендована література:

Основна

1. Астахов В. М. Теорія ймовірностей і математична статистика : навчальний посібник / В. М. Астахов, Г. С. Буланов, В. О. Паламарчук. – Краматорськ : ДДМА, 2009. – 64 с.
2. Барковський В. В. Теорія ймовірностей та математична статистика ; 5-те вид. : навчальний посібник . – К. : Центр учбової літератури, 2010. – 424 с

Додаткова

3. Волощенко А. Б. Теорія ймовірностей і математична статистика : навчально-методичний посібник для самостійного вивчення дисципліни / А. Б. Волощенко, І. А. Джалладова. – К. : КНЕУ, 2003. – 256 с.

Інформаційні ресурси в інтернеті

4. Жильцов О.Б. Теорія ймовірностей та математична статистика у прикладах і задачах: навч. посіб.— К. : Київ. ун-т ім. Б. Грінченка, 2015. - 336 с.
http://elibrary.kubg.edu.ua/id/eprint/13578/1/O_Zhyltsov_KUBG_TY_UN.pdf
5. Федоров М.В., Хренов О.М. Теорія ймовірностей і математична статистика: Конспект лекцій. – Х.: ХНАМГ, 2011. – 68 с.
https://docviewer.yandex.ua/view/14307034/?page=3&*=OMbtom834MHYY%2B0pcfIKdrchzlx7%7D }}

Текст лекції 30

1. Первинна обробка результатів спостережень.

Види похибок.

Відхилення виміряного значення x величини, що визначається, від істинного її значення $M(X)$ називається похибкою i -го вимірювання Δx_i . Похибки вимірювання прийнято розділяти на два типи – систематичні та випадкові. *Систематичними* називають такі похибки, які викликають фактори, що діють однаковим чином при багатократному повторенні одних й тих самих вимірів. Систематичні похибки, в принципі, можуть бути усунуті або враховані, хоча виявлення джерела цих похибок є справою надзвичайно складною. Будь – який експеримент не гарантований від наявності неврахованої систематичної похибки.

Випадкові похибки (відхилення) завжди присутні в експерименті. При відсутності систематичних похибок вони слугують причиною розкиду результатів повторних вимірів як між собою, так і відносно істинного

значення величини, що вимірюється. Природа випадкових похибок може бути різноманітною. Випадкові похибки послідовних вимірів як правило незалежні та характеризуються тим чи іншим законом розподілу. Вони мають властивість концентрації, тобто малі за абсолютною величиною випадкові помилки зустрічаються частіше, ніж великі. Якщо вдається знизити до достатньо низького рівня систематичні похибки, то точність вимірювань визначається тепер випадковими похибками, а, отже, може бути піддана статистичному аналізу.

В курсі статистики доведено, що найкращою оцінкою величини X , що вимірюється, є середня вибірка, а найкращою оцінкою середнього квадратичного відхилення є виправлене середнє квадратичне відхилення. Відповідно з цим маємо наступну схему:

Схема первинної обробки результатів спостереження є наступною:

а) визначення середнього значення отриманих результатів, тобто обчислення середньої арифметичної \bar{y} за формулою:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

б) визначення відхилення від середнього значення для кожного результату за формулою: $\Delta y_i = y_i - \bar{y}$. Ці відхилення характеризують абсолютну помилку вимірювання. Випадкові помилки мають різні знаки. Коли значення результату дослідження перевищує середнє значення, похибка дослідження вважається позитивною. Коли ж значення результату дослідження менше середнього значення, похибка вважається негативною. Чим точніше проведені виміри, тим ближче значення окремих результатів і середнє значення;

в) обчислення дисперсії S^2 за формулою:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

г) обчислення стандартного відхилення окремого значення за формулою:

$$S_{(y_i)} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

і стандартного відхилення середнього результату за формулою

$$S_{(\bar{y})} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (y_i - \bar{y})^2}$$

д) перевірка надійності отриманих результатів за критерієм Стюдента та для проведеної кількості дослідів n при обраній надійній імовірності α . У більшості випадків при дослідженнях приймають $\alpha = 0,95$ або $0,99$.

Це означає, що 95 % або 99 % абсолютних відхилень результатів лежать в означених межах. Критерій t_α з надійною ймовірністю α показує, у скільки разів модуль різниці між істинним значенням величини y , що визначається, і середнім значенням менше стандартного відхилення середнього результату відповідно t_α

$$t_{\alpha} = \frac{|y - \bar{y}|}{S_{(\bar{y})}}$$

Знайшовши за таблицею значення t_{α} та використовуючи раніше отримане значення $S_{(\bar{y})}$, розраховують помилку отриманого середнього результату

$$\varepsilon = t_{\alpha} S_{(\bar{y})}$$

е) Встановлення інтервалу, в якому з надійною ймовірністю буде знаходитися середній результат $\bar{y} \pm \varepsilon$.

ж) Визначення відносної помилки за формулою: $\Delta y = \frac{\varepsilon \cdot 100}{\bar{y}}$.

Якщо значення Δy велике порівняно зі значенням y , то результати, що обробляються, перевіряють за одним з вищеописаних способів (за критерієм максимального відхилення Стюдента або t_{α} на наявність грубих помилок. Для такої перевірки можуть застосовуватися й інші методи.

2. Вторинна обробка результатів спостережень.

Після виключення грубих помилок проводиться вторинна обробка вже без виключених експериментальних даних.

Вторинні методи застосовуються для того, щоб на базі зібраних даних виявити приховані в них статистичні закономірності. Наприклад, методи дисперсійного, регресійного аналізу застосовуються для того, щоб визначити динаміку зміни окремих параметрів вибірки. Кореляційний, факторний аналіз здійснюється для встановлення статистичних зв'язків, що існують між змінними величинами, вимірюваними в експерименті. Методи тестування статистичних гіпотез та визначення довірчих інтервалів дозволяють теоретично обґрунтувати висновки стосовно параметрів та вигляду отриманого розподілу. Статистичні методи дають можливість довести, що отримано дійсно не випадкові результати і підтвердити існування виявлених.

За допомогою вторинних методів статистичної обробки експериментальних даних безпосередньо перевіряються, доводяться або спростовуються гіпотези, пов'язані з експериментом. Ці методи, як правило, складніші, ніж методи первинної статистичної обробки.

Обговорювану групу методів можна розділити на кілька підгруп:

- 1) Регресійне числення.
2. Методи порівняння між собою двох або декількох елементарних статистик (середніх, дисперсій і т.п.), що відносяться до різних вибірок.
3. Методи встановлення статистичних взаємозв'язків між змінними, наприклад їх кореляції одна з одною.
4. Методи виявлення внутрішньої статистичної структури емпіричних даних (наприклад, факторний аналіз).

Щонайкраще відтворити шукану залежність факторів часто допомагає метод найменших квадратів (МНК), при використанні якого вимога найбільш ймовірного узгодження встановлюваної залежності з даними експерименту

зводиться до того, щоб сума квадратів відхилень експериментальних значень від заданої лінії зверталася в мінімум:

$$\sum_{i=1}^n (y_i - \varphi(x))^2 = \min, \text{ де } y_i - \text{експериментальні значення, } \varphi(x) - \text{шукана залежність.}$$

Розглянемо приклад криволінійної кореляції, а саме параболічну кореляцію. Нехай зі зростанням однієї випадкової величини умовні середні другої зростають (спадають), досягають максимуму (мінімуму), а потім спадають (зростають). Тоді можна вважати, що між ними існує параболічна залежність виду:

$$\bar{y}_x = a_2 x^2 + a_1 x + a_0.$$

За методом найменших квадратів для визначення значень параметрів a_2 , a_1 , a_0 потрібно скласти і розв'язати систему рівнянь:

$$\begin{cases} a_2 \sum_{i=1}^m x_i^2 n_{x_i} + a_1 \sum_{i=1}^m x_i n_{x_i} + a_0 \sum_{i=1}^m n_{x_i} = \sum_{i=1}^m \bar{y}_{x_i} n_{x_i}, \\ a_2 \sum_{i=1}^m x_i^3 n_{x_i} + a_1 \sum_{i=1}^m x_i^2 n_{x_i} + a_0 \sum_{i=1}^m x_i n_{x_i} = \sum_{i=1}^m x_i \bar{y}_{x_i} n_{x_i}, \\ a_2 \sum_{i=1}^m x_i^4 n_{x_i} + a_1 \sum_{i=1}^m x_i^3 n_{x_i} + a_0 \sum_{i=1}^m x_i^2 n_{x_i} = \sum_{i=1}^m x_i^2 \bar{y}_{x_i} n_{x_i}. \end{cases}$$

Приклад 1. Дано кореляційну таблицю

X	Y					n_x
	2	3	5	7	8	
12	5	–	–	2	6	13
14	1	8	–	12	3	24
17	–	6	11	10	1	28
18	–	4	15	8	–	27
20	–	–	7	1	–	8
n_y	6	18	33	33	10	$n = 100$

Знайти вибіркове рівняння параболічної регресії Y на X.

Розв'язання. Складемо розрахункову таблицю для обчислення коефіцієнтів вищенаведеної системи рівнянь.

X	n_{x_i}	\bar{y}_{x_i}	$n_{x_i}x_i$	$n_{x_i}x_i^2$	$n_{x_i}x_i^3$	$n_{x_i}x_i^4$	$n_{x_i}\bar{y}_{x_i}$	$n_{x_i}\bar{y}_{x_i}x_i$	$n_{x_i}\bar{y}_{x_i}x_i^2$
12	13	5,538	156	1872	22464	269568	72	864	10368
14	24	5,583	336	4704	65856	921984	134	1876	26264
17	28	5,393	476	8092	137564	2338588	151	2567	43639
18	27	5,296	486	8748	157464	2834352	143	2574	46332
20	8	5,25	160	3200	64000	1280000	42	840	16800
Σ	100		1614	26616	447348	7644492	542	8721	143403

Підставивши числа з останнього рядка таблиці у систему рівнянь, маємо систему рівнянь відносно невідомих коефіцієнтів a_0, a_1, a_2 :

$$\begin{cases} 26616a_2 + 1614a_1 + 100a_0 = 542, \\ 447348a_2 + 26616a_1 + 1614a_0 = 8721, \\ 7644492a_2 + 447348a_1 + 26616a_0 = 143403. \end{cases}$$

Розв'язавши одержану систему рівнянь, маємо:

$$a_2 \approx -0,00424; \quad a_1 \approx 0,08868; \quad a_0 \approx 5,1434.$$

Отже, рівняння параболічної регресії $\bar{y}_x = a_2x^2 + a_1x + a_0$ має вигляд:

$$y = -0,0434x^2 + 0,08868x + 5,1434.$$