МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ЧЕРКАСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ БОГДАНА ХМЕЛЬНИЦЬКОГ

ФАКУЛЬТЕТ ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ,

ІНТЕЛЕКТУАЛЬНИХ ТА УПРАВЛЯЮЧИХ СИСТЕМ

Кафедра інформаційних технологій

**ІНДИВІДУАЛЬНЕ ЗАВДАННЯ**

з дисципліни

«Теорія ймовірностей та математична статистика»

на тему: «Статистична обробка результатів експерименту»

Студента 2 курсу групи КН-22

напряму підготовки Компютерні

науки

Бавченко В.О.

Керівник канд. техн. наук, доцент

Косенюк Г.В.

Оцінка:

за універститетською шкалою \_\_\_\_\_\_\_\_\_\_\_\_\_

за шкалою ECTS \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

за національною шкалою \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

«\_\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_ 201\_\_ р.

Члени комісії \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(підпис) (прізвище та ініціали)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(підпис) (прізвище та ініціали)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(підпис) (прізвище та ініціали

Черкаси – 2023

# Зміст

[Зміст 4](#_Toc120641792)

[1. Побудова гістограм частот 5](#_Toc120641793)

[2. Побудова нормальних кривих 8](#_Toc120641794)

[3. Перевірка гіпотези про нормальний розподіл генеральних сукупностей 11](#_Toc120641795)

[4. Знаходження точкових оцінок математичних сподівань і дисперсій генеральних сукупностей 14](#_Toc120641796)

[6. Перевірка гіпотез про рівність нулю математичних сподівань генеральних сукупностей 16](#_Toc120641797)

[7. Перевірка гіпотези про рівність дисперсій генеральних сукупностей 18](#_Toc120641798)

[9. Представлення теоритичних моделей генеральних сукупностей 19](#_Toc120641799)

[10. Висновки 22](#_Toc120641800)

[Список використаних джерел 24](#_Toc120641801)

# 1. Побудова гістограм частот

**Завдання:** Побудувати гістограми частот.

Знайдемо максимальне і мінімальне значення.

Визначимо розмах варіації R із формули:

*Rx=XmaxXmin* =2,79+2,64=5,43,

*Ry=YmaxYmin=*2,36+4,43=6,79.

Тепер потрібно вибрати кількість часткових інтервалів N. Для цього застосовується формулу Стреджеса:

N=1+3,322 *lg(n)*,

де *n* – число варіант вибіркової сукупності (в даному випадку, n = 50). Для визначення довжини часткового інтервалу використовуємо формулу:

Побудуємо спершу гістограму для X. Для початку першого інтервалу візьмемо наступне значення:

Xmin–0,5h=-2,64–0,5·0,817=-3,0485.

Таблиця 1

Дані для гістограми частот вибірки X

|  |  |  |  |
| --- | --- | --- | --- |
| Номер  інтервалу, *i* | Частковий  інтервал,  *xi – xi+1* | Сума частот  варіант  інтервалу, *ni* | Густина  частоти, *ni/h* |
| 1 | -3,0485 – -2,2315 | 2 | 2,44798 |
| 2 | -2,2315 – -1,4145 | 2 | 2,44798 |
| 3 | -1,4145 – -0,5975 | 10 | 12,2399 |
| 4 | -0,5975 – 0,2195 | 12 | 14,68788 |
| 5 | 0,2195 – 1,0365 | 11 | 13,46389 |
| 6 | 1,0365 – 1,8535 | 9 | 11,01591 |
| 7 | 1,8535 – 2,6705 | 3 | 3,671971 |
| 8 | 2,6705 – 3,3875 | 1 | 1,22399 |

Рис. 1. Гістограма частот вибірки X

Зовнішній вигляд гістограми свідчить про те, що величина з досить високою імовірністю може мати нормальний розподіл.

Тепер побудуємо гістограму для Y. Для початку першого інтервалу візьмемо наступне значення:

Ymin – 0,5h = -4,43 – 0,5·1,022= -4,941.

Таблиця 2

Дані для гістограми частот вибірки Y

|  |  |  |  |
| --- | --- | --- | --- |
| Номер  інтервалу, *i* | Частковий  інтервал,  *yi – yi+1* | Сума частот  варіант  інтервалу, *ni* | Густина  частоти, *ni/h* |
| 1 | -4,941 – -3,919 | 1 | 0,978474 |
| 2 | -3,919 – -2,897 | 2 | 1,956947 |
| 3 | -2,897 – -1,875 | 3 | 2,935421 |
| 4 | -1,875 – -0,853 | 6 | 5,870841 |
| 5 | -0,853 – 0,169 | 21 | 20,54795 |
| 6 | 0,169 – 1,191 | 11 | 10,76321 |
| 7 | 1,191 – 2,213 | 5 | 4,892368 |
| 8 | 2,213– 3,235 | 1 | 0,978474 |

Рис. 2. Гістограма частот вибірки Y

***Висновок:*** Зовнішній вигляд гістограми не дуже нагадує нормальний розподіл. Це може пояснюватися як недостатнім обсягом вибірки або невдалим способом групування числових значень ознаки, так і тим, що насправді величина має інший розподіл.

# 2. Побудова нормальних кривих

**Завдання:** За вибірками з генеральних сукупностей X і Y побудувати нормальні криві.

**Для вибірки X:**

Спочатку знаходимо середнє квадратичне відхилення:

Далі, обчислюємо вирівнюючі частоти і заносимо результати розрахунків у таблицю.

Таблиця 3

Таблиця для обчислення вирівнюючих частот X

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| -2,64 | 2 | -2,64 | -2,1 | 0,044 | 1,445 |
| -1,823 | 2 | -1,82 | -1,46 | 0,1374 | 4,51 |
| -1,006 | 10 | -1,01 | -0,8 | 0,2897 | 9,514 |
| -0,189 | 12 | -0,19 | -0,15 | 0,2059 | 6,762 |
| 0,628 | 11 | 0,63 | 0,5 | 0,3521 | 11,56 |
| 1,445 | 9 | 1,45 | 1,16 | 0,2036 | 6,687 |
| 2,262 | 3 | 2,26 | 1,82 | 0,0761 | 2,499 |
| 3,029 | 1 | 3,03 | 2,44 | 0,203 | 6,667 |
|  |  |  |  |  | 49,644 |

Будуємо нормальні криві Х:

Рис. 3. Нормальна крива (Ряд2) та полігон частот (Ряд1) вибірки X

Відхилення значень від нормальної кривої несуттєві.

**Для вибірки Y:**

Переходимо від інтервального ряду до дискретного:

Знаходимо середнє квадратичне відхилення:

Таблиця 4

Таблиця для обчислення вирівнюючих частот Y

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| -4,43 | 1 | -4,43 | -3,12 | 0,0031 | 0,11 |
| -3,408 | 2 | -3,41 | -2,4 | 0,0224 | 0,81 |
| -2,386 | 3 | -2,39 | -1,68 | 0,0973 | 3,51 |
| -1,364 | 6 | -1,36 | -0,96 | 0,2516 | 9,07 |
| -0,342 | 21 | -0,34 | -0,24 | 0,3876 | 13,98 |
| 0,68 | 11 | 0,68 | 0,48 | 0,3555 | 12,81 |
| 1,702 | 5 | 1,70 | 1,2 | 0,1942 | 6,998 |
| 2,724 | 1 | 2,72 | 1,92 | 0,0632 | 2,28 |
|  |  |  |  |  | 49,568 |

Будуємо нормальні криві Y:

Рис. 4. Нормальна крива (Ряд2) та полігон частот (Ряд1) вибірки Y

***Висновок:*** Хоча в цілому криві схожі, але деякі значення розходяться досить сильно. Це може пояснюватися як недостатнім обсягом вибірки або невдалим способом групування числових значень ознаки, так і тим, що теоретичні частоти обчислені виходячи із неправильної гіпотези про нормальний розподіл генеральної сукупності.

# 3. Перевірка гіпотези про нормальний розподіл генеральних сукупностей

**Завдання:** Перевірити гіпотези про нормальний розподіл генеральних сукупностей X та Y, використовуючи критерій погодженості Пірсона.

У якості критерію перевірки нульової гіпотези приймемо величину:

,

де - теоретичні частоти, а - емпіричні, і порівняти її з

де – рівень значущості, k=s–3 – число ступенів вільності, де *s* – кількість часткових інтервалів. Якщо – нульову гіпотезу відхиляють.

Проводимо перевірку гіпотези про нормальний розподіл для вибірки X.

Таблиця 5

Розрахунок емпіричного критерію для X

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 2 | 1,445 | 0,555 | 0,308025 | 0,213166 | 4 | 2,768166 |
| 2 | 4,51 | -2,51 | 6,3001 | 1,396918 | 4 | 0,886918 |
| 10 | 9,514 | 0,486 | 0,236196 | 0,024826 | 100 | 10,51083 |
| 12 | 6,762 | 5,238 | 27,43664 | 4,057475 | 144 | 21,29547 |
| 11 | 11,56 | -0,56 | 0,3136 | 0,027128 | 121 | 10,46713 |
| 9 | 6,687 | 2,313 | 5,349969 | 0,800055 | 81 | 12,11306 |
| 3 | 2,499 | 0,501 | 0,251001 | 0,100441 | 9 | 3,601441 |
| 1 | 6,667 | -5,667 | 32,11489 | 4,816993 | 1 | 0,149993 |
|  |  |  |  |  |  | 61,793 |

Для перевірки правильності обчислюємо значення критерію за другою формулою:

Відповіді співпали майже повністю, отже, розрахунки правильні.

За рівень значущості приймемо 0,05. Число груп вибірки (число різних варіант) s=8, тоді число ступенів вільності *k=*8*–*3*=*5. Значення критичної точки:

Оскільки , то нульову гіпотезу відхиляємо. Те, що гіпотеза не підтвердилася, може свідчити як про те, що насправді розподіл не нормальний, так і про те, що вибірка надто мала або невдала (особливо зважаючи на те, що гістограма і полігон частот давали результати, досить сильно схожі на нормальний розподіл).

Проводимо перевірку гіпотези про нормальний розподіл для вибірки Y.

Таблиця 6

Розрахунок емпіричного критерію для Y

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 1 | 0,09 | 0,91 | 0,8281 | 9,201111 | 1 | 11,11111 |
| 2 | 0,715 | 1,285 | 1,651225 | 2,309406 | 4 | 5,594406 |
| 3 | 3,35 | -0,35 | 0,1225 | 0,036567 | 9 | 2,686567 |
| 6 | 9,1 | -3,1 | 9,61 | 1,056044 | 36 | 3,956044 |
| 21 | 14,26 | 6,74 | 45,4276 | 3,185666 | 441 | 30,92567 |
| 11 | 13,05 | -2,05 | 4,2025 | 0,322031 | 121 | 9,272031 |
| 5 | 6,9 | -1,9 | 3,61 | 0,523188 | 25 | 3,623188 |
| 1 | 2,113 | -1,113 | 1,238769 | 0,586261 | 1 | 0,473261 |
|  |  |  |  |  |  | 67,64227 |

Для перевірки правильності обчислюємо значення критерію за другою формулою:

Відповіді повністю не співпали, але приблизно співмірні.

За рівень значущості приймемо 0,05. Число груп вибірки (число різних варіант) s=8, тоді число ступенів вільності *k=*8*–*3*=*5. Значення критичної точки: .

Оскільки , то нульову гіпотезу відхиляємо. Те, що гіпотеза не підтвердилася, може свідчити як про те, що насправді розподіл не нормальний, так і про те, що вибірка надто мала або невдала (особливо зважаючи на те, що гістограма і полігон частот давали результати, досить сильно схожі на нормальний розподіл).

# 4. Знаходження точкових оцінок математичних сподівань і дисперсій генеральних сукупностей

**Завдання:** Знайти оцінки математичних сподівань і дисперсій генеральних сукупностей методом найбільшої правдоподібності. Для спрощення розрахунків використати метод добутків.

Для розрахунків доцільно скласти розрахункову таблицю. За фальшивий нуль (С) беремо значення 0,628.

Таблиця 7

Таблиця для розрахунку точкових оцінок параметрів вибірки X

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| -2,64 | 2 | -4 | 8 | 32 | 18 |
| -1,823 | 2 | -3 | -6 | 18 | 8 |
| -1,006 | 10 | -2 | -20 | 40 | 10 |
| -0,189 | 12 | -1 | -12 | 12 | 0 |
| 0,628 | 11 | 0 | 0 | 0 | 11 |
| 1,445 | 9 | 1 | 9 | 9 | 36 |
| 2,262 | 3 | 2 | 6 | 12 | 27 |
| 3,029 | 1 | 2,9388 | 2,9388 | 8,636548 | 15,51415 |
|  | Σ=n=50 |  | Σ=  -28,0612 | Σ=131,6365 | Σ=125,5141 |

Тепер треба перевірити правильність розрахунків.

Розрахунки проведені правильно. Тепер обчислюємо умовні моменти першого і другого порядків:

Далі обчислюємо вибіркові середню і дисперсію за формулами:

Тепер побудуємо таблицю для другої вибірки. За фальшивий нуль (С) беремо значення 0,2545.

Таблиця 8

Таблиця для розрахунку точкових оцінок параметрів вибірки Y

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| -4,43 | 1 | -4 | -4 | 16 | 9 |
| -3,408 | 2 | -3 | -6 | 18 | 8 |
| -2,386 | 3 | -2 | -6 | 12 | 3 |
| -1,364 | 6 | -1 | -6 | 6 | 0 |
| -0,342 | 21 | 0 | 0 | 0 | 21 |
| 0,68 | 11 | 1 | 11 | 11 | 44 |
| 1,702 | 5 | 2 | 10 | 20 | 45 |
| 2,724 | 1 | 3 | 3 | 9 | 16 |
|  | Σ=n=50 |  | Σ=2 | Σ=92 | Σ=146 |

Тепер треба перевірити правильність розрахунків.

Розрахунки проведені правильно. Тепер обчислюємо умовні моменти першого і другого порядків:

Далі обчислюємо вибіркові середню і дисперсію за формулами:

***Висновок:*** знайдені оцінки генеральних середніх не вказують на відсутність систематичних похибок, адже не дорівнюють нулю. Це може бути пов’язано як з наявністю похибок, так і з невеликим обсягом вибірки. Дисперсії вибірок хоч і не сильно, але відрізняються, хоча мали б бути однаковими, адже умови однакові. Це теж пояснюється невеликим обсягом вибірки.

# 6. Перевірка гіпотез про рівність нулю математичних сподівань генеральних сукупностей

**Завдання:** Перевірити гіпотезу про рівність нулю математичних сподівань генеральних сукупностей X і Y.

Спочатку потрібно обчислити спостережене значення критерію

і за таблицею критичних точок розподілу Стьюдента, за заданим рівнем значущості *α* і числом ступенів вільності *k*=n–1 знайти двосторонню критичну точку .

Якщо – немає підстав відхилити нульову гіпотезу.

Якщо – нульову гіпотезу відхиляють.

Перевіримо нульову гіпотезу для вибірки X:

За вибіркою обсягом , взятої із нормальної генеральної сукупності, ми знайшли вибіркове середнє . Шукаємо «виправлене» середнє квадратичне:

При рівні значущості перевіряємо нульову гіпотезу при конкуруючій гіпотезі .

Обчислимо спостережне значення критерію

Конкуруюча гіпотеза має вигляд , тому критична область двостороння.

За таблицею критичних точок розподілу Стьюдента, за рівнем значущості a=0,05 і за числом ступенів вільності знаходимо критичну точку .

Оскільки – немає підстав відхилити нульову гіпотезу, тобто вибіркове середнє несуттєво відрізняється від гіпотетичної генеральної середньої, тобто

Перевіримо нульову гіпотезу для вибірки Y:

За вибіркою обсягом , взятої із нормальної генеральної сукупності, ми знайшли вибіркове середнє . Шукаємо «виправлене» середнє квадратичне:

При рівні значущості перевіряємо нульову гіпотезу при конкуруючій гіпотезі .

Обчислимо спостережне значення критерію

Конкуруюча гіпотеза має вигляд , тому критична область двостороння.

За таблицею критичних точок розподілу Стьюдента, за рівнем значущості a=0,05 і за числом ступенів вільності знаходимо критичну точку .

Оскільки – немає підстав відхилити нульову гіпотезу, тобто вибіркове середнє несуттєво відрізняється від гіпотетичної генеральної середньої.

***Висновок:*** є підстави вважати, що системи працюють без систематичних похибок керування.

# 7. Перевірка гіпотези про рівність дисперсій генеральних сукупностей

**Завдання:** запропонувати просту гіпотезу про рівність дисперсій генеральних сукупностей *H0:D(X)=D(Y)* при конкуруючій гіпотезі *H1:D(X)≠D(Y)*. Прийняти рівень значущості α = 0,1. Перевірити запропоновану гіпотезу.

Спочатку необхідно знайти «виправлені» вибіркові дисперсії.

Далі, знайдемо емпіричне значення критерію як відношення більшої виправленої дисперсії до меншої:

За умовою, конкуруюча гіпотеза має вигляд , тому критична область двостороння.

За таблицею критичних точок розподілу Фішера-Снедекора, за рівнем значущості вдвічі меншим заданого () і числом ступенів вільності знаходимо критичну точку:

.

***Висновок:*** Оскільки , то немає підстав відхилити нульову гіпотеза про рівність генеральних дисперсій. Іншими словами, вибіркові виправлені дисперсії відрізняються несуттєво. Рівність дисперсій свідчить про те, що системи автоматичного керування ідентичні.

# 9. Представлення теоритичних моделей генеральних сукупностей

**Завдання:** за одержаними результатами обробки даних вибіркових сукупностей для кожної із генеральних сукупностей представити ймовірнісну теоретичну модель. Представимо теоретичні моделі генеральних сукупностей у вигляді графіків, на підставі того, що гіпотези про нормальний розподіл генеральних сукупностей не були відхилені. Отже, обидві наші генеральні сукупності мають нормальний розподіл.

Нормальний розподіл характеризується густиною ймовірності:

Для вибірки Х:

Підставляємо значення ;

Будуємо розрахункову таблицю для вибірки X:

Таблиця 9

|  |  |
| --- | --- |
| *Xi* | *F(xi)* |
| -2,64 | 0,0518861685 |
| -1,823 | 0,1139881811 |
| -1,006 | 0,1915733379 |
| -0,189 | 0,2463073259 |
| 0,628 | 0,2422627332 |
| 1,445 | 0,1822900300 |
| 2,262 | 0,1049315869 |
| 3,029 | 0,0489617715 |

Будуємо гістограму за даними таблиці:

Рис 7.1 Теоретична модель вибірки X

Підставляємо значення ;

Будуємо розрахункову таблицю для вибірки Y:

Таблиця 7.2

|  |  |
| --- | --- |
| *Yi* | *F(Yi)* |
| -3,408 | 0,0579046220 |
| -2,386 | 0,1155847807 |
| -1,364 | 0,1757557388 |
| -0,342 | 0,2035822051 |
| 0,68 | 0,1796352455 |
| 1,702 | 0,1207437689 |
| 2,724 | 0,0618243191 |
| 1,6175 | 0,1260697931 |

Будуємо гістограму за даними таблиці:

Рис. 7.2. Теоретична модель вибірки Y

***Висновок:*** Mи побудували теоретичну модель для вибірок X та Y за обробленими даними про генеральну сукупність.

# 10. Висновки

Виконавши всі необхідні розрахунки, підведемо підсумок.

У ході роботи, ми побудували гістограми частот, визначивши максимальне та мінімальне значення у вибірках, щоб знайти розмах варіації *R* для кожної з них та довжину інтервалів *h*. Побудувавши гістограми частот для кожної з вибірок, ми зробили висновок, що припущення, що генеральні сукупності, представлені вибірками X та Y, розподілені нормально може бути правильним.

Побудували нормальні криві за дослідними даними і виявили, що відхилення від нормальної кривої існують, але вони в цілому не суттєві. Дізналися, яка різниця між емпіричними частотами та вирівнюючими (теоретичними) частотами.

Перевірили гіпотези про нормальний розподіл генеральних сукупностей і виявили, що оскільки , то є підстави відхилити нульову гіпотезу. Отже, генеральні сукупності не мають нормальний розподіл.

Знайшли точкові оцінки математичних сподівань і дисперсій генеральних сукупностей, скориставшись логарифмічною функцією правдоподібності. Обчислили вибіркове середнє та вибіркову дисперсію.

Перевірили гіпотези про рівність нулю генеральних середніх нормальних генеральних сукупностей і виявили, що оскільки , то немає підстав відхилити нульову гіпотезу, тобто вибіркове середнє незначуще відрізняється від гіпотетичної генеральної середньої.

Перевірили гіпотези про рівність дисперсій генеральних сукупностей і виявили, що оскільки Fемп < Fкр, то нульова гіпотеза про рівність генеральних дисперсій не відхиляється. Іншими словами, вибіркові виправлені дисперсії відрізняються несуттєво.

Тож, дивлячись на результати наших обчислень, та на те, що математичне сподівання рівне нулю, можна зробити висновок, що системи автоматичного керування працюють якісно, а відхилення є наслідками випадкових факторів

# Список використаних джерел

1. Косенюк Г. В. Теорія ймовірностей, імовірнісні процеси та математична статистика: Методичні рекомендації щодо виконання розрахунково – графічної роботи – Черкаси: Видавництво ЧНУ 2005. – 79 с.

2. Сеньо П.С. Теорія ймовірностей та математична статистика: Підручник. – Київ: Центр навчальної літератури, 2004. – 448 с.