­­МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

­­



Лабораторна робота №1

«Статистичні методи в проектуванні сос.

Попередня обробка експериментальних даних.

Дисперсійний аналіз»

з курсу «Основи автоматизованого проектування складних об'єктів і систем ІІ»

для студентів базового напрямку 6.08.04 "Комп’ютерні науки"

(заочна форма навчання)

Варіант 11

Виконав студент гр. КНз-41:

Чалий Михайло

Перевірив:

Матвійків О.М

­

Львів 2015

## Мета роботи

* вивчити основні методи статистичного аналізу
* здійснити попередню обробку експериментальних данних
* визначити основні характеристики факторів та контрольованої величини
* побудувати гістограму для контрольованої величини.

## Теоретичні відомості

Математично задача формулюється наступним чином: необхідно отримати модель функціональної залежності



де *y* – параметр процесу, котрий необхідно оптимізувати;

*(х1 , х2 , ... , хn)* - незалежні змінні, які можна змінювати в процесі експериментів.

Змінні *(х1 , х2 , ... , хn)* називають **факторами**, а координатний простір з координатами *х1 , х2 , ... , хn* – **факторним простором**.

**Генеральна сукупність** – сукупність усіх можливих значень експериментів, які могли б бути при даному комплексі умов. Результати обмеженого ряду спостережень розглядаються як **вибірка** з даної генеральної сукупності.

Попередня обробка результатів спостережень необхідна для того, щоб в подальшому з найбільшою ефективністю, а головне - коректно, використовувати для побудови емпіричних залежностей статистичні методи. Зміст попередньої обробки, в основному, полягає в тому, щоб відсіяти грубі похибки, які можуть зявлятись в експериментальних даних.

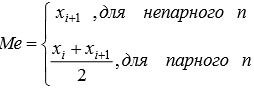
Для обробки наших експерементальних даних ми використаємо **метод Ст’юдента**, який полягає у тому, що розраховується τ за наступною формулою:

,

де Xi – крайне найбільший чи найменший елемент вибірки, по якій підраховувалися математичні сподівання ; τ1-p – табличне значення статистики, τ обчислене при довірчій ймовірності q=1-p.

Якщо умова виконується, то спостереження не відсіюють; якщо не виконується його необхідно відсіяти.

Визначення характеристик є основним для проведення аналізу. Важливими характеристиками є:

* Мінімальне та максимальне значення;
* Середнє значення ;
* Дисперсія ;
* Середньо-квадратичне відхилення ;
* Медіана ;
* Мода.

При оцінці результатів вимірювань важливим моментом обробки даних є перевірка гіпотези про те, що розподіл контрольованої випадкової величини Y наближається до теоретичної кривої нормального розподілу. Далі ця гіпотеза перевіряється з допомогою обчислення вирівнюючих частот вибраного розподілу і одного з критеріїв узгодження.

Для не дуже великих виборок (n<120) перевірка гіпотези на нормальність розподілу є простою і для її реалізації необхідно обчислити середнє абсолютне відхилення САВ:



Для вибірки, яка має приблизно нормальний закон розподілу, буде справедливим наступний вираз:



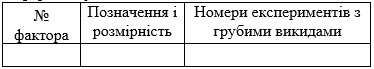
## Завдання

Вивчити основні поняття статистичного аналізу та принципи їх застосування для експериментальних даних.

2) Побудувати матрицю експериментів згідно індивідуального завдання.

3) Дослідити отриману вибірку:

a) Провести попередню обробку результатів спостережень і відсіяти грубі похибки, якщо вони існують. Результати оформити у вигляді таблички:



b) Обчислити min, max та середні значення фактора (Хі) та контрольованої величини (Y). Результати оформити у вигляді таблички:





c) Обчислити дисперсію, середньо-квадратичне відхилення, медіану та моду для фактора (Хі) та контрольованої величини (Y). Результати оформити у вигляді таблички:



d) перевірити гіпотезу про нормальність розподілу контрольованої величини (Y) і, при необхідності, перетворити розподіл в нормальний.

e) Побудувати гістограму розподілу контрольованої величини (Y), розбивши діапазон значень на 5-10 рівних інтервалів.

4) Оформити та захистити звіт.

## Індивідуальне завдання

Варіант – 11(1)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Y** | **X1** | **X2** | **X3** | **X4** |
| 12.1597 | 0.8250 | 0.6710 | 4.7140 | 236.8771 |
| 23.0246 | 0.9240 | 0.7200 | 5.1280 | 246.6777 |
| 20.3394 | 0.7770 | 0.6930 | 4.6870 | 266.0465 |
| 21.9445 | 0.7410 | 0.5225 | 2.1980 | 293.9535 |
| 18.8692 | 0.6958 | 0.5684 | 3.6552 | 196.6777 |
| 26.8312 | 0.8505 | 0.6195 | 4.2145 | 203.2558 |
| 17.2552 | 0.7931 | 0.6180 | 4.4852 | 210.4485 |
| 37.5470 | 0.8806 | 0.8092 | 5.3786 | 257.2093 |
| 19.2391 | 0.6497 | 0.5429 | 3.5495 | 210.8970 |
| 9.4840 | 0.6900 | 0.5336 | 3.7688 | 249.0698 |
| 15.3467 | 0.7008 | 0.5664 | 4.0504 | 236.8106 |
| 5.8934 | 0.8268 | 0.5830 | 4.7682 | 190.0000 |
| 28.0439 | 0.8307 | 0.7956 | 5.3383 | 210.6312 |
| 21.0815 | 0.6936 | 0.6528 | 4.3348 | 196.0133 |
| 20.7983 | 0.6417 | 0.5301 | 3.6385 | 204.0698 |
| 17.8155 | 0.6790 | 0.5723 | 4.0008 | 177.3754 |
| 18.4680 | 0.7614 | 0.5358 | 4.0214 | 196.4286 |
| 14.6182 | 0.7425 | 0.6237 | 5.0836 | 242.8073 |
| 13.8825 | 0.7200 | 0.6400 | 4.5300 | 265.5150 |
| 32.6991 | 0.8775 | 0.7488 | 4.8118 | 302.3256 |
| 9.0713 | 0.7738 | 0.6360 | 4.3654 | 257.9568 |
| 18.3992 | 0.8137 | 0.6283 | 4.1041 | 245.6811 |
| 7.7686 | 0.8480 | 0.6148 | 4.6834 | 212.4917 |
| 8.4340 | 0.8118 | 0.5841 | 4.1926 | 255.9468 |
| 4.7122 | 0.8300 | 0.5800 | 3.9400 | 210.7973 |
| 4.3110 | 0.7526 | 0.5936 | 4.7046 | 205.0000 |
| 7.7139 | 0.8400 | 0.6720 | 4.6520 | 219.0698 |
| 14.2807 | 0.9605 | 0.6441 | 4.6302 | 275.1163 |
| 3.3644 | 0.7371 | 0.5551 | 3.5495 | 237.6412 |
| 8.4438 | 0.8170 | 0.6175 | 4.2405 | 207.1429 |
| 1.6951 | 0.7007 | 0.5733 | 4.0591 | 179.9003 |
| 8.9907 | 0.8664 | 0.6726 | 5.0176 | 202.7741 |
| 0.9159 | 0.6336 | 0.5104 | 3.6272 | 207.7243 |
| 14.6174 | 0.7268 | 0.5428 | 3.6768 | 229.2359 |
| 3.3561 | 0.8056 | 0.6042 | 2.4680 | 186.0465 |
| 2.6947 | 0.7597 | 0.6527 | 4.4648 | 210.7973 |
| 14.1335 | 0.9102 | 0.6882 | 5.0944 | 218.5382 |
| 2.7850 | 0.6900 | 0.6400 | 4.3900 | 217.9402 |
| 3.2105 | 0.7140 | 0.5916 | 4.3348 | 248.5880 |
| 9.5913 | 0.8904 | 0.6572 | 4.2064 | 268.6047 |
| 10.4054 | 0.6699 | 0.5742 | 3.4237 | 246.0465 |
| 15.0554 | 0.6825 | 0.6097 | 3.4767 | 204.3189 |
| 11.5569 | 0.7524 | 0.5841 | 3.9847 | 206.5116 |
| 11.9127 | 0.6674 | 0.5640 | 3.8334 | 218.5382 |
| 7.1533 | 0.8148 | 0.6402 | 4.2918 | 183.5714 |
| 16.3610 | 0.4590 | 0.4950 | 3.8650 | 215.9468 |
| 9.5823 | 1.0062 | 0.6903 | 4.6597 | 230.0000 |
| 19.2406 | 0.8512 | 0.7392 | 4.8872 | 228.0399 |
| 4.9806 | 0.7380 | 0.5490 | 4.0000 | 251.8605 |
| 6.8708 | 0.8000 | 0.6200 | 4.0500 | 209.9668 |

## Реалізація

Лістінг 1. l1.r

D = read.csv('data.csv')

cat('Table 0: Summary', '\n')

summary(D)

cat('Table 1: Outliers', '\n')

outliers <- function (d) {

abs(d - mean(d)) / sd(d) > 2.086

}

for (v in names(D)) {

cat(v, ':', which(outliers(D[,v])), '\n')

}

D = subset(D, !outliers(D$X1) & !outliers(D$X2) & !outliers(D$X3) & !outliers(D$X4))

cat('Table 2: Min, Max, Mean', '\n')

# summary(D)

cat('-', 'Min', 'Max', 'Mean', '\n')

for (v in names(D)) {

cat(v, ':', min(D[,v]), max(D[,v]), mean(D[,v]), '\n')

}

cat('Table 3: Var, SD, Median, Mode', '\n')

mode <- function(x) {

ux <- unique(x)

ux[which.max(tabulate(match(x, ux)))]

}

cat('-', 'Var', 'SD', 'Median', 'Mode', '\n')

for (v in names(D)) {

cat(v, ':', var(D[,v]), sd(D[,v]), median(D[,v]), mode(D[,v]), '\n')

}

cat('Table 4: Normal distribution', '\n')

cat('Y p-value:', shapiro.test(D$Y)$p.value, '\n')

hist(D$Y)e

## Результат

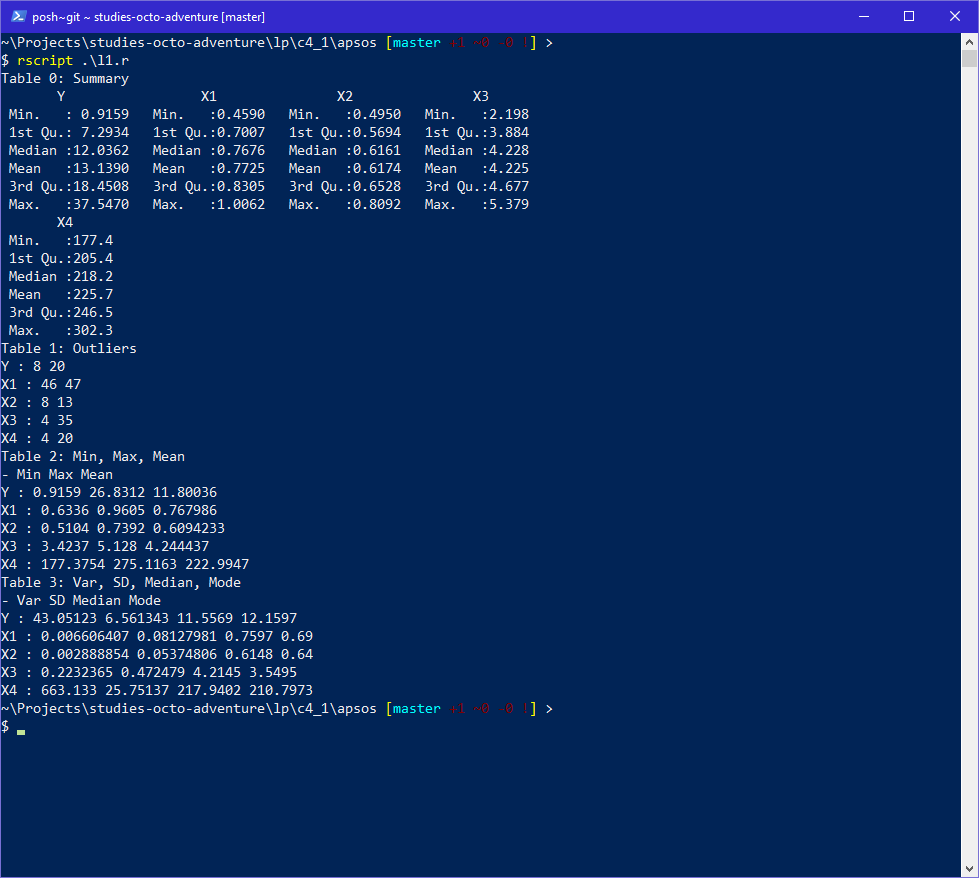


Рис. . Результат виконання

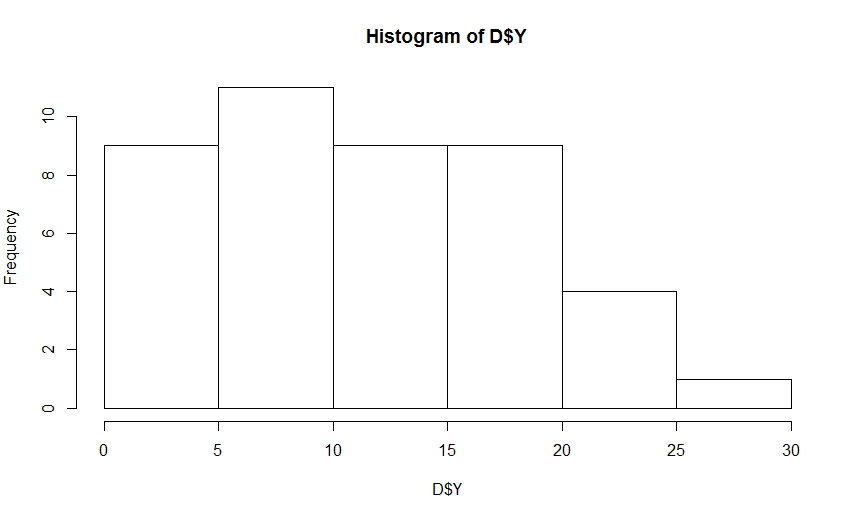


Рис. . Гістограма

## Висновки

Ознайомився під час виконання лабораторної роботи з попередньою обробкою статистичних даних. Розраховано характеристики, які описують експериментальні величини, а саме: мінімальне, максимальне та середнє значення, дисперсію, середньоквадратичне відхилення, моду та медіану. Гіпотеза, щодо подібності розподілу величини Y до нормального, підтвердилася. Побудовано гістограму розподілу контрольованої величини. Отримані результати збігаються з очікуваними.