НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ

імені Ігоря Сікорського»

Кафедра прикладної математики

Звіт

із лабораторної роботи №*3*

із дисципліни «Аналіз даних»

|  |  |  |
| --- | --- | --- |
| Виконав: | | Керівник: |
| студент групи КМ-63 | | *Пашко Анатолій Олексійович* |
| *Артеменко Я.К.* |  | |

Київ — 2019

**Тема:** Алгоритми однофакторного і двохфакторного дисперсійного аналізу.

**Мета роботи:** здобути практичні навички проведення і аналізу даних

однофакторного та двохфакторного дисперсійного аналізу.

**Завдання:**

1. Провести дисперсійний аналіз даних, відповідно до варіанту, при довірчій ймовірності а=0.95.
2. Провести двох факторний дисперсійний аналіз даних, відповідно до варіанту, при довірчій ймовірності а=0.95.
3. За результатами оформити звіт. В звіті відобразити особливості реалізації алгоритмів в середовищі R Studio.

**Код програми та результати його виконання:**

**1) Однофакторний аналіз(при варіанті = 1):**

> prodd <- function(x) {

+ sum = 0

+ for (i in 1: length(x)) {

+ sum = sum + x[i]\*x[i]

+ }

+ return (sum)

+ }

> N = 1000

> GTEOR = 0.33

> k = 5

> variant = 1

> FTEOR = 2.37;

> cat("f = ", FTEOR, "\n")

f = 2.37

> alfa = 0.95;

> koef = c(1, 0.5,0.8,1.4,2)

> S = c(0,0,0,0,0)

> mas = matrix(0, nrow = 5, ncol = N)

> for (i in 1:5) {

+ for (j in 1:N) {

+ mas[i,j] = variant\*koef[i] + runif(1)

+ }

+ }

> maxS = 0

> for (i in 1:5) {

+ S[i] = var(mas[i,])

+ maxS = max(S[i], maxS);

+ }

>

> G = max(S) / (S[5] + S[1] + S[2] + S[3] + S[4]);

> cat("g = ", G, "\n");

g = 0.2035133

>

> if (GTEOR < G) {

+ cat("END")

+ } else {

+ sum1 = 0

+ sum2 = 0

+ sum3 = sum(mas)\*\*2

+ for (i in 1:k) {

+ sum1 = sum1 + prodd(mas[i,])

+ sum2 = sum2 + sum(mas[i,])\*\*2

+ }

+ sum2 = sum2 / N

+ S0\_2 = 1/(k\*(N-1))\*(sum1 - sum2)

+ S2 = 1/(k\*N - 1) \* (sum1 - 1/(k\*N)\*sum3)

+

+ x\_\_ = 0

+

+ for (i in 1:k) {

+ x\_\_ = x\_\_ + mean(mas[i,])

+ }

+ x\_\_ = x\_\_/k

+ sum4 = 0

+ for (i in 1:k) {

+ sum4 = sum4 + (mean(mas[i,]) - x\_\_)\*\*2

+ }

+ SA\_2 = N/(k-1)\*sum4

+ cat("SA^2 = " , SA\_2, "\n")

+ cat("S0^2 = " , S0\_2, "\n");

+ cat("SA^2/S0^2 = " , SA\_2 / S0\_2, "\n");

+ if (SA\_2 / S0\_2 > FTEOR) cat("A is good criteria")

+ else cat("A is bad criteria");

+ }

SA^2 = 338.8739

S0^2 = 0.08398875

SA^2/S0^2 = 4034.754

A is good criteria

**2) Двохфакторний аналіз(при варіанті = 1)**

> N = 100

> variant = 1

> k = 5

> m = 4

> FTeora = 3.26;

> FTeorb = 3.49;

> FTeor1 = 1.75;

> mas <- array(0, dim=c(m, k, N))

> ar <- array(0, c(m,k))

> sum = 0;

> koefData = c(1, 3.5, 3.8, 1.4, 2,

+ 1, 2.5, 2.8, 2.4, 3,

+ 1, 1.5, 1.8, 3.4, 4,

+ 1, 0.5, 0.8, 4.4, 5)

> koef<-matrix(koefData, nrow=4, ncol=5,byrow = TRUE)

> array <- matrix(0,nrow = m, ncol = k)

>

> for ( i in 1:m) {

+ for ( j in 1:k) {

+ sum = 0;

+ for (q in 1 : N ) {

+ s <- variant \* koef[i,j] + runif(1)

+ mas[i,j,q] = s

+ sum = sum + mas[i,j,q];

+ }

+ array[i,j] = sum / N;

+ }

+ }

>

>

> sumOfMasPow <- function() {

+ return (sum(mas\*mas))

+ }

>

> sumOfXiPow <- function() {

+ sum = sum(array)

+ return (sum \* sum )

+ }

>

> sumOfXj <- function() {

+ sum = 0;

+ for (i in 1 : m ) {

+ localSum = 0;

+ for (j in 1 : k ) {

+ localSum = localSum + array[i,j];

+ }

+ sum = sum + (localSum \* localSum);

+ }

+ return (sum);

+ }

>

> sumOfXi <- function() {

+ sum = 0;

+ for (j in 1 : k ) {

+ localSum = 0

+ for (i in 1 : m ) {

+ localSum = localSum + array[i,j];

+ }

+ sum = sum + (localSum \* localSum);

+ }

+ return (sum);

+ }

>

> sumArrayAvg <- function() {

+ sum = 0

+ for ( i in 1 : m ) {

+ for (j in 1 : k ) {

+ sum = sum + (array[i,j] \* array[i,j]);

+ }

+ }

+ return (sum);

+ }

>

> Q = sumArrayAvg();

> Q2 = sumOfXi() / m;

> Q3 = sumOfXj() / k;

> Q4 = sumOfXiPow() / (m \* k);

> Q5 = sumOfMasPow();

> S0 = (Q + Q4 - Q2 - Q3) / ((k - 1) \* (m - 1));

> Sa = (Q2 - Q4) / (k - 1);

> Sb = (Q3 - Q4) / (m - 1);

> Sab = (Q5 - N \* Q) / (m \* k \* (N - 1));

> cat("Q1 = " , Q, "\n")

Q1 = 195.3557

> cat("Q2 = " , Q2,"\n")

Q2 = 175.4672

> cat("Q3 = " , Q3,"\n")

Q3 = 161.324

> cat("Q4 = " , Q4,"\n")

Q4 = 161.3233

> cat("Q5 = " , Q5,"\n")

Q5 = 19700.83

> cat("S0^2 = " , S0,"\n")

S0^2 = 1.657312

> cat("Sa^2 = " , Sa,"\n")

Sa^2 = 3.535983

> cat("Sb^2 = " , Sb,"\n")

Sb^2 = 0.0002585886

> cat("Sab^2 = " , Sab,"\n");

Sab^2 = 0.08346263

> cat("Sa^2 / S0^2 = " , Sa / S0 ,

+ "\n Sb^2 / S0^2 = " , Sb / S0 ,

+ "\n N\*S0^2/Sab^2 = ", N \* S0 / Sab, "\n")

Sa^2 / S0^2 = 2.133565

Sb^2 / S0^2 = 0.0001560289

N\*S0^2/Sab^2 = 1985.694

>

> cat("Fa = ", FTeora, "\n")

Fa = 3.26

> cat("Fb = ", FTeorb, "\n")

Fb = 3.49

> cat("Fab = ", FTeor1, "\n")

Fab = 1.75

> if (Sa/S0 > FTeora) {

+ cat("A is good criteria\n")

+ } else {cat("A is bad criteria\n") }

A is bad criteria

>

>

> if (Sb / S0 > FTeorb) {

+ cat("B is good criteria\n")

+ } else

+ {cat("B is bad criteria\n")}

B is bad criteria

> if (N \* S0 / Sab > FTeor1) {

+ cat("Both criterias are good\n")

+ } else

+ {cat("Both criterias are bad\n")}

Both criterias are good