НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ

імені Ігоря Сікорського»

Кафедра прикладної математики

Звіт

із лабораторної роботи №*6*

із дисципліни «Аналіз даних»

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

Київ — 2019

**Тема:** Регресійний аналіз

**Мета:** здобути практичні навички проведення регресійного аналізу даних.

**Завдання.** Результати спостереження за деякою функціональною залежністю задані таблицею (відповідно варіанту). Методом найменших квадратів знайти найкращу функціональну залежність: лінійна, поліноміальна(другого та третього порядків). Провести статистичний аналіз отриманих коефіцієнтів для довірчої ймовірності а=0.95.

N – номер студента в списку групи(в даному випадку було обрано 1)

NG – номер групи

**Код програми та результати його виконання:**

|  |
| --- |
| > library(ggplot2)  > library(pracma)  > library(polynom)  > N = 1;  > NG = 63;  > n = 1000;  > FisherConst = 1.03;  > x = array(dim = n)  > y = array(dim = n)  > e = array(dim = n)  > s = array(dim = n)  > dividionOfEAndS = array(n)  > for (i in 1:n) {  + x[i] = i+( runif(1) \* N / NG);  + y[i] = N \* runif(1) \* x[i] + NG \* runif(1) + N;  + }  > b = (mean(x\*y) - mean(x)\*mean(y)) / (mean(x\*\*2) - mean(x)\*\*2)  > a = mean(y) - mean(x)\*b  >  > t = seq(min(x),max(x),0.1)  > z = a + b \* t  > cat("a = ",a,"\n")  a = 30.56824  > cat("b = ",b,"\n")  b = 0.5265452  > plot(x,y,type="l",col="red")  > points(t,z,col="green")  > while (1) {  + e = y - (a + b \* x)  + S\_i\_2 = (sum(e\*\*2) / (n - 2)) \* (1 - (1 / n) - ((x - mean(x))\*\*2) / sum( (x - mean(x))\*\*2))  + R = max((e / sqrt(S\_i\_2)))  + yNew = array(dim = n)  + for (i in 1:n) {  + if (e[i] / sqrt(S\_i\_2[i]) == R) {  + maxIndex = i;  + }  + }  + k = 1;  + for (i in 1:n - 1) {  + if (k != maxIndex) {  + yNew[i] = y[k];  + } else { i = i - 1 }  + k=k+1  + }  + R\_delta = 4.0  + if (R > R\_delta) {  + cat("Вкиди є ", y[maxIndex], "\n")  + y = yNew;  + n = n - 1;  +  + } else {  + cat("Викидів нема", "\n")  + break  + }  +  +  + }  Викидів нема  > S\_2 = sum((y - a - b \* x)\*\*2) / (n - 2)  > S\_2\_x = sum((x - mean(x))\*\*2) / (n - 1)  > S\_beta = sqrt(S\_2) / (sqrt(S\_2\_x) \* sqrt(n - 1))  > t\_delta = 1.645  > if (abs(b) > t\_delta \* S\_beta) {  + cat('b значимое\n')  + } else  + cat('b не значимо\n')  b значимое  > S\_alpha = sqrt(S\_2) \* sqrt((1/n) + (mean(x)\*\*2) / ((n-1) \* S\_2\_x))  > if (abs(a) > t\_delta \* S\_alpha){  + cat('a значимое\n')  + } else  + cat('а не значимое\n')  a значимое  > S\_2\_y = sum((y - mean(y)) \*\* 2) / (n - 1)  > F\_delta = 1.0  > if (S\_2 / S\_2\_y < F\_delta) {  + cat('Модель адекватна')  + } else  + cat('Модель не адекватна')  Модель адекватна  > cat("\n")  > z\_1 = polyfit(x, y, 1)  > z\_2 = polyfit(x, y, 2)  > z\_3 = polyfit(x, y, 3)  > la = polynomial(rev(z\_1))  > Sa = polynomial(rev(z\_2))  > Ca = polynomial(rev(z\_3))  > cat('\n Лінійна апрксимація: ',la)  Лінійна апрксимація: 30.56824 0.5265452> cat('\n Квадратична апроксимація: ',Sa)  Квадратична апроксимація: 24.90811 0.5604372 -3.385759e-05> cat('\n Кубічна апрксимація: ',Ca)  Кубічна апрксимація: 39.35719 0.3876623 0.0003974226 -2.872285e-07>  >  > S\_2\_1 = (sum((y - predict(la,x))\*\*2)) / (n - 2)  > S\_2\_2 = (sum((y - predict(Sa,x))\*\*2)) / (n - 2)  > S\_2\_3 = (sum((y - predict(Ca,x))\*\*2)) / (n - 2)  > cat('\nS\_2 для лінійної апрксимації: ', S\_2\_1)  S\_2 для лінійної апрксимації: 27048.23> cat('\nS\_2 для квадратичної апрксимації: ',S\_2\_2)  S\_2 для квадратичної апрксимації: 27041.85> cat('\nS\_2 для кубічної апрксимації: ',S\_2\_3)  S\_2 для кубічної апрксимації: 27012.33>  >  >  > plot(x,y,col="red")  > points(x, predict(la,x), col="green")  > points(x, predict(Sa,x), col="yellow")  > points(x, predict(Ca,x), col="blue") |
|  |
| |  | | --- | |  | |

