

Laborator 4 și 5

De citit: slide-urile: II Model de regresie liniara simpla.ppt si III Model liniar simplu Excel Tema 1.ppt

Să se genereze niște date pe baza ecuației unei drepte, să se perturbe cu o eroare si apoi să se aproximeze punctele cu o dreapta.

Exemplu:

- Se genereaza nr $x=1:10$
- Se genereaza eroarea folosind functia `rand()`
- Se calculeaza $y = a*x+b+err$
- Se reprezinta graphic x,y
- Se analizeze influenta erorii
- Se calculeaza coeficientul de corelatie

Exemplu Matlab

Sa se implementeze in Matlab o funcție care genereaza $y=a*x+b+err$. Punctele (x,y) se vor reprezenta grafic, împreună cu dreapta din a cărei ecuație au fost generate aceste puncte. Funcția va avea ca și date de intrare: n – numărul de puncta, a și b parametrii dreptei. Sa se calculeze corelatia dintre x si y .

Temă de trimis – se notează

Să se implementeze în:

- Python
- C++

un program care calculeaza parametrii alfa si beta ai dreptei de regresie pentru un set de date citit dintr-un fișier cu extensia .csv. Și mai departe face predicții folosind modelul de regresie liniară simplă.

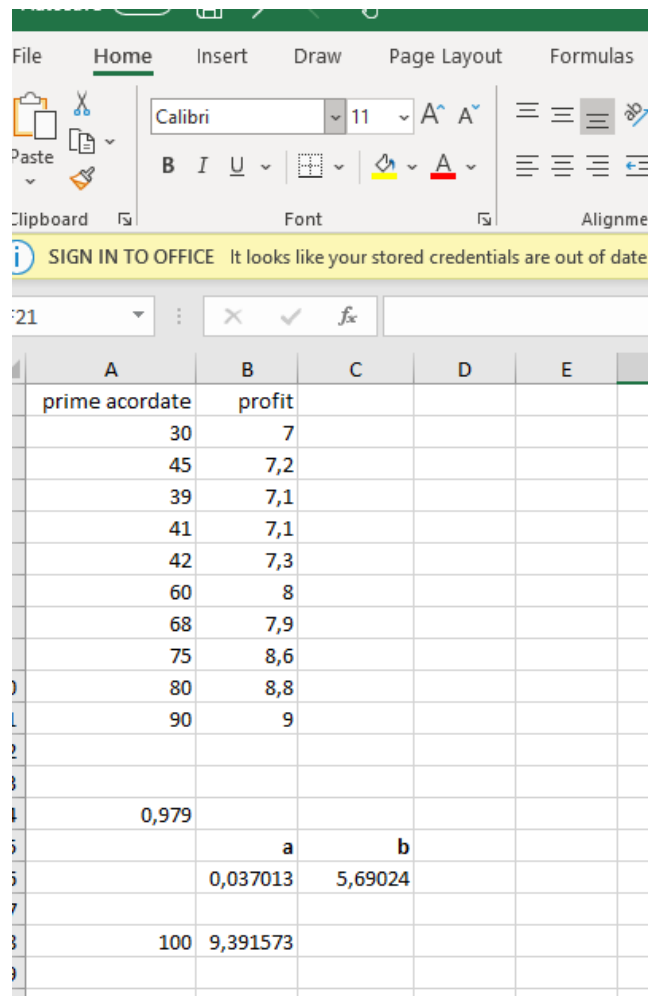
Exemplu:

Pe aceste date:

```
date - Notepad
File Edit Format View Help
prime,profit
30,7
45,7.2
39,7.1
41,7.1
42,7.3
60,8
68,7.9
75,8.6
80,8.8
90,9
```

Se obțin: $a=0.037$ și $b=5.6904$

Si pentru o valoare noua de 100 pentru prima acordata se previzioneaza un profit de 9.39.



A	B	C	D	E
prime acordate	profit			
30	7			
45	7,2			
39	7,1			
41	7,1			
42	7,3			
60	8			
68	7,9			
75	8,6			
80	8,8			
90	9			
0,979				
	a	b		
	0,037013	5,69024		
100	9,391573			

Estimatorii de cele mai mici pătrate ai parametrilor de regresie (valorile reale pentru *alfa* și *beta* nu pot fi determinate exact deoarece, dispunem doar de eșantionul de observații, $(x_i, y_i) i = 1, n$):

$$y = \alpha + \beta x,$$

$$y_i = \alpha + \beta x_i + \varepsilon_i.$$

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x},$$
$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Unde \bar{x} reprezinta media.

Trimiteti:

- codul in Python si C++
- 2 printscreen-uri care sa evidentieze rulara si Python si C++ (returnarea parametrilor regresiei si predictia pentru o noua valoare, atat in Python cat si in C++)

la adresa: adriana.birlutiu@uab.ro. Termen limita: 31.03.20 ora 12:00. După aceasta data tema se penalizeaza cu 4 puncte.

Aceasta tema se va nota, nota va conta în nota finala de la laborator!

Pentru a discuta această temă, ne întâlnim online pe platforma Microsoft Teams unde am creat o echipă pentru laboratorul de Modelare în care am adăugat toți studenții din Anul III, Informatică. Întâlnirile vor fi în zilele și orele în care avem în orar laboratorul de Modelare, adică, marți, miercuri și joi de la ora 14. Intrați așadar pe platforma Microsoft Teams accesând link-ul webmail.uab.ro și vă conectați cu user-ul și parola pe care le-ați primit, și acolo găsiți informații despre laborator și întâlnirile online.

Precizare: Dacă din cauza situației de urgență în care ne aflăm aveți probleme sau cineva apropiat are probleme și nu puteți să respectați termenul inițial de trimitere a temei, vă rog să îmi trimiteți un email cu o scurtă explicație a situației și o să o iau în considerare.