Facebook | 泣きたい時に聴く、切なく悲しい夏のピ | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (unsaved changes) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [1]:
```
import numpy as np
import pandas as pd
from numpy import *
```

In [2]:
```
#Câu 1:

data = np.loadtxt("bostonh.dat")
print(data)
```

```
[[6.3200e-03 1.8000e+01 2.3100e+00 ... 3.9690e+02 4.9800e+00 2.4000e+01]
 [2.7310e-02 0.0000e+00 7.0700e+00 ... 3.9690e+02 9.1400e+00 2.1600e+01]
 [2.7290e-02 0.0000e+00 7.0700e+00 ... 3.9283e+02 4.0300e+00 3.4700e+01]
 ...
 [6.0760e-02 0.0000e+00 1.1930e+01 ... 3.9690e+02 5.6400e+00 2.3900e+01]
 [1.0959e-01 0.0000e+00 1.1930e+01 ... 3.9345e+02 6.4800e+00 2.2000e+01]
 [4.7410e-02 0.0000e+00 1.1930e+01 ... 3.9690e+02 7.8800e+00 1.1900e+01]]
```

In [3]:
```
#Câu 2:

X =np.append(np.ones((len(data),1)),data[:,:13],axis =1)
print(X)
print("-----")
Y = data.transpose()[-1].reshape(len(data),1)
print(Y)
```

```
[[1.0000e+00 6.3200e-03 1.8000e+01 ... 1.5300e+01 3.9690e+02 4.9800e+00]
 [1.0000e+00 2.7310e-02 0.0000e+00 ... 1.7800e+01 3.9690e+02 9.1400e+00]
 [1.0000e+00 2.7290e-02 0.0000e+00 ... 1.7800e+01 3.9283e+02 4.0300e+00]
 ...
 [1.0000e+00 6.0760e-02 0.0000e+00 ... 2.1000e+01 3.9690e+02 5.6400e+00]
 [1.0000e+00 1.0959e-01 0.0000e+00 ... 2.1000e+01 3.9345e+02 6.4800e+00]
 [1.0000e+00 4.7410e-02 0.0000e+00 ... 2.1000e+01 3.9690e+02 7.8800e+00]]
-----
[[24. ]
```

32.9% ENG 7:05 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏のピ | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [4]:

```
#Câu 3:

x_T = X.transpose()
print(x_T)
#Tính vecto B
beta = linalg.inv(x_T.dot(X)).dot(x_T).dot(Y)
print(beta)
```

```
[[1.0000e+00 1.0000e+00 1.0000e+00 ... 1.0000e+00 1.0000e+00 1.0000e+00]
 [6.3200e-03 2.7310e-02 2.7290e-02 ... 6.0760e-02 1.0959e-01 4.7410e-02]
 [1.8000e+01 0.0000e+00 0.0000e+00 ... 0.0000e+00 0.0000e+00 0.0000e+00]
 ...
 [1.5300e+01 1.7800e+01 1.7800e+01 ... 2.1000e+01 2.1000e+01 2.1000e+01]
 [3.9690e+02 3.9690e+02 3.9283e+02 ... 3.9690e+02 3.9345e+02 3.9690e+02]
 [4.9800e+00 9.1400e+00 4.0300e+00 ... 5.6400e+00 6.4800e+00 7.8800e+00]]
[[ 3.64594884e+01]
 [-1.08011358e-01]
 [ 4.64204584e-02]
 [ 2.05586264e-02]
 [ 2.68673382e+00]
 [-1.77666112e+01]
 [ 3.80986521e+00]
 [ 6.92224640e-04]
 [-1.47556685e+00]
 [ 3.06049479e-01]
 [-1.23345939e-02]
 [-9.52747232e-01]
 [ 9.31168327e-03]
 [-5.24758378e-01]]
```

In [5]:

```
#Câu 4:
#Tính C:

C = linalg.inv(x_T.dot(X))
print(C)
```

32.2% ENG 7:06 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏の | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [5]:

```
#Câu 4:
#Tính C:

C = linalg.inv(x_T.dot(X))
print(C)
```

```
[[ 1.15665067e+00 -4.79507297e-04 -7.27912374e-05  7.81007312e-04
  -5.30636011e-03 -4.72633728e-01 -6.72009691e-02  2.52574657e-04
  -1.60858766e-02  4.27955175e-03 -9.32179147e-05 -1.77832398e-02
  -1.77463251e-04 -3.44062107e-03]
 [-4.79507297e-04  4.79667291e-05 -1.74622146e-06  3.33058300e-06
   6.29163676e-05  3.32353579e-04  1.60589501e-05 -6.44104667e-08
   3.40052584e-05 -2.58500572e-05  6.50408184e-08  3.99960618e-06
   4.60269987e-07 -1.13029256e-05]
 [-7.27912374e-05 -1.74622146e-06  8.36861245e-06  3.98676259e-06
  -7.75453928e-06  8.33368670e-05 -3.96939822e-05  9.71684642e-07
  -4.89624081e-05  4.27993705e-06 -5.11685476e-07  2.48176992e-05
  -1.86831132e-08 -1.64268535e-06]
 [ 7.81007312e-04  3.33058300e-06  3.98676259e-06  1.67943163e-04
  -2.42364262e-04 -2.75364026e-03  1.04925166e-04 -3.97527513e-08
   1.14987074e-04  5.04398943e-05 -4.49769287e-06 -4.64149258e-05
   2.43976935e-07 -1.05525682e-05]
 [-5.30636011e-03  6.29163676e-05 -7.75453928e-06 -2.42364262e-04
   3.29658256e-02 -4.90317797e-03 -4.98480956e-04 -2.56627859e-05
   1.20331081e-04 -2.73145702e-04  1.74383419e-05  4.68840845e-04
  -5.27318555e-06  1.15327313e-04]
 [-4.72633728e-01  3.32353579e-04  8.33368670e-05 -2.75364026e-03
  -4.90317797e-03  6.47949909e-01  6.62608200e-03 -5.94124576e-04
   9.53889054e-03 -1.59989671e-03 -4.75854409e-05  7.31483621e-03
   3.41113655e-05 -5.91605940e-04]
 [-6.72009691e-02  1.60589501e-05 -3.96939822e-05  1.04925166e-04
  -4.98480956e-04  6.62608200e-03  7.75657890e-03 -5.06941561e-05
   4.94416767e-04 -1.94649180e-04  5.18444533e-06  3.87223280e-04
   5.22111421e-06  5.03507223e-04]
 [ 2.52574657e-04 -6.44104667e-08  9.71684642e-07 -3.97527513e-08
  -2.56627859e-05 -5.94124576e-04 -5.06941561e-05  7.74933231e-06
   3.40849783e-05  2.95027780e-06 -6.60690922e-08 -6.00107894e-06
```

31.0% ENG 7:08 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏の | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [6]:
```
#Tính y mũ:

y_mu = X.dot(beta)
print(y_mu)
```

```
[[30.00384338]
 [25.02556238]
 [30.56759672]
 [28.60703649]
 [27.94352423]
 [25.25628446]
 [23.00180827]
 [19.53598843]
 [11.52363685]
 [18.92026211]
 [18.99949651]
 [21.58679568]
 [20.90652153]
 [19.55290281]
 [19.28348205]
 [19.29748321]
 [20.52750979]
 [16.91140135]
 [16.17801106]
 [18.40613603]
```

In [7]:
```
#Tính e mũ:

e = Y - y_mu
print(e)
```

```
[[-6.00384338e+00]
 [-3.42556238e+00]
 [ 4.13240328e+00]
 [ 4.79296351e+00]
 [ 8.25647577e+00]
 [ 3.44371554e+00]
```

30.5% ENG 7:09 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏の | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [8]:
```
#Tính sigma bình:
sigma = (transpose(e).dot(e)) / (len(X) - (13+1))
print(sigma)
```

```
[[22.51785483]]
```

In [9]:
```
#Tính SE_beta:

SE_beta = sigma*np.diag(C)
print(SE_beta)
```

```
[[2.60452918e+01 1.08010784e-03 1.88443200e-04 3.78171976e-03
  7.42319676e-01 1.45904420e+01 1.74661518e-01 1.74498340e-04
  3.97821912e-02 4.40185014e-03 1.41416344e-05 1.71156401e-02
  7.21440767e-06 2.57203944e-03]]
```

In [10]:
```
#Câu 5: Xây dựng KTC cho beta:

import scipy.stats as scipy
t = (1 - 0.99)/2
t_value = scipy.t.ppf(t,len(X) - (13+1))
VT = beta - t_value * SE_beta
VP = beta + t_value * SE_beta
print(VT,VP)
```

```
[[ 1.03808938e+02  3.64622814e+01  3.64599757e+01  3.64692674e+01
   3.83790223e+01  7.41883125e+01  3.69111384e+01  3.64599396e+01
   3.65623595e+01  3.64708709e+01  3.64595250e+01  3.65037470e+01
   3.64595070e+01  3.64661393e+01]
 [ 6.72414380e+01 -1.05218351e-01 -1.07524070e-01 -9.82323642e-02
   1.81152257e+00  3.76208127e+01  3.43638680e-01 -1.07560130e-01
  -5.14022526e-03 -9.66287945e-02 -1.07974790e-01 -6.37527279e-02
  -1.07992702e-01 -1.01360427e-01]
 [ 6.73958698e+01  4.92134648e-02  4.69077459e-02  5.61994520e-02
   1.96595439e+00  3.77752445e+01  4.98070496e-01  4.68716865e-02
   1.49291591e-01  5.78030217e-02  4.64570266e-02  9.06790883e-02
```

30.0% ENG 7:09 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏の | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [11]:
```
#Câu 6:

#Tính y mũ của câu 6: với x mũ là hàm tự cho
t = (1 - 0.95)/2
t_value = scipy.t.ppf(t,len(X) - (13+1))
data1 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]
x_mu_nho = np.array(data1)
y_mu_moi = x_mu_nho.dot(beta)
print(y_mu_moi)
```

```
[-59.04026202]
```

In [12]:
```
#Tính epsilon:
import math

canbachai = sigma*(1+transpose(x_mu_nho).dot(C).dot(x_mu_nho))
epsilon = abs(t_value*canbachai)
print(epsilon)

VT1 = float(y_mu_moi - epsilon)
VP1 = float(y_mu_moi + epsilon)
print("Khoảng tin cậy của y là:")
(VT1,VP1)
```

```
[[957.24609699]]
Khoảng tin cậy của y là:
```

Out[12]: (-1016.2863590120626, 898.2058349725091)

In [13]:
```
#Tính các giá trị của bảng ANOVA:

#Tính SST:
sum_y = 0
for i in Y:
    sum_y = sum_y + i**2
SST = transpose(Y).dot(Y) - (1/len(Y))*(sum_y)
```

29.3% ENG 7:10 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏の | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [13]:

```
#Tính các giá trị của bảng ANOVA:

#Tính SST:
sum_y = 0
for i in Y:
    sum_y = sum_y + i**2
SST = transpose(Y).dot(Y) - (1/len(Y))*(sum_y)
print("SST là: ",SST)

#Tính SSR:
SSR = transpose(beta).dot(transpose(X)).dot(Y) - (1/len(Y))*(sum_y)
print("SSR là: ",SSR)

#Tính SSE:
SSE = SST - SSR
print("SSE là: ",SSE)

#Tính trung bình bình phương MSR và MSE:
MSR = (SSR/13)
print("MSR là: ",MSR)

MSE = (SSE/len(X)-(13+1))
print("MSR là: ",MSR)

#Tính F0:
F0 = MSR/MSE
print("F0 là: ",F0)

```

```
SST là:  [[299034.193083]]
SSR là:  [[287955.40850503]]
SSE là:  [[11078.78457797]]
MSR là:  [[22150.41603885]]
MSR là:  [[22150.41603885]]
F0 là:  [[2805.68583784]]
```

In [14]:

```
#Cách lập bảng ANOVA:
```

28.9% ENG 7:11 PM

Facebook | 泣きたい時に聴く、切なく悲しい夏の | TNHien/Xử lý số liệu thống kê/Bài tập | 19110315_TrinhNgocHien_BTTuan4

localhost:8888/notebooks/TNHien/Xử%20lý%20số%20liệu%20thống%20kê/Bài%20tập/BT%20Tuần%204/19110315_TrinhNgocHien_BTTua...

Library Genesis | Electronic library. D... | Nhập môn AI - Bai... | Thư - TRỊNH NGỌC... | Zalo Web

jupyter 19110315_TrinhNgocHien_BTTuan4 Last Checkpoint: 21 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

```
MSR là:  [[22150.41603885]]
MSR là:  [[22150.41603885]]
F0 là:  [[2805.68583784]]
```

In [14]:

```
#Cách lập bảng ANOVA:

import pandas as pd
s = {'Tên đại lượng': pd.Series(['Hồi quy bội','Sai số (hay thặng dư)','Tổng']),
    'Tổng bình phương': pd.Series([SSR,SSE,SST]),
    'Bậc tự do': pd.Series([13,len(X)-(13+1),len(X)]),
    'Trung bình bình phương': pd.Series([MSR,MSE,' ']),
    'F0': pd.Series([F0,' ',' '])}

df = pd.DataFrame(s)
print(df)

```

```
           Tên đại lượng        Tổng bình phương  Bậc tự do  \
0            Hồi quy bội  [[287955.40850503143]]         13
1  Sai số (hay thặng dư)  [[11078.784577972605]]        492
2                  Tổng  [[299034.19308300404]]        506

   Trung bình bình phương                    F0
0  [[22150.416038848572]]  [[2805.685837844506]]
1    [[7.89483118176404]]
2
```

In [ ]:

28.8% ENG 7:11 PM