Jaime Frade

STA5707

HW1: exercise 1.14 and 1.15

1.14

a)

Plot of x2\*x4. Symbol used is '+'.

240 ˆ

‚ + +

‚

‚

‚ +

‚

‚

220 ˆ

‚

‚ +

‚ +

‚

‚ +

‚ +

200 ˆ +

‚

‚ +

x2 ‚ +

‚

‚ +

‚

180 ˆ

‚ +

‚ +

‚ + + + +

‚ + +

‚ + + +

‚ + +

160 ˆ +

‚ + + + + + + +

‚ +++ +

‚ + + ++ + +

‚ ++ + +

‚ + +++ ++ +

‚ ++++ ++

140 ˆ +++++ +++ + + +

‚ + ++++ +

‚ + +

‚ +

‚

‚ +

‚

120 ˆ

Šˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒˆ

160 180 200 220 240 260 280 300 320 340

x4

Comments:

There seems to exist a positive relationship between x2 (total response of both eyes to stimulus S1 ) and x4, (total reponse of both eyes to stimulus S2).

1.14  
(b)

Compute 

The MEANS Procedure

Variable N **Mean** Std Dev Minimum Maximum

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ

x1 98 **39.1938776** 15.2678209 18.0000000 79.0000000

x2 98 **156.4571429** 22.9033649 125.4000000 238.4000000

x3 98 **4.7326531** 10.8170139 0 90.2000000

x4 98 **207.8326531** 28.8099595 169.2000000 328.0000000

x5 98 **5.0122449** 11.4298512 0 83.0000000

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ

Simple Statistics

Variable N Mean Std Dev Sum Minimum Maximum

x1 98 39.19388 15.26782 3841 18.00000 79.00000

x2 98 156.45714 22.90336 15333 125.40000 238.40000

x3 98 4.73265 10.81701 463.80000 0 90.20000

x4 98 207.83265 28.80996 20368 169.20000 328.00000

x5 98 5.01224 11.42985 491.20000 0 83.00000

Compute ****

Covariance Matrix, DF = 97

x1 x2 x3 x4 x5

x1 233.1063539 108.7455081 7.1307174 127.9049442 3.5171891

x2 108.7455081 524.5641237 141.6686303 607.7403829 106.9976436

x3 7.1307174 141.6686303 117.0077888 161.8438712 112.0255754

x4 127.9049442 607.7403829 161.8438712 830.0137681 124.7756785

x5 3.5171891 106.9976436 112.0255754 124.7756785 130.6414980

1.14 (continued)

(b) continued

Compute **R**

Pearson Correlation Coefficients, N = 98

Prob > |r| under H0: Rho=0

x1 x2 x3 x4 x5

x1 1.00000 0.31098 0.04318 0.29078 0.02015

0.0018 0.6729 0.0037 0.8438

x2 0.31098 1.00000 0.57183 **0.92104** 0.40873

0.0018 <.0001 <.0001 <.0001

x3 0.04318 0.57183 1.00000 0.51933 0.90609

0.6729 <.0001 <.0001 <.0001

x4 0.29078 **0.92104** 0.51933 1.00000 0.37892

0.0037 <.0001 <.0001 0.0001

x5 0.02015 0.40873 0.90609 0.37892 1.00000

0.8438 <.0001 <.0001 0.0001

Comments: The correlation matrix above confirms the answer obtained from part a. There also exist a high positive correction between x3 and x5. Also there seems to exist from slight to high positive correlations between all variables of input as well.

CODE

**data** Multiplesclerosis;

infile 'C:\Documents and Settings\Jaime\Desktop\FALL07\STA5707\T1-6.dat';

input x1 x2 x3 x4 x5;

**proc** **print**;

**run**;

**proc** **plot** data=Multiplesclerosis;

plot x2 \* x4 = '+';

**run**;

**proc** **means** data=Multiplesclerosis;

**run**;

**proc** **corr** covar;

**run**;

1.15

a)

4.0 ˆ +

‚

‚ +

‚

‚ + +

‚ +

3.5 ˆ

‚

‚

‚ +

‚ +

‚

3.0 ˆ + +

‚ +

‚

x2 ‚ + + +

‚ + +

‚ + + + + + +

2.5 ˆ

‚ + +

‚ + +

‚ + +

‚ + +

‚ +

2.0 ˆ + + + + + + +

‚ + + + + + +

‚

‚ + + +

‚ + +

‚ + +

1.5 ˆ +

‚ + + + +

‚ + + + + +

‚ + + + + + +

‚ + + ++

‚ ++ + + + + + +

1.0 ˆ + + +++ ++ + + + + **+**

‚ +

‚

‚

‚

‚

0.5 ˆ

Šƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒƒƒƒƒƒƒƒƒƒˆƒƒ

0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0

x3

Histogram for x2. (Amount of activity)

Used the analyst dialog in SAS.



Histogram for x3. (amount of sleep)

Used the analyst dialog in SAS.



Comments:

There seems to exist a no relationship between x2 (amount of activity) and x3, (amount of sleep).

From the charts above, there seems to be no error in the variable x3. There may be an outlier within the data set. With the scatterplot above, (x2,x3), (1,4). This observation may be an outlier, however, it does not seem to be significant.

1.15  
(b)

Compute 

The MEANS Procedure

Variable N Mean Std Dev Minimum Maximum

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ

x1 98 **3.5423469** 2.1574872 0 10.4610000

x2 98 **1.8093571** 0.7828289 0.9410000 4.0000000

x3 98 **2.1376020** 0.7559291 0.6660000 4.0000000

x4 98 **2.2090000**  0.3322786 1.2860000 2.9370000

x5 98 **2.5748265** 0.9285324 1.0000000 5.0000000

x6 98 **1.2755102** 0.9281465 0 3.0000000

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ

Simple Statistics

Variable N Mean Std Dev Sum Minimum Maximum

x1 98 3.54235 2.15749 347.15000 0 10.46100

x2 98 1.80936 0.78283 177.31700 0.94100 4.00000

x3 98 2.13760 0.75593 209.48500 0.66600 4.00000

x4 98 2.20900 0.33228 216.48200 1.28600 2.93700

x5 98 2.57483 0.92853 252.33300 1.00000 5.00000

x6 98 1.27551 0.92815 125.00000 0 3.00000

Compute ****

Covariance Matrix, DF = 97

x1 x2 x3 x4 x5 x6

x1 4.654750889 0.931345370 0.589699088 0.276915309 1.074885659 0.158150852

x2 0.931345370 0.612821160 0.110933412 0.118469052 0.388886434 -0.024851988

x3 0.589699088 0.110933412 0.571428861 0.087004959 0.347989910 0.110131391

x4 0.276915309 0.118469052 0.087004959 0.110409072 0.217405649 0.021814433

x5 1.074885659 0.388886434 0.347989910 0.217405649 0.862172372 -0.008817694

x6 0.158150852 -0.024851988 0.110131391 0.021814433 -0.008817694 0.861455923

1.15 (continued)

(b) continued

Compute **R**

Pearson Correlation Coefficients, N = 98

Prob > |r| under H0: Rho=0

x1 x2 x3 x4 x5 x6

x1 1.00000 0.55144 0.36158 0.38627 0.53656 0.07898

<.0001 0.0003 <.0001 <.0001 0.4395

x2 0.55144 1.00000 **0.18746** 0.45544 0.53501 -0.03420

<.0001 0.0645 <.0001 <.0001 0.7381

x3 0.36158  **0.18746** 1.00000 0.34639 0.49578 0.15697

0.0003 0.0645 0.0005 <.0001 0.1227

x4 0.38627 0.45544 0.34639 1.00000 **0.70465** 0.07073

<.0001 <.0001 0.0005 <.0001 0.4889

x5 0.53656 0.53501 0.49578 **0.70465** 1.00000 -0.01023

<.0001 <.0001 <.0001 <.0001 0.9204

x6 0.07898 -0.03420 0.15697 0.07073 -0.01023 1.00000

0.4395 0.7381 0.1227 0.4889 0.9204

Comments

Interpret the pair wise correlations

There exists a moderate positive and significant correlation between x4 (amount of food consumption) and x5 (appetite). As from the scatter plot above and the correlations above, the variables x2 and x3 exhibit a non-significant positive correlation. Between a number of vairiables, such as x2 and x1, there exist a slight positive (>0.5) correlation.

Analyzing the negative correlations, there seems to exist no strong significant negative correlation.

CODE

**data** Radio;

infile 'C:\Documents and Settings\Jaime\Desktop\FALL07\STA5707\T1-7.dat';

input x1 x2 x3 x4 x5 x6;

**proc** **print**;

**run**;

**proc** **plot** data=Radio;

plot x2 \* x3 = '+';

**run**;

**proc** **means** data=Radio;

**run**;

**proc** **corr** covar;

**run**;