IT3030 quiz#5 6/16/2020   ID:_____   Name:_____

1. (50%) The following table summarizes the relationship between hypertension and coronary artery disease (CAD) for people over 65-year-old.

Hypertension	CAD Yes	CAD No	Total
Yes	1102	87	1189
No	1018	106	1124
Total	2120	193	2313

(a) Determine the expected frequency table.
(b) Compute the chi-square value.
(c) Compute the p-value of the hypothesis test, assuming that no correlation exists between hypertension and CAD, using $\alpha=0.05$.
(d) Your conclusion of the relationship between hypertension and coronary artery based on what you have computed?

Answer:

Hypertension	CAD Yes	CAD No	Total
Yes	1089.8	99.2	1189
No	1030.2	93.8	1124
Total	2120	193	2313

>> O = [1102   87   1018   106];

>> E = [2120*1189/2313 193*1189/2313 2120*1124/2313 193*1124/2313]

= 1.0e+003 *

1.0898   0.0992   1.0302   0.0938

>> (O-E).^2./E = 0.1368   1.5031   0.1448   1.5901

>> chi2=sum(ans) = 3.3748

>> 1-chi2cdf(chi2,1) =   0.0662

Since p-value is greater than 0.05, we do not reject the null hypothesis that no correlation exists between the two mentioned variables. That is, no statistically significant correlation exists between hypertension and CAD.

2. (50%) The following table describes the relationship between the quiz scores and the exam scores for 10 students in Biostatistics. Answer the following questions.

ID	Quiz	Exam
1	46	68
2	59	57
3	26	68
4	84	91
5	70	76
6	79	84
7	57	83
8	75	84
9	51	80
10	61	70

(a) Graph an X-Y scatter plot for visually assessing whether the quiz scores are correlated to the exam scores. Explain whether your data is weakly or strongly correlated, positively or negatively correlated, or not correlated at all.

(b) Find the Pearson correlation coefficient of the two variables. Let X be the quiz scores and Y be the exam scores.

(c) Is such correlation significant? (Find the p-value and use 0.05 as the level of significance for the hypothesis testing.)

(d) Find the linear regression equation between X and Y.

(e) Given one's quiz score 60, what will be the predicted exam score based on the equation found in (d)?

Answer:
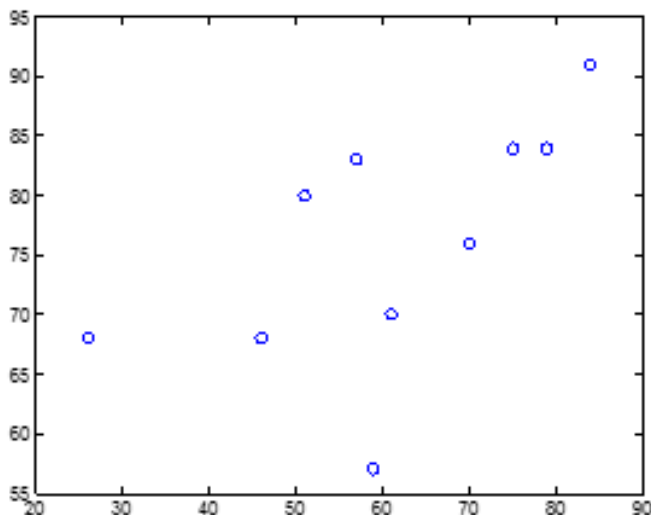
X =[46   59   26   84   70   79   57   75   51   61];

Y =[68   57   68   91   76   84   83   84   80   70];

>> >> [R,P]=corrcoef(X,Y)

R =   1.0000   0.6087
      0.6087   1.0000

P =   1.0000   0.0618
      0.0618   1.0000

The coefficient **R = 0.6087**. The p-value is **0.0618**, which is greater than 0.05. Thus the correlation is not significant.

>> polyfit(X,Y,1) =   0.3598   54.2268

>>

This gives **Y=0.3598*X + 54.2268** as the regression equation.

>> >> polyval(polyfit(X,Y,1),60)=   **75.8122**

>>

This gives an exam score of 75.8122 with one having a quiz score of 60.