Correlation and Linear Regression

PRANAVCHENDUR T K

15BCE1097

Faculty: ARUN PRASATH G M

Aim:

Model fitting and investigating relationship between two variables within a regression framework.

Program:

> #PROBLEM 1

AGE GROUP	REPRESENTATIVE AGE	HOURS SPENT IN THE LOCAL LIBRARY
10-19	15	302.28
20-29	25	193.63
30-39	35	185.46
40-49	45	198.49
50-59	55	224.30
60-69	65	288.71

Spearman's Rank Correlation Coefficient
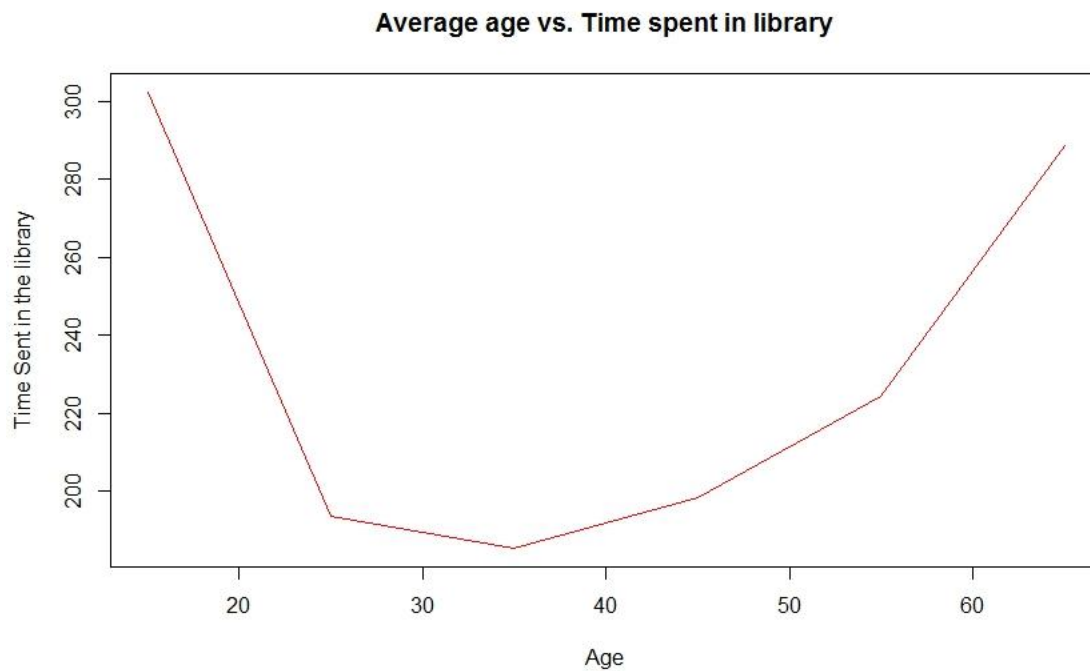Rho = 1-[6Sigma(d^2)/n(n^2-1)]

Spearman's Rank Correlation Coefficient for a data

> x<-c(15,25,35,45,55,65)
> y<-c(302.28,193.63,185.46,198.49,224.30,288.71)

> plot(x,y,type = "l",main="Average age vs. Time spent in library",xlab="Age",ylab = "Time Sent in the library",col="red")

Average age vs. Time spent in library


```
> #PROBLEM 2
> var(x)
[1] 350

> var(y)
[1] 2595.123

> r=var(x,y)/sqrt(var(x)*var(y)) #Spearmans

> r
[1] 0.03902233

> cor(x,y)
[1] 0.03902233

> cor.test(x,y)

        Pearson's product-moment correlation

data:  x and y
t = 0.078104, df = 4, p-value = 0.9415
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.7978046  0.8244734
sample estimates:
     cor
0.03902233
```

```
> cor.test(x,y,method = "pearson")

        Pearson's product-moment correlation

data:  x and y
t = 0.078104, df = 4, p-value = 0.9415
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.7978046  0.8244734
sample estimates:
     cor
0.03902233


> #Import Hello.CCSV
> View(hello)

> hello<-hello[c("BIRTHS2015","DEATHS2015")]

> cor(hello)
           BIRTHS2015 DEATHS2015
BIRTHS2015   1.000000   0.997816
DEATHS2015   0.997816   1.000000

> cor(hello)^2
           BIRTHS2015 DEATHS2015
BIRTHS2015  1.0000000  0.9956367
DEATHS2015  0.9956367  1.0000000
```

Problem :-
The bod weight and the BMI of 12 school going children are given in the following table

Weight	15	26	27	25	25.5	27	32	18	22	20	26	24
BMI	13.35	16.12	16.74	16.00	13.59	15.73	15.65	13.85	16.07	12.8	13.65	14.42

```
> weight=c(15,26,27,25,25.5,27,32,18,22,20,26,24)

> bmi=c(13.35,16.12,16.74,16,13.59,15.73,15.65,13.85,16.07,12.8,13.65,14.42)

> cor(weight,bmi)
```

[1] 0.5790235

> mode1<-lm(bmi~weight)

> summary.lm(mode1)

Call:
lm(formula = bmi ~ weight)

Residuals:
    Min   1Q Median   3Q   Max
-1.52988 -0.75527 0.04426 0.95286 1.57397

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 10.73487   1.85405   5.790 0.000175 ***
weight   0.17096   0.07612   2.246 0.048524 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.155 on 10 degrees of freedom
Multiple R-squared: 0.3353,	Adjusted R-squared: 0.2688
F-statistic: 5.044 on 1 and 10 DF, p-value: 0.04852