Statist.ics Practical 5

**Q Find Correlation and Regression.**

**Program:**

x<-c(68,42,15,75,91,25,95,30,58,82,50,99)

y<-c(170,80,32,200,275,40,347,60,141,208,97,378)

#find covariance using cov function

a<-cov(x,y)

a

#find standard deviation

b<-sd(x)

b

c<-sd(y)

c

#karl pearson's correlation coefficient

cr<-(a/(b\*c))

cr

#karl pearson's correlation coefficient

cr<-cor(x,y)

cr

#Scatter plot of x and y

plot(x,y)

**Output:**

x<-c(68,42,15,75,91,25,95,30,58,82,50,99)

> y<-c(170,80,32,200,275,40,347,60,141,208,97,378)

>

> #find covariance using cov function

> a<-cov(x,y)

> a

[1] 3211.455

>

> #find standard deviation

> b<-sd(x)

> b

[1] 28.68269

> c<-sd(y)

> c

[1] 116.8776

>

> #karl pearson's correlation coefficient

> cr<-(a/(b\*c))

> cr

[1] 0.9579673

>

> #karl pearson's correlation coefficient

> cr<-cor(x,y)

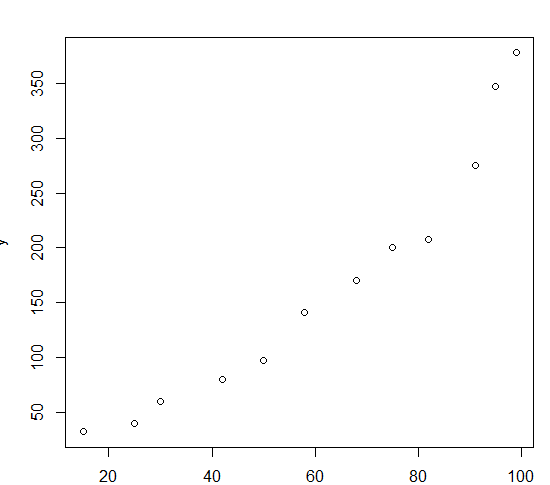
> cr

[1] 0.9579673

>

> #Scatter plot of x and y

> plot(x,y)



**Q Find the correlation between age of husband and wife for 6 couples.**

**age of husbands 28, 30, 27, 26, 32, 29**

**age of wives 24, 25, 24, 20, 28, 26**

**Program:**

x<-c(28,30,27,26,32,29)

y<-c(24,25,24,20,28,26)

#find covariance using cov function

a<-cov(x,y)

a

#find standard deviation

b<-sd(x)

b

c<-sd(y)

c

#karl pearson's correlation coefficient

cr<-(a/(b\*c))

cr

#karl pearson's correlation coefficient

cr<-cor(x,y)

cr

#Scatter plot of x and y

plot(x,y)

**Output:**

x<-c(28,30,27,26,32,29)

> y<-c(24,25,24,20,28,26)

>

> #find covariance using cov function

> a<-cov(x,y)

> a

[1] 5.2

>

> #find standard deviation

> b<-sd(x)

> b

[1] 2.160247

> c<-sd(y)

> c

[1] 2.664583

>

> #karl pearson's correlation coefficient

> cr<-(a/(b\*c))

> cr

[1] 0.9033806

>

> #karl pearson's correlation coefficient

> cr<-cor(x,y)

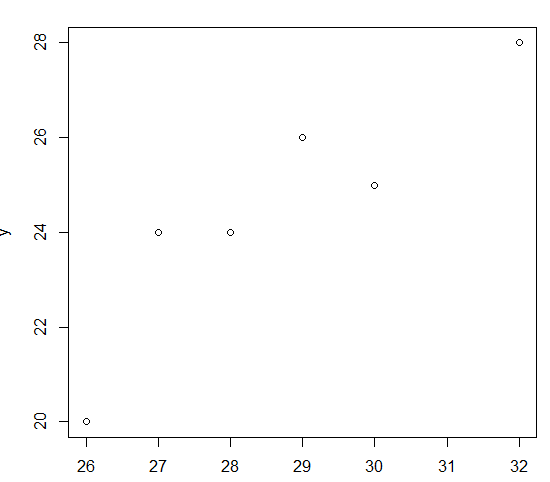
> cr

[1] 0.9033806

>

> #Scatter plot of x and y

> plot(x,y)



**Q Find out regretion of y on x regression coefficients and coeffient of determination**

**Program:**

#regression on line, coefficient of determination

x<-c(2,4,8,11,14,15)

y<-c(10,15,18,20,24,30)

z<-data.frame(x,y)

z

#bivariate regression

#regresson on y on x

#regression coefficient

byx<-(cov(x,y)/var(x))

mean(x)

mean(y)

#we directly get regression using lm function

rg<-lm(y~x)

rg

#to find ceofficient of determination

cdr<-(cor(x,y)\*cor(x,y))

cdr

summary(rg)

#regression of x on y

bxy<-(cov(x,y)/var(y))

mean(x)

mean(y)

#we directly get regression using lm function

rg<-lm(x~y)

rg

#to find ceofficient of determination

cdr<-(cor(x,y)\*cor(x,y))

cdr

summary(rg)

**Output:**

> #regression on line, coefficient of determination

> x<-c(2,4,8,11,14,15)

> y<-c(10,15,18,20,24,30)

>

> z<-data.frame(x,y)

> z

x y

1 2 10

2 4 15

3 8 18

4 11 20

5 14 24

6 15 30

>

> #bivariate regression

>

> #regresson on y on x

>

> #regression coefficient

> byx<-(cov(x,y)/var(x))

>

> mean(x)

[1] 9

> mean(y)

[1] 19.5

>

> #we directly get regression using lm function

> rg<-lm(y~x)

> rg

Call:

lm(formula = y ~ x)

Coefficients:

(Intercept) x

8.121 1.264

>

> #to find ceofficient of determination

> cdr<-(cor(x,y)\*cor(x,y))

> cdr

[1] 0.9190085

>

> summary(rg)

Call:

lm(formula = y ~ x)

Residuals:

1 2 3 4 5 6

-0.6500 1.8214 -0.2357 -2.0286 -1.8214 2.9143

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 8.1214 1.9168 4.237 0.01330 \*

x 1.2643 0.1877 6.737 0.00253 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.22 on 4 degrees of freedom

Multiple R-squared: 0.919, Adjusted R-squared: 0.8988

F-statistic: 45.39 on 1 and 4 DF, p-value: 0.002529

>

> #regression of x on y

>

> bxy<-(cov(x,y)/var(y))

>

> mean(x)

[1] 9

> mean(y)

[1] 19.5

>

> #we directly get regression using lm function

> rg<-lm(x~y)

> rg

Call:

lm(formula = x ~ y)

Coefficients:

(Intercept) y

-5.1745 0.7269

>

> #to find ceofficient of determination

> cdr<-(cor(x,y)\*cor(x,y))

> cdr

[1] 0.9190085

>

> summary(rg)

Call:

lm(formula = x ~ y)

Residuals:

1 2 3 4 5 6

-0.09446 -1.72895 0.09035 1.63655 1.72895 -1.63244

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -5.1745 2.2134 -2.338 0.07957 .

y 0.7269 0.1079 6.737 0.00253 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.684 on 4 degrees of freedom

Multiple R-squared: 0.919, Adjusted R-squared: 0.8988

F-statistic: 45.39 on 1 and 4 DF, p-value: 0.002529

**Q Find the regression of prize on supply and supply on prize**

**and the coefficient of determination for both**

**Program:**

supply<-c(80,84,86,88,92,96,97)

price<-c(12,11,15,15,18,16,18)

bsupplypirce<-(cov(supply,price)/var(price))

bsupplypirce

bpricesupply<-(cov(supply,price)/var(supply))

bpricesupply

mean(supply)

mean(price)

#we directly get regression using lm function

sp<-lm(supply~price)

sp

ps<-lm(price~supply)

ps

summary(sp)

summary(ps)

#to find ceofficient of determination

cdr<-(cor(supply,price)\*cor(supply,price))

cdr

**Output**:

> supply<-c(80,84,86,88,92,96,97)

> price<-c(12,11,15,15,18,16,18)

>

> bsupplypirce<-(cov(supply,price)/var(price))

> bsupplypirce

[1] 1.977273

>

> bpricesupply<-(cov(supply,price)/var(supply))

> bpricesupply

[1] 0.3655462

>

> mean(supply)

[1] 89

> mean(price)

[1] 15

>

> #we directly get regression using lm function

> sp<-lm(supply~price)

> sp

Call:

lm(formula = supply ~ price)

Coefficients:

(Intercept) price

59.341 1.977

>

> ps<-lm(price~supply)

> ps

Call:

lm(formula = price ~ supply)

Coefficients:

(Intercept) supply

-17.5336 0.3655

>

> summary(sp)

Call:

lm(formula = supply ~ price)

Residuals:

1 2 3 4 5 6 7

-3.068 2.909 -3.000 -1.000 -2.932 5.023 2.068

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 59.3409 8.3284 7.125 0.000845 \*\*\*

price 1.9773 0.5476 3.611 0.015371 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.633 on 5 degrees of freedom

Multiple R-squared: 0.7228, Adjusted R-squared: 0.6673

F-statistic: 13.04 on 1 and 5 DF, p-value: 0.01537

>

> summary(ps)

Call:

lm(formula = price ~ supply)

Residuals:

1 2 3 4 5 6 7

0.28992 -2.17227 1.09664 0.36555 1.90336 -1.55882 0.07563

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -17.5336 9.0299 -1.942 0.1098

supply 0.3655 0.1012 3.611 0.0154 \*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.562 on 5 degrees of freedom

Multiple R-squared: 0.7228, Adjusted R-squared: 0.6673

F-statistic: 13.04 on 1 and 5 DF, p-value: 0.01537

>

> #to find ceofficient of determination

> cdr<-(cor(supply,price)\*cor(supply,price))

> cdr

[1] 0.7227846

**Q What is the probability that a 5 card hand selected**

**from a well shuffled pack of card containing 3 aces**

**Program:**

prA<-((choose(4,3)\*choose(48,2))/choose(52,5))

prA

**Output:**

> prA<-((choose(4,3)\*choose(48,2))/choose(52,5))

> prA

> 0.001736079

**Q A commitee of university teachers consists of 3 professros 5 readers and 2 lecturerers .A sub commitee of 6 is selected, what is the probability that the sub-commitee composed of**

**2 professors 3 readers and 1 lecturerer**

**Program:**

prAC<-((choose(3,2)\*choose(5,3)\*choose(2,1))/choose(10,6))

prAC

**Output:**

> prAC<-((choose(3,2)\*choose(5,3)\*choose(2,1))/choose(10,6))

> prAC

> 0.2857143