

8

Ch. 10

- Scatter Plots and Correlation.
- Regression.

Ch. 13 - Sec. 6

- The Spearman Rank
Correlation Coefficient.

STAT. 110

جمال السعدي
رياضيات - إحصاء



Ch. 10 Part. 1

Correlation

→ الارتباط

Is a statistical method to determine whether a relationship between variables exists.

Scatter plots is show the relationship between the independent and dependent variables.

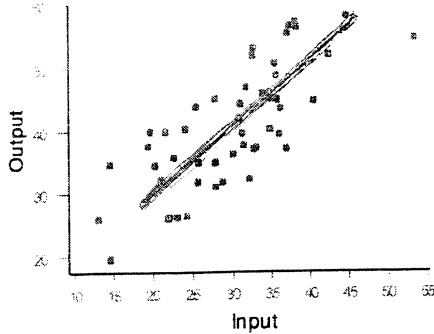
← أشكال الانتشار

Types of correlation

→ انواع الارتباط

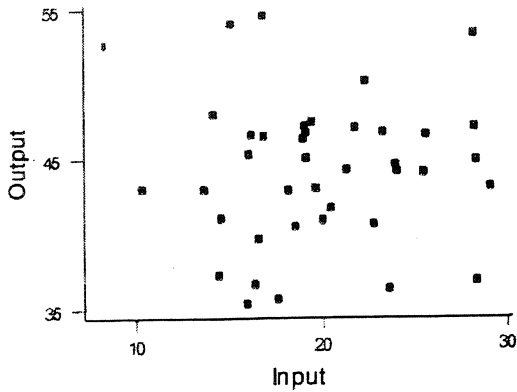
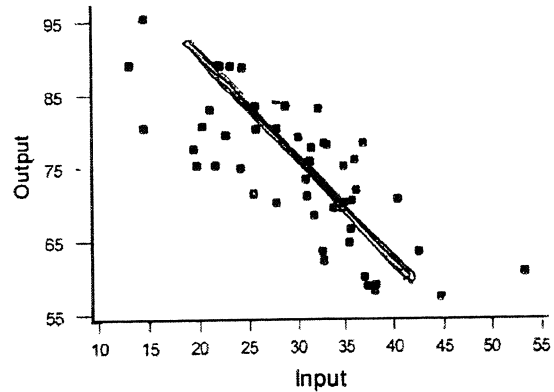
Positive correlation

As x increases, y also increases



Negative correlation

As x increases, y decreases



الرسم يوضح نوع ومدى العلاقة بين متغيرين
(ظاهرتين) X, Y من حيث

مدى العلاقة	نوع العلاقة
Strong قوية	positive طردية
Weak ضعيفة	negative عكسية
Zero منعدمة	no relation لا توجد علاقة

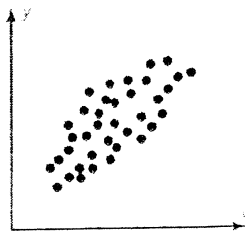
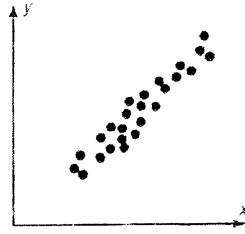
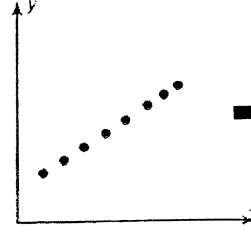
A
L
S
A
A
D
I

معامل الارتباط

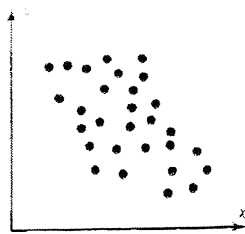
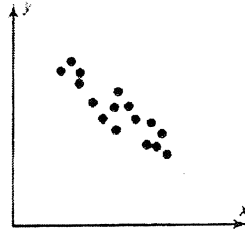
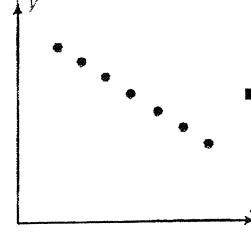
The correction coefficient computed from the sample data measures the strength and direction of linear relationship between two variables. The symbol for the sample correlation coefficient is r . the symbol for the population correlation coefficient is ρ .

The range of the correlation coefficient is $[-1, 1]$.

- If $r = -1$: There is a perfect(complete) negative linear relationship.
- If $r = 1$: There is a perfect(complete) positive linear relationship.
- If $r = 0$: There is no linear relationship (Does not exist).

(a) $r = 0.50$ (b) $r = 0.90$ (c) $r = 1.00$

➔ Positive

(d) $r = -0.50$ (e) $r = -0.90$ (f) $r = -1.00$

➔ Negative

weak
ضعيف

strong
قوى

Perfect (complete)
تام

A
L
S
A
A
D
I

* **The Pearson Correlation coefficient r** : معامل ارتباط بيرسون

$$r = \frac{n \cdot \sum xy - \sum x \cdot \sum y}{\sqrt{[n \cdot \sum x^2 - (\sum x)^2] \times [n \cdot \sum y^2 - (\sum y)^2]}}$$

Note

مدى العلاقة	نوع العلاقة
• $r = 0$	no relation
• $0.01 \leq r \leq 0.29$	Very weak ضعيف جدًا
• $0.30 \leq r \leq 0.49$	Weak ضعيف
• $0.50 \leq r \leq 0.69$	Moderate متوسط
• $0.70 \leq r \leq 0.89$	Strong قوي
• $0.90 \leq r \leq 0.99$	Very strong قوي جدًا
• $r = 1$	Perfect (complete)

- n is the number of data pairs. (عدد خانات الجدول)
- Round the value of r to two decimal places

Example

Compute the value of the correlation coefficient for the data obtained in the study of the number of absences and the final grade of the seven students

Sample**Solution**

Student	Number of Absences x	Final grad y (%)	x . y	x ²	y ²
A	6	82	492	36	6724
B	2	86	172	4	7396
C	15	43	645	225	1849
D	9	74	666	81	5476
E	12	58	696	144	3364
F	5	90	450	25	8100
G	8	78	624	64	6084
$\sum x = 57$ $\sum y = 511$			$\sum xy = 3745$	$\sum x^2 = 579$	$\sum y^2 = 38993$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

$$= \frac{(7)(3745) - (57)(511)}{\sqrt{[(7)(579) - (57)^2][(7)(38993) - (511)^2]}} = -0.94$$

- **Strong negative relationship** between a student's Final grade and the number of absences a student has.

Example

Compute the value of the correlation coefficient for the data obtained in the study of age and blood pressure.

Sample**Solution**

Subject	Age x	Pressure y	x . y	x ²	y ²
A	43	128	5504	1849	16384
B	48	120	5760	2304	14400
C	56	135	7560	3136	18225
D	61	143	8723	3721	20449
E	67	141	9447	4489	19881
F	70	152	10640	4900	23104
$\sum x=345$ $\sum y=819$			$\sum xy=47634$	$\sum x^2=20399$	$\sum y^2=112443$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

$$= \frac{(6)(47634) - (345)(819)}{\sqrt{[(6)(20399) - (345)^2][(6)(112443) - (819)^2]}} = \underline{\underline{0.89}}$$

The correlation coefficient is a strong positive relationship between age and blood pressure.

A
L
S
A
A
D
I

Use the following to answer questions

In the study of relationship between the number of absences X and the final grade Y of 8 students in the statistic class, the data are shown as follows

$$n = 8$$

$$\sum X = 42, \sum Y = 470, \sum XY = 3143, \sum X^2 = 354 \text{ and } \sum Y^2 = 37358$$

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2] \cdot [n \sum y^2 - (\sum y)^2]}}$$

$$= \frac{8(3143) - (42)(470)}{\sqrt{[8(354) - (42)^2][8(37358) - (470)^2]}}$$

$$= \boxed{0.592}$$

The value of the Pearson correlation coefficient means that there is a ... linear relationship between the number of absences and the final grade.

A) strong negative B) strong positive C) moderate negative D) moderate positive

$$\therefore r = 0.592$$

موجب

من صورته بين 0.5 و 0.69

∴ moderate positive

The range of the Pearson correlation coefficient value (r) for the positive linear relationship is ...

A) $0 < r \leq 1$ B) $0 \leq r \leq 1$ C) $0 \leq r < 1$ D) $-1 \leq r \leq 1$

موجب

The range of positive r is $0 < r \leq 1$

If the value of the correlation coefficient equals -0.95, then the type of the relationship is ...
 A) weak negative B) strong negative C) strong positive D) weak positive

عندما تكون $r = -0.95$
 فإنها تكون strong negative

If the value of the correlation coefficient equals -0.19, then the type of the relationship is ...
 A) weak negative B) strong negative C) strong positive D) weak positive

If Pearson correlation coefficient (r_p) equals 0.45, then the relationship can be described as
 A) positive, strong and non linear C) positive, moderate and non linear
 B) weak and linear D) positive, moderate and linear

$r_p = 0.45$ is weak and linear

- A positive relationship exists when both variables increase or decrease at the same time.
- A negative relationship exists when one variable increases and the other variable decreases.

If there is a strong positive linear relationship between the variables, the value of r will be close to +1.

If there is a strong negative linear relationship between the variables, the value of r will be close to -1.

كمية الدهون

الوزن

The correlation coefficient between the amount of fats which a person eats and his or her weight may be

- A) close to -1 B) close to 2 C) close to 1 D) 0

* العلاقة بين كمية الدهون التي يتناولها الشخص ووزنه الشخص
علاقة طردية قوية \leftarrow تقترب من +1
 \Rightarrow close to 1

الارتباط

الدخل السنوي الوزن

If the correlation between body weight and annual income were high and positive, we could conclude that high income people

- A) are eating more food C) are losing weight
B) are eating less food D) are gaining weight

* إذا كانت العلاقة بين الوزن والدخل السنوي
high and positive علاقته قوية بوجبه (طردية)

∴ أصحاب الدخل العالي يزدادون من الوزن
أما كلما زاد الدخل كلما زاد الوزن .
 \leftarrow الاختيار الصحيح هو D

A negative relationship between two variables means that for the most part, as the X variable decreases, the Y variable

- A) stays the same B) increases C) decreases D) equals X

Negative relationship

= X decreases \Rightarrow y increases

x = y عكس بعضهم

CH. 10 Part 2

Regression

الانحدار

Regression is a statistical method used to describe the nature of the relationship between variables.

In a simple relationship, there are only two types of variables under study; an independent variable (explanatory variable) or predictor variable and a dependent variable (outcome variable) or response variable

Equation of regression Line is: $\hat{y} = a + b x$

- Another name is line of best fit.

Where b: is the slope of the line.

a: is y – Intercept.

$$b = \frac{n \sum xy - \sum x \cdot \sum y}{n \sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y \cdot \sum x^2 - \sum y - \sum xy}{n \sum x^2 - (\sum x)^2}$$

- round the values of a and b to three decimal places

Example

Instructor is interested in finding the relationship between the final exam grades of students enrolled in calculus I (x)

And calculus II (y). use the following information to answer the following question:

$$n = 5, \sum x = 400, \sum y = 380, \sum x^2 = 32656$$

$$\sum y^2 = 29546, \sum xy = 31023, \text{slope } (b) = 0.95$$

(1) The Pearson correlation coefficient is :

- (a) 0.94 (b) 0.81 (c) 0.68 (d) 0.34

Solution

$$r = \frac{n \cdot \sum xy - \sum x \cdot \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \cdot \sum y^2 - (\sum y)^2]}}$$

$$= \frac{5 \cdot (31023) - (400)(380)}{\sqrt{[5 \cdot (32656) - (400)^2][5 \cdot (29546) - (380)^2]}} = \underline{\underline{0.94}}$$

(2) Interpret the value of the person correlation coefficient (r).

- a. Weak positive linear correlation.
 b. Weak negative linear correlation.
 ✓ c. Strong positive linear correlation.
 d. Strong negative linear correlation.

(3) The equation of the regression line is:

a. $y' = 0.154 + 1.48x$

b. $y' = 0.924 + 0.925x$

c. $y' = 0.024 + 0.95x$

d. $y' = 0.95 + 0.024x$

Solution

• $b = 0.95$ (the slope)

$$a = \frac{\sum y \cdot \sum x^2 - \sum x \cdot \sum xy}{n \cdot \sum x^2 - (\sum x)^2}$$

$$= \frac{(380)(32656) - (400)(31023)}{5(32656) - (400)^2} = 0.024$$

Equation of the regression line

$$y' = a + b x$$

there for $y' = 0.024 + 0.95x$

(4) predict a calculus II exam score for student who get 80 in calculus I.

a. 85

b. 80

c. 70

d. 76

Solution

$$\therefore \hat{y} = 0.024 + 0.95 x$$

$$= 0.024 + 0.95 (80) = 76.024 \cong \underline{\underline{76}}$$

Example:

Age and sick days

Age x	18	26	39	48	53	58
Days y	16	12	9	5	6	2

- Find y' when $x = 47$ years

Solution

$$n = 6, \sum x = 242, \sum y = 50$$

$$\sum x^2 = 10998, \sum y^2 = 546, \sum xy = 1625$$

$$b = \frac{n \sum xy - \sum x \cdot \sum y}{n \sum x^2 - (\sum x)^2} = \frac{6 \cdot (1625) - (242) \cdot (50)}{6 \cdot (10998) - (242)^2} = \underline{\underline{-0.317}}$$

$$a = \frac{n \sum y \cdot \sum x^2 - \sum y \cdot \sum xy}{n \sum x^2 - (\sum x)^2} = \frac{(50)(10998) - (242)(1625)}{6 \cdot (10998) - (242)^2} = \underline{\underline{21.1}}$$

Equation of regression line:

$$\hat{y} = a + b x$$

$$\hat{y} = 21.1 + (-0.317) x$$

when $x = 47$

$$\hat{y} = 21.1 + (-0.317)(47) = 6.201 \cong 6 \text{ days}$$

الميل

The slope of the regression line $y' = -2x + 5$ is:

- A) 2 B) 5 C) -5 D) -2

$$y' = -2x + 5$$

↓
الميل
slope
↓
الجزء المقطوع من محور y
y-intercept

$$\therefore \text{slope} = \underline{\underline{-2}}$$

عمر الرجل x ضعف عمر الزوجه y

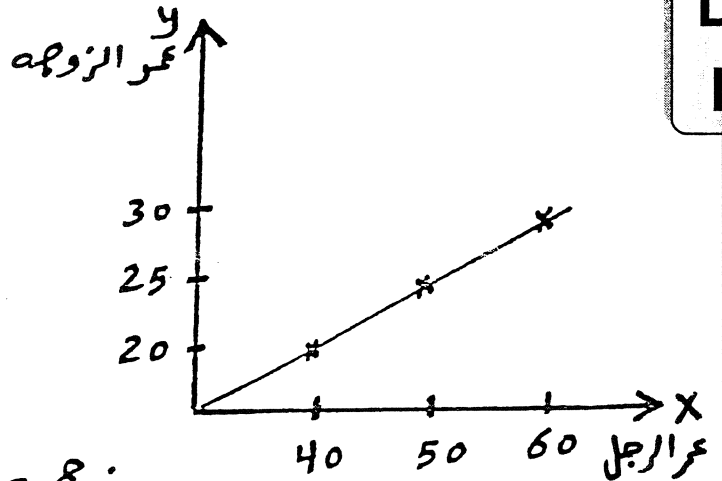
Suppose that every man's age (x) is exactly twice his wife's age (y). Then the Pearson correlation coefficient (r) between x and y is

- A) 0.5 B) -0.5 C) 1 D) -1

$$x = 2y$$

$$y = \frac{1}{2}x$$

على فرضه أم :
عمر الرجل x
ضعف
عمر الزوجه y



∴ جميع النقاط تقع على استقامه واحد

وبالتالي يكون معامل الارتباط $r = 1$

Use the following to answer questions

عمر الشخص

عدد ساعات
ممارسة الرياضةThe equation of the regression line between a person's age in years (x) and the number of hours he exercises per week (y) is given by $\hat{y} = 10.944 - 0.18x$ العلاقة تمثل الجملة الصحيحة
The correct statement that represents the relationship between (x) and (y) is:

- A) When the number of hours he exercises increases by 1 hour, his age increases by 10.944 on average.
- B) When the number of hours he exercises decreases by 1 hour, his age decreases by 10.944 on average.
- C) When a person's age increases by 1 year, the number of hours he exercises decreases by 0.18 on average.
- D) When a person's age increases by 1 year, the number of hours he exercises increases by 0.18 on average.

العلاقة بين العمر x ، عدد ساعات التمارين y
(ممارسة الرياضة)

مثله بالمعادلة

$$\hat{y} = 10.944 - 0.18x$$

* العلاقة عكسية بدليل وجود إشارة الب أمام x * كلما زاد العمر x بقدر ا سنةكما قل عدد ساعات التمارين الرياضية y بقدر ا ساعة 0.18 ∴ الإجابة الصحيحة هي **C**

توقع

 y ?

Predict the number of hours a person exercises per week when his age is 50 years.

A) 1.49

B) 10.04

C) 19.9

D) 1.944عندما $x = 50$

توقع قيمة

 y ?

$$\therefore \hat{y} = 10.944 - 0.18x$$

$$\hat{y} = 10.944 - 0.18(50) \quad \text{بالإدخال}$$

$$\hat{y} = 1.944$$

∴ نضع $x = 50$

في المعادلة

ونوجد قيمة y

There is an approximate linear relationship between the height (y in cm) and age (x from 5 to 18 years) described by $y = 50 + 6 * x$. Which of the following is NOT CORRECT?

- A) Children reach the height of 50 cm when they are $50/6=8.33$ years old. اي اعمل
الأشياء
ليتها صححة.
- B) Children's heights increase by 6 cm for each year they grow older
- C) The estimated height of a 10 years old child is 110 cm
- D) My son is 7 years old and is 110 cm tall. He is taller than average

العلاقة الخطية بين y height (y) و x age (x)

$$y = 50 + 6x$$

A) يكون الطول 50 عندها يكون العمر $\frac{50}{6}$

$$x = \frac{50}{6}$$

فنجد أنه قيمته y هي

$$y = 50 + 6 \left(\frac{50}{6} \right) = 100$$

∴ الطول 100 وليس 50 ∴ الجملة A خاطئة
Not correct

When the correlation coefficient (r) equals zero, the linear relationship between the variables

- A) is strong B) is weak C) is moderate D) does not exist غير موجوده

إذا كانت قيمة معامل الارتباط $r = 0$
معنا معناه لا توجد علاقة بين المتغيرين

Use the following to answer questions

The following data is recorded to determine the relationship between the number of hours a person goes without sleeping (x) and the number of mistakes he makes on a simple test (y):

$$n = 10, \sum x = 46, \sum y = 60, \sum xy = 303, \sum x^2 = 238, \sum y^2 = 616$$

The Pearson correlation coefficient (r) is

- A) 0.24 B) 0.33 C) -0.24 D) -0.33

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

قانونه

يعطى من نهاية

ورقة الأسئلة .

تعويم مباشر من المعطيات

$$= \frac{10(303) - (46)(60)}{\sqrt{[10(238) - (46)^2][10(616) - (60)^2]}}$$

بالأله \approx 0.33

A
L
S
A
A
D
I

The value of the Pearson correlation coefficient (r) means that there is

- A) strong negative linear relationship C) strong positive linear relationship
B) moderate negative linear relationship D) weak positive linear relationship

The equation of the regression line is

- A) $y' = 1.02 - 1.3x$ B) $y' = 1.3 - 1.02x$ C) $y' = 1.02 + 1.3x$ D) $y' = 1.3 + 1.02x$

$b < a$ تيم حساب eq. of regression line. *

$$*b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$= \frac{10(303) - (46)(60)}{10(238) - (46)^2} = \boxed{1.02}$$

$$*a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$= \frac{(60)(238) - (46)(303)}{10(238) - (46)^2} = \boxed{1.3}$$

$$\rightarrow y' = a + bx$$

$$\boxed{y' = 1.3 + 1.02x}$$

The correct statement that represents the relationship between (x) and (y) is:

- A) When the number of hours a person goes without sleeping increases by 1 hour, his number of mistakes increases by 1.02 on average.
 B) When the number of hours a person goes without sleeping increases by 1 hour, his number of mistakes decreases by 1.02 on average.
 C) When the number of mistakes increases by 1, the number of hours a person goes without sleeping decreases by 1.3 on average.
 D) When the number of mistakes increases by 1 year, the number of hours a person goes without sleeping increases by 1.3 on average.

من السؤال السابق

$$\hat{y} = 1.3 + 1.02x$$

number of hours without sleeping x increasing by 1
 " " mistakes y increasing by 1.02
 (ساعات x)

توقع

y ?

x

Predict the number of mistakes for a person who goes 9.51 hours without sleeping.

A) -8

B) 11

C) 13

D) -11

توقع قيمه y ? عندها $x = 9.51$

ذعوهم من المعادله عنده $x = 9.51$

$$\hat{y} = 1.3 + 1.02x$$

$$\hat{y} = 1.3 + 1.02(9.51)$$

$$\hat{y} \approx 11$$

Use the following to answer questions

In the study of relationship between the number of absences X and the final grade Y of 6 students in the statistic class, the data are shown as follows

$$\sum X = 42, \sum Y = 470, \sum XY = 3143, \sum X^2 = 354 \text{ and } \sum Y^2 = 37358$$

$$\begin{array}{c} \downarrow \\ n = 6 \end{array}$$

The correlation coefficient is

- A) 0.82 B) 0.92 C) -0.82 D) 1

by التعويض substitute in the الصيغة formula ;

$$\begin{aligned} r &= \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \\ &= \frac{6 (3143) - (42)(470)}{\sqrt{[6(354) - (42)^2][6(37358) - (470)^2]}} = -0.815 \\ &\approx \underline{\underline{-0.82}} \end{aligned}$$

The slope of the regression line is

- A) 2.45 B) -3.45 C) 3.45 D) -2.45

The ميل slope of خط الانحدار regression line is

$$\begin{aligned} b &= \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} \\ &= \frac{6 (3143) - (42)(470)}{6 (354) - (42)^2} = \underline{\underline{-2.45}} \end{aligned}$$

A negative relationship between two variables means that for the most part ,as the X variable increases, the Y variable

- A) decreases B) increases C) equals X D) remains the same

*negative relation
 X increas → y decreas
 حل السؤال

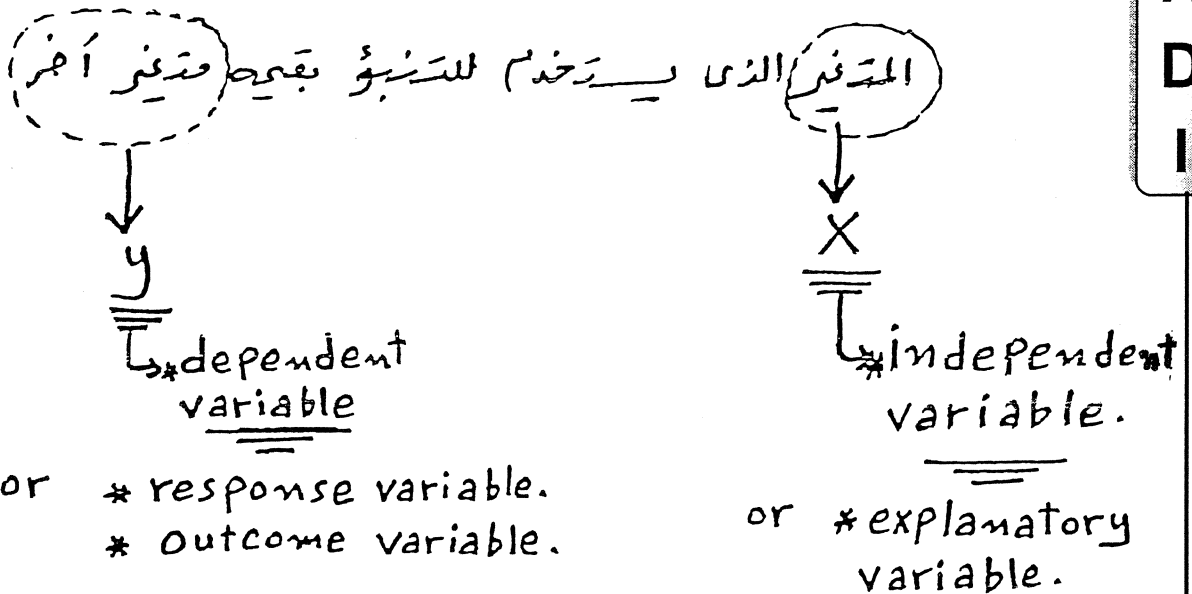
* positive relation
 X increas → y increas
 X decreas → y decreas
 شرح
 زيادة

In the relationship between the number of studying hours and an exam grade, the number of studying hours is assumed to be

- A) continuous variable B) dependent variable C) nominal variable D) independent variable

The variable that is used to predict the value of another is called ...

- A) independent variable. C) correlation variable.
 B) dependent variable. D) variable of determination.



The variable that is used to predict the value of another is called ...

- A) response variable. C) correlation variable.
 B) explanatory variable. D) variable of determination.

A
L
S
A
A
D
I

A home owner wanted to determine if there was a relationship between the size (in 100 square feet) of a new home and the price (in thousands of dollars) of the home. He found the equation of the regression line and

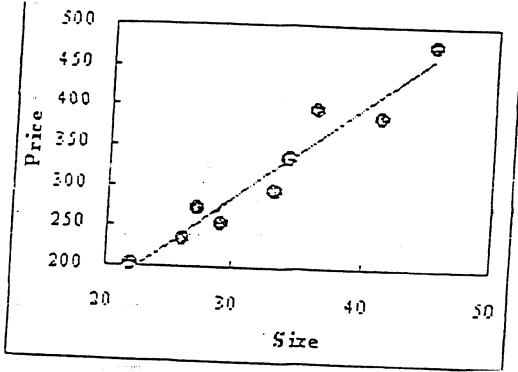
graph the line on a scatter plot as presented below.

$$y' = -58.767 + 11.535x$$

Price

Size

Use this graph to answer the following four questions



What would be the predicted home price y if $x=40$

عوضه عمم $x=40$
من المعادله اعطاه

- A) 461.4 B) 402.63 C) 350 D) 3419.15

$$y' = -58.767 + 11.535(40) = 402.63$$

When the size of the home increases one unit (100 square feet), the expected change in home price is

عندما تزداد x بمقدار one unit
فاره y تزداد بمقدار معامل x
اما تزداد y بمقدار 11.535

- A) 11.535 B) -58.767 C) 0 D) 58.767

The size of the home is called

- A) Explanatory variable C) Outcome variable
B) Response variable D) Confounding variable

Size of home (x) is called:

→ Explanatory variable
or → independent variable.
تغير مستقل

From the graph, the relationship between size and prize is

- A) Strong negative B) Weak positive C) Weak negative D) Strong positive

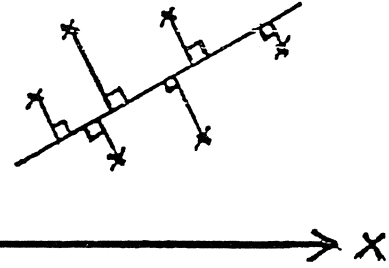
A
L
S
A
A
D
I

The line of best fit means the line that the sum of the ... distances from each point to this line is at a minimum.

- A) squares of the vertical B) vertical C) squares of the horizontal D) horizontal

مجموع مربعات المسافات العمودية
بين النقاط والنقطة (بُعد النقاط عن الخط)
يكون أقل ما يمكن.

تعريف



A graph of the independent variable, X , and the dependent variable, Y , is called ...

- A) scatter plot. B) pie graph. C) histogram. D) frequency polygon.

تعريف

حفظ

In the $y' = a + bx$, what is y' ?

- A) Slope of the regression line.
B) Intercept of the regression line.
C) Predicted value of y , given a specific x value.
D) Value of y when $x = 0$.

من معادله خط الانحدار
eq. of regression line

y' تعني: توقع قيمه y عند قيمه محدده لـ x

$y = a - bx$
من حاله b السالبة المتغير سيم عكس اتجاههم
اقدامهم يزداد الاخر ينقص.

$y = a + bx$
من حاله b الموجبه المتغير سيم مثل اتجاههم
سواء بالزيادة او النقصان.

Complete the following statements with the best answer.

1. The regression line is called the line of best fit

2. If all the points fall on a straight line the value of r will be 1 or -1.

3. The strength and direction of the relationship between two variables is determined by the value of the correlation coefficient.

4. The equation of the regression line used in statistics is $y' = a + bx$

5. The sign of r and b will always be the same.

6. A statistical graph of two variables is called scatter plot.

7. The range of the correlation coefficient is from -1 to 1.

راندري 10 Ch.

كل التمنيات بالانجاح والتوفيق

السعدى

CH. 13 sec. 6**Section 13-6****The spearman rank correlation coefficient**

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

Where: n is number of data pairs.

: d is difference in ranks.

Note

- The range of $r_s = [-1, 1]$ or $-1 \leq r_s \leq 1$
- $d = \text{zero}$ if both sets of data have the same ranks $r_s = \underline{1}$.
- If the sets of data are ranked in the opposite way $r_s = \underline{-1}$.
- If there is no relationship between the two variables $r_s = \underline{0}$.

A
L
S
A
A
D
I

Example:

A statistics instructor wishes to see whether is a relationship between the number of homework exercise a student completes and her or his exam score. The data are shown here.

Using the table to find spearman rank correlation coefficient.

Homework problems x	63	55	58	87	89	52	46	75	105
Exam score y	85	71	75	98	93	63	72	89	100

Solution

Arrange X and y From Low to high

H.W. Problems x	Exam score y	Rank x	Rank y	$d = R_x - R_y$	d^2
63	85	5	5	0	0
55	71	3	2	1	1
58	75	4	4	0	0
87	98	7	8	-1	1
89	93	8	7	1	1
52	63	2	1	1	1
46	72	1	3	-2	4
75	89	6	6	0	0
105	100	9	9	0	0
					$\sum d^2 = 8$

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6(8)}{9(81 - 1)} = \underline{\underline{0.93}}$$

Strong Positive

Example:**For this data:**

X	50	60	24	30	25	35	44	56	37	30
Y	40	37	20	25	19	25	25	42	30	20

Find the spearman rank correlation coefficient.

Solution

Arrange x and y from high to low.....

X	Y	R _x	R _y	d = R _x - R _y	d ²
50	40	3	2	1	1
60	37	1	3	-2	4
24	20	10	8.5	1.5	2.25
30	25	7.5	6	1.5	2.25
25	19	9	10	-1	1
35	25	6	6	0	0
44	25	4	6	-2	4
56	42	2	1	1	1
37	30	5	4	1	1
30	20	7.5	8.5	-1	1
					$\sum d^2 = 17.5$

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} = 1 - \frac{6(17.5)}{10(100 - 1)} = \underline{\underline{0.89}}$$

Strong Positive

Use the following to answer questions

If the differences between the ranks of two variables are $(-1, -3, 1, -2, 1, 3, 2, -1)$ then answer the following three questions:

عجم العينة
The sample size is

- A) 0 B) 30 C) 16 D) 8

عدد العتق = 8 = عجم العينة Sample size

The value of the correlation coefficient is

- A) 0.357 B) -0.357 C) -0.643 D) 0.643

d : -1, -3, 1, -2, 1, 3, 2, -1

d^2 : 1, 9, 1, 4, 1, 9, 4, 1

$$\sum d^2 = 30$$

* Correlation coefficient

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

قانونه سببر مانه
يدعمل من فهايه ورده الاسله

$$= 1 - \frac{6(30)}{8(64 - 1)} \approx \boxed{0.643}$$

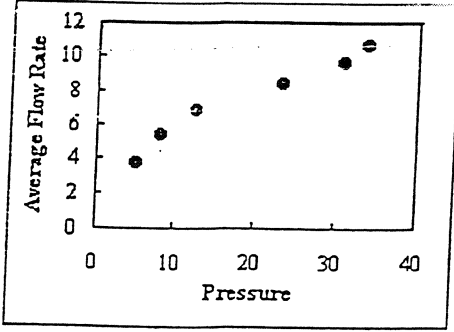
The correlation coefficient value means that there is

- A) very weak linear relationship C) moderate linear relationship
B) very strong linear relationship D) no linear relationship

$r \approx 0.64$ is moderate.

- The regression line can be used to predict a value for the dependent variable (y) for a given value of the independent variable (X).

Determine the type of relationship shown in the figure below.



- A) there is no relationship B) positive C) negative D) curvilinear

واضح أنه العلاقة موجبة (موجبة)

حيث أنه

pressure * كلما زاد
Average flow Rate زاد

If the Spearman rank correlation coefficient (r_s) equals 0.6, then the relationship can be described as

- A) positive, strong and non linear
B) weak and linear
C) positive, moderate and linear
D) moderate and non linear

Use the following to answer questions

X	1	2	3	4
Y	2	5	1	3

The sample size is ...

- A) 4 B) 8 C) 0 D) 18

Sample size = 4 (عدد العينات باكمدول)

The Spearman rank correlation coefficient (r_s) is ...

- A) 1 B) 0 C) -1 D) -0.98

$$n = 4$$

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$= 1 - \frac{6(10)}{4(16 - 1)} = 0$$

X	Y	R _x	R _y	d	d ²
1	2	4	3	1	1
2	5	3	1	2	4
3	1	2	4	-2	4
4	3	1	2	-1	1
					$\sum d^2 = 10$

The value of the Spearman rank correlation coefficient (r_s) means that there is ...

- A) strong negative linear relationship C) strong positive linear relationship
B) moderate linear relationship D) no linear relationship

$\therefore r_s = \text{zero}$ \therefore no linear relationship.
لا توجد علاقة بين المتغيرين X و Y

كل الأمنيات بالنجح والتوفيق