Statistics

Spring 2023

Lecture 12

Feb 19-8:47 AM

Class QZ 1

Consider the Sample below

15 17 20 28 32

25 30 20 18 10

Clear all list.

Store this data in L1

STAT → CALC 1:1-Var Stats

use L1

VARS 5: Statistics 3: Sx x^2 MATH 1:▸Frac Enter

Find

1) $\bar{x} = \boxed{21.5}$ ✓

2) $S = \boxed{7.059}$ ✓ Round to 3-decimal

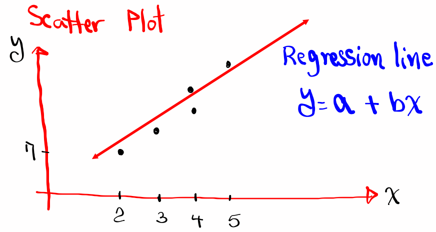
3) $n = \boxed{10}$ ✓

4) $S^2 = \boxed{\frac{299}{6}}$ ✓ Reduced fraction

Feb 23-8:07 AM

Consider the chart below

x	y
2	7
3	10
4	12
4	15
5	20



Clear all lists

$x \to L1, \; y \to L2$ $\Rightarrow$ $\boxed{\text{STAT}}$ CALC $\boxed{\text{2: 2-VarStats}}$ L1, L2

$\Sigma x = 18$

$\Sigma x^2 = 70$

$n = 5$

$\Sigma y = 64$

$\Sigma y^2 = 918$

$\Sigma xy = 252$

Compute $\dfrac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$

$$= \frac{5 \cdot 252 - 18 \cdot 64}{5 \cdot 70 - 18^2} = \frac{108}{26} \approx \boxed{4.154}$$

Compute $\dfrac{\Sigma y \cdot \Sigma x^2 - \Sigma x \cdot \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2} = \dfrac{64 \cdot 70 - 18 \cdot 252}{5 \cdot 70 - 18^2}$

$$= \frac{-56}{26} \approx \boxed{-2.154}$$

Feb 27-7:23 AM

Consider the chart below:

	Study time	QZ Score	
L1	1	7	L2
	2	8	
	2	10	
	3	9	
	4	10	

Study time $\to x \to$ L1

QZ Score $\to y \to$ L2

2-Var Stats

$\Sigma x = 12$

$\Sigma x^2 = 34$

$n = 5$

$\Sigma y = 44$

$\Sigma y^2 = 394$

$\Sigma xy = 110$

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

$$= \frac{44 \cdot 34 - 12 \cdot 110}{5 \cdot 34 - 12^2} = \frac{176}{26} \approx \boxed{6.769}$$

≈ 6.8

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

$$= \frac{5 \cdot 110 - 12 \cdot 44}{5 \cdot 34 - 12^2} = \frac{22}{26} \approx \boxed{.846} \approx .8$$

Regression line

$y = a + bx$

$\boxed{y \approx 6.8 + .8x}$

Feb 27-7:36 AM

Finding Regression line using TI:

STAT → CALC

8: Lin Reg(a+bx)

with Menu:

xList: L1

Ylist: L2

Clear

Calculate

No Menu:

Lin Reg(a+bx)

L1, L2

Enter

$y = a + bx$

$a = 6.769$

$b = .846$

$r^2 = .548$

$r = .740$

If r^2 & r are missing

2nd 0 ↓↓↓ ... ↓ DiagnosticOn

Enter Enter

Feb 27-7:48 AM

r is linear Correlation Coefficient.

$-1 \le r \le 1$

when r is close to 1 or -1,

the linear Correlation is Significant.

when r is close to 0,

the linear Correlation is not Significant.

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

x	y
5	7
6	10
6	12
4	10
5	12
8	15

x → L1, y → L2

use 2-Var stats with L1 & L2

$\sum x = 34$ $\quad \sum y = 66$

$\sum x^2 = 202$ $\quad \sum y^2 = 762$

$n = 6$ $\quad \sum xy = 387$

use STAT CALC 8: LinReg(a+bx) with L1 & L2

$a = 3.107$

$r^2 = .503$

$b = 1.393$

$r = .709$

$y = a + bx$

$y \approx 3.1 + 1.4x$

Feb 27-7:57 AM

Use formula to find linear Correlation Coef.:

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

$\sum x = 34$
$\sum x^2 = 202$
$n = 6$
$\sum y = 66$
$\sum y^2 = 762$
$\sum xy = 387$

$$= \frac{6 \cdot 387 - 34 \cdot 66}{\sqrt{6 \cdot 202 - 34^2}\sqrt{6 \cdot 762 - 66^2}}$$

$$= \frac{78}{\sqrt{56}\sqrt{216}} = \frac{78}{\sqrt{12096}} = \boxed{.709} \leftarrow r$$

78 $\boxed{\div}$ $\boxed{2nd}$ $\boxed{x^2}$ 12096 $\boxed{enter}$

r appears to be close to 1, Linear Correlation is Significant.

Find $(.709)^2 \to .503$

r^2

Always in % $\Rightarrow r^2 \approx 50\%$

Coef. of determination. It tells us % of Y-values are explained by x-values.

Feb 27-8:08 AM

QZ Scores	Exam Scores
8	85
7	80
10	95
9	90
6	75
5	60

QZ Scores $\to x \to L1$

Exam Scores $\to y \to L2$

Use LinReg(a+bx) with L1 & L2 to find

$a = 32.619$ $r^2 = .938$

$b = 6.429$ $r = .969$

$a \approx 33$, $b \approx 6$

$\boxed{y = 33 + 6x}$

r^2 (%)

$\boxed{r^2 \approx 94\%}$

r appears to be very close to 1, therefore It is significant

94% of exam Scores are explained by QZ Scores.

6% unexplained

Feb 27-8:20 AM

How to make Predictions:

If r is significant

$\Rightarrow$ use the regression line

$y = a + bx$

If r is not significant

$\Rightarrow$ use $\bar{y}$

$\bar{y} = \frac{\sum y}{n}$ or $\boxed{\text{VARS}}$ $\boxed{\text{5: Statistics}}$ $\boxed{5\text{: } \bar{y}}$ $\boxed{\text{Enter}}$

Feb 27-8:30 AM