Statistics

Lecture 8

Feb 19-8:47 AM

Class QZ 3

Consider the chart below

class MP	class F
17	3
25	7
33	15
41	10
49	5

Find

1) $\bar{x} = 34.4 \approx 34$

2) $S = S_x = 8.854 \approx 9$

Round to whole numbers

3) $n = 40$

4) $S^2 = \frac{392}{5}$

Reduced fraction

Class MP → L1

Class F → L2

STAT → CALC

1:1-Var Stats

With Menu

List: L1

FreqList: L2

Calculate

No Menu

1-Var Stats

L1 , L2

Enter

VARS | 5: Statistics | 3: S_x

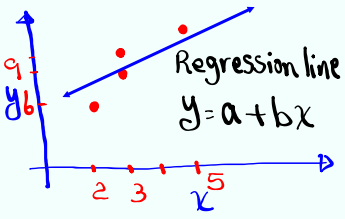
x^2 | math | 1: ▸frac | Enter

Sep 29-7:47 AM

Complete the chart below

x	y	x^2	y^2	xy
2	6	4	36	12
3	9	9	81	27
3	10	9	100	30
5	12	25	144	60

draw Scatter plot



Clear all lists

$x \to L1$, $y \to L2$

STAT → CALC

2: 2-Var Stats

$\Sigma x = 13$ $\Sigma y = 37$

$\Sigma x^2 = 47$ $\Sigma y^2 = 361$

$n = 4$ $\Sigma xy = 129$

with Menu

Xlist: L1

Ylist: L2

FreqList: clear

Calculate

No Menu

2-Var Stats

L1 , L2

1

enter

Sep 29-10:52 AM

How to find eqn of the regression line

STAT → CALC

8: LinReg(a+bx)

with Menu

Xlist: L1

Ylist : L2

clear

Calculate

No Menu

L1 , L2

1

enter

You Should See

$a = 3.263 \approx 3$

$b = 1.842 \approx 2$ $\Rightarrow$ Regression line $y = 3 + 2x$

$r^2 = .860$

$r = .927$

Sep 29-11:06 AM

$\Sigma x = 13$ $\Sigma y = 37$ Formulas for a & b

$\Sigma x^2 = 47$ $\Sigma y^2 = 361$

$n = 4$ $\Sigma xy = 129$

$$a = \frac{\Sigma y \Sigma x^2 - \Sigma x \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2} = \frac{37 \cdot 47 - 13 \cdot 129}{4 \cdot 47 - 13^2} = \frac{62}{19}$$

$= 3.263$

$\approx \boxed{3}$

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

$$= \frac{4 \cdot 129 - 13 \cdot 37}{4 \cdot 47 - (13)^2} = \frac{35}{19} = 1.842 \approx \boxed{2}$$

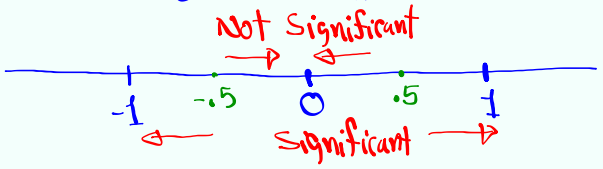
Sep 29-11:12 AM

r → Linear Correlation Coefficient

$$-1 \le r \le 1$$

If r is close to 1 or −1,
the linear Correlation is Significant.

If r is close to 0,
the linear Correlation is not Significant



From Last example

$r = .927$ It is close to 1,
Linear Correlation is Significant

Sep 29-11:18 AM

Formula for r:

$\Sigma x = 13$ $\quad \Sigma y = 37$

$\Sigma x^2 = 47$ $\quad \Sigma y^2 = 361$

$n = 4$ $\quad \Sigma xy = 129$

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

$$= \frac{4 \cdot 129 - 13 \cdot 37}{\sqrt{4 \cdot 47 - 13^2}\ \sqrt{4 \cdot 361 - 37^2}} = \frac{35}{\sqrt{19}\ \sqrt{75}}$$

$$= \frac{35}{\sqrt{1425}} = .927$$

35 ÷ [2nd] [x^2] 1425 [enter]

Sep 29-11:23 AM

Now what about r^2?

r^2 → Always in whole %.

→ Coefficient of determination

r^2 tells us what % of Y-values are explained by x-values.

From Last example

$r = .927$

$r^2 = (.927)^2 = .859 \approx 86\%$

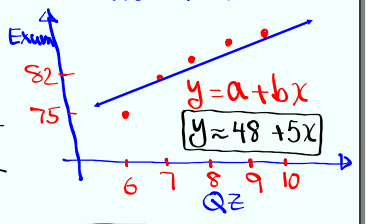
86% of Y-values are explained by x-values.

Sep 29-11:29 AM

5 students were randomly selected, chart below shows Qz Scores & Exam Scores.

Qz Scores	Exam Scores
7	82
8	85
9	90
10	94
6	75

Scatter Plot



$y = a + bx$

$\hat{y} \approx 48 + 5x$

Clear all lists

Qz Scores → X → L1

Exam Scores → Y → L2

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

$a = 48.4 \approx 48$

$b = 4.6 \approx 5$

$r^2 \approx 99\%$

99% of exam scores are explained by Qz scores.

$r^2 = .985$

$r = .993$

r is close to 1, Linear Correlation is Significant.

Sep 29-11:33 AM

How to make predictions:

If r is Significant

⇒ use the regression line to make predictions.

If r is not Significant

⇒ use $\bar{y}$ as the prediction value

$\bar{y} = \frac{\Sigma y}{n}$ OR [VARS] 5: Statistics 5: $\bar{y}$ Enter

$\bar{y} = 85.2$

≈ 85

Sep 29-11:46 AM

Jose got 8 on the quiz, Predict his exam score

1) Assume r is Significant

$y = a + bx$

$= 48 + 5(8) = 48 + 40 = \boxed{88}$

2) Assume r is not Significant.

use $\bar{y}$

VARS

5: Statistics

5: $\bar{y}$ Enter

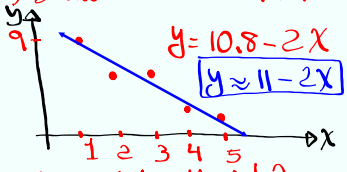
$\bar{y} = 85.2 \approx \boxed{85}$

Sep 29-11:51 AM

Clear all lists

x	y
1	9
2	6
3	6
4	2
5	1

1) Draw Scatter Plot



$y = 10.8 - 2x$

$y \approx 11 - 2x$

2) $x \to L1$, $y \to L2$

3) Use LinReg(a+bx) with L1 & L2

$a = 10.8$

$b = -2$

$r^2 = .9345 \approx 93\%$

93% of Y-Values are explained by X-Values.

$r = -.967$

r is close to -1, Linear Correlation is Significant

Sep 29-11:55 AM

Predict Y if x is 3.

1) Assume r is Significant.
Use regression line
$\hat{y} \approx 11 - 2(3) = 11 - 6 = \boxed{5}$

2) Assume r is not Significant.
Use $\bar{y}$
$\bar{y} = 4.8$
$\approx \boxed{5}$

VARS
5: Statistics
5: $\bar{y}$
Enter

Sep 29-12:06 PM

use 2-Var Stats with L1 & L2 to find

$\sum x = 15$ $\quad \sum y = 24$

$\sum x^2 = 55$ $\quad \sum y^2 = 158$

$n = 5$ $\quad \sum xy = 52$

use formulas to find a & b.

$$a = \frac{\sum y \cdot \sum x^2 - \sum x \cdot \sum xy}{n\sum x^2 - (\sum x)^2} = \frac{24 \cdot 55 - 15 \cdot 52}{5 \cdot 55 - 15^2} = \frac{540}{50} = \boxed{10.8}$$

$$b = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{5 \cdot 52 - 15 \cdot 24}{5 \cdot 55 - 15^2} = \frac{-100}{50} = \boxed{-2}$$

Sep 29-12:09 PM

Now use formula to find r

$\sum x = 15$ $\quad \sum y = 24$

$\sum x^2 = 55$ $\quad \sum y^2 = 158$

$n = 5$ $\quad \sum xy = 52$

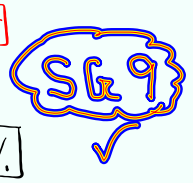
$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}} = \frac{5 \cdot 52 - 15 \cdot 24}{\sqrt{5 \cdot 55 - 15^2}\sqrt{5 \cdot 158 - 24^2}}$$

$$= \frac{-100}{\sqrt{50}\sqrt{214}} = \frac{-100}{\sqrt{10700}} \approx \boxed{-.967}$$

-100 ÷ end x^2 10700 enter

Find r^2 in whole percent

$r^2 = (-.967)^2 = .935 \approx \boxed{94\%}$



Sep 29-12:19 PM