Statistics

Lecture 7

Feb 19-8:47 AM

Class Quiz 5

use the table below

class MP	class F
16	8
25	12
34	20
43	10

Find

1) $\bar{x} = 30.76 \approx \boxed{31}$ } Round to whole #

2) $S = 8.863 \approx \boxed{9}$ }

3) $n = 50$

4) $S^2 = \frac{96228}{1225}$ } Reduced fraction

Clear All lists

class MP → L1

class F → L2

STAT → CALC

1: 1-Var Stats

with Menu

List: L1

Freq List: L2

Calculate

No Menu

L1, L2

Enter

VARS | 5: Statistics

3: S_x | x^2 | MATH | 1: ▶Frac

Enter

Sep 18-12:08 PM

Consider the Stem Plot below

3	0 2 5
4	2 3 5 5 7 8
5	0 2 4 4 8 8 8 9
6	2 3 4 5 5 5 6 8 9
7	0 3 5 8 8 9
8	0 5 5

1) $n = 35$

2) Mode = 58 & 65

3) Range = 85 − 30 = 55

4) Estimate $S = \frac{\text{Range}}{4} = \frac{55}{4} = 13.75$

5) Find P_{20}

$L = \frac{20}{100} \cdot 35 = 7$ $P_{20} = \frac{7\text{th} + 8\text{th}}{2} = \frac{45+46}{2} = 46$

20%, 80%
46

6) Find Median

Median = P_{50} $= P_{50} = 18\text{th element} = 62$

$L = \frac{50}{100} \cdot 35 = 17.5$ $L = 18$

7) Find K such that $P_K = 70$

74%, 26%
70

$K = \frac{B}{n} \cdot 100 = \frac{26}{35} \cdot 100 = 74.2857\ldots \approx 74$

Sep 18-12:30 PM

Complete the Chart below

x	y	x^2	y^2	xy
2	3	4	9	6
3	5	9	25	15
4	6	16	36	24
5	10	25	100	50
8	12	64	144	96

$\Sigma x = 22$ $\Sigma y = 36$

$\Sigma x^2 = 118$ $\Sigma y^2 = 314$

$n = 5$ $\Sigma xy = 191$

Clear all lists

$x \rightarrow L1$

$y \rightarrow L2$

STAT → CALC

2: 2-Var Stats

With Menu

Xlist: L1

Ylist: L2

FreqList: clear

Calculate

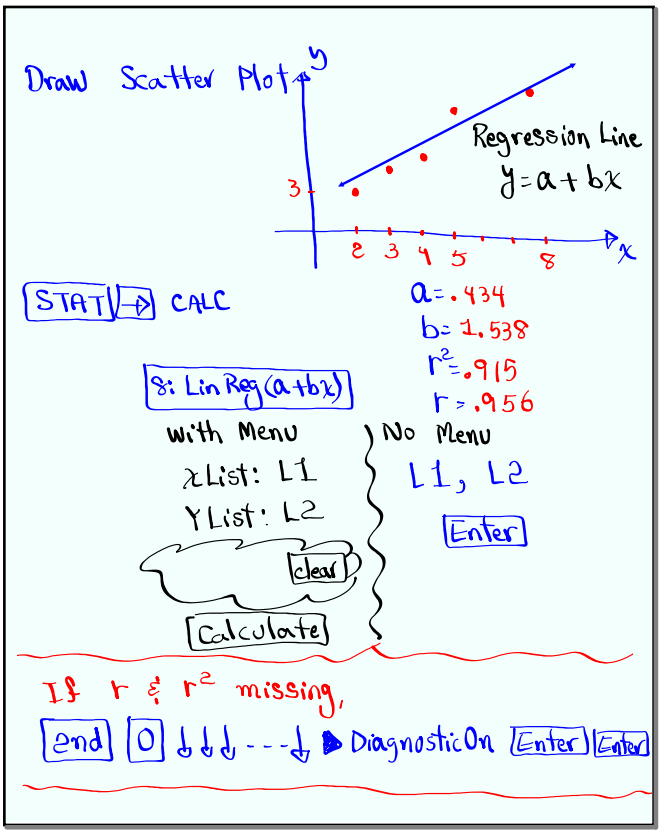
No Menu

L1, L2

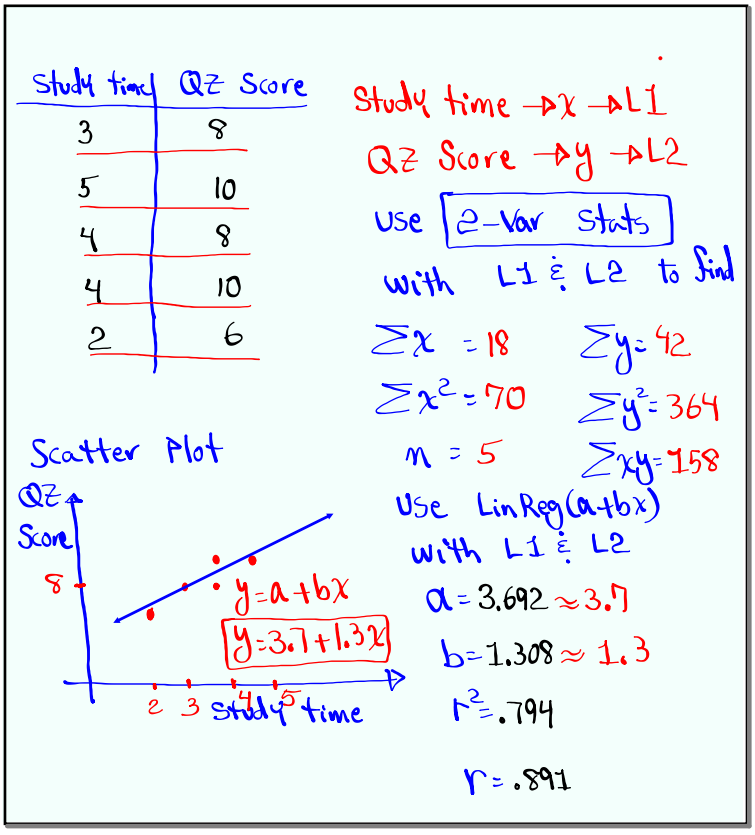
[ ]

Enter

Sep 18-12:44 PM



Sep 18-12:54 PM



Sep 18-1:05 PM

How to find a & b for the regression line:

$y = a + bx$

x	y
1	3
2	8
4	10
5	10

Find

$\Sigma x = 12$ $\quad \Sigma y = 31$

$\Sigma x^2 = 46$ $\quad \Sigma y^2 = 273$

$n = 4$ $\quad \Sigma xy = 109$

$$a = \frac{\Sigma y \, \Sigma x^2 - \Sigma x \, \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2} = \frac{31 \cdot 46 - 12 \cdot 109}{4 \cdot 46 - 12^2} = \frac{118}{40} = 2.95 \checkmark$$

$$b = \frac{n\Sigma xy - \Sigma x \, \Sigma y}{n\Sigma x^2 - (\Sigma x)^2} = \frac{4 \cdot 109 - 12 \cdot 31}{4 \cdot 46 - 12^2} = \frac{64}{40} = 1.6 \checkmark$$

Use Lin Reg(a+bx) with L1 & L2

$a = 2.95$ ✓

$b = 1.6$ ✓

$r^2 = .782$

$r = .884$ ✓

Sep 18-1:33 PM

what is r? what does it to us?

$r \rightarrow$ Linear Correlation Coefficient

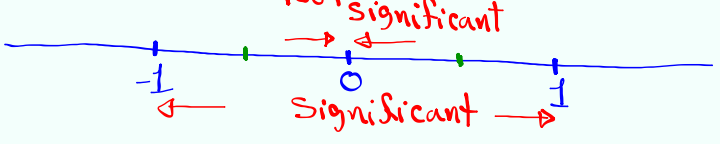
$$-1 \leq r \leq 1$$

when r is close to 1 or -1,

Linear Correlation is Significant

when r is close to 0,

Linear Correlation is not Significant



Sep 18-1:43 PM

How to find r:

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

$\Sigma x = 12$ $\quad \Sigma y = 31$

$\Sigma x^2 = 46$ $\quad \Sigma y^2 = 273$

$n = 4$ $\quad \Sigma xy = 109$

$$r = \frac{4 \cdot 109 - 12 \cdot 31}{\sqrt{4 \cdot 46 - 12^2}\ \sqrt{4 \cdot 273 - 31^2}}$$

64 ÷ [end] [x²] 5240 [Enter] $= \frac{64}{\sqrt{40}\sqrt{131}}$

Since .884 close to 1,

Linear Correlation is Significant

$= \frac{64}{\sqrt{5240}} = .884$

Sep 18-1:48 PM

What about r^2? What does it say to us?

r^2 is the Coefficient of determination.

Always round to whole %

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

Last example

$r^2 = .782 \approx 78\%$

78% of Y-values are explained by x-values

Sep 18-1:55 PM

Quiz Score	Exam Score
8	85
7	80
10	90
6	70

Regression Line

$y \approx 45 + 5x$

$r^2 \approx 89\%$

89% of exam scores are explained by Quiz Scores

Quiz Scores → x → L1

Exam Scores → y → L2

use LinReg(a+bx) with L1 & L2

$a = 44.714 \approx 45$

$b = 4.714 \approx 5$

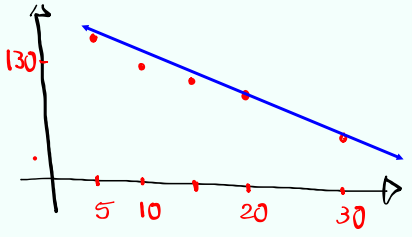
$r^2 = .889$

$r = .943$

↳ close to 1 Linear Correlation is significant.

Sep 18-1:59 PM

Walk time	BS level
10	130
20	115
30	100
15	125
5	140



Walk time → x → L1

BS level → y → L2

use Lin Reg (a+bx) with L1 & L2

$y = 147 - 2x$

$a = 147.297 \approx 147$

$b = -1.581 \approx -2$

$r^2 = .9945 \approx 99\%$

Close to -1 ⟶ $r = -.997$

Linear Correlation is Significant

99% of BS level are explained by walking time

Sep 18-2:08 PM

How to make prediction:

If r is significant,

Use regression line

Plug in x, find y

If r is not significant,

Use $\bar{y}$

$\bar{y} = \frac{\Sigma y}{n}$

VARS

Statistics

5: $\bar{y}$

Enter

$\bar{y} = 122$

Sep 18-2:17 PM

Given $n = 10$, $\Sigma y = 125$,

$\hat{y} = 8 + 2.5x$

Predict y when $x = 4$

1) Assume Linear Correlation is Significant.

Use Regression line $\hat{y} = 8 + 2.5x$

$= 8 + 2.5(4) = \boxed{18}$

2) Assume Linear Correlation is not Significant.

$\bar{y} = \frac{\Sigma y}{n} = \frac{125}{10} = \boxed{12.5}$

Sep 18-2:21 PM

Exam 1	Final Exam
75	88
70	80
65	70
90	90
100	95
85	80

Exam 1 → x → L1

Final Exam → y → L2

use Lin Reg(a+bx)

with L1 & L2

$a = 38.3$ $r^2 = 69\%$

$b = .6$ $r = .831$

69% of final exam scores are explained by exam 1 scores.

Sep 18-2:25 PM

Predict final exam score for someone who got 80 on exam 1.

1) Linear Correlation is Significant

$y = 38.3 + .6x = 38.3 + .6(80) \approx 86$

2) Linear Correlation is not Significant — use $\bar{y}$

VARS

5: Statistics

5: $\bar{y}$ Enter

83.8 ≈ 84

Sep 18-2:31 PM

Class Quiz 5

x	y
2	8
3	10
4	10
5	15
8	20

Find

$a \approx \boxed{3.5}$

$b \approx \boxed{2.1}$

} Round to 1-dec.

$r^2 \approx \boxed{95\%}$

} whole %.

$r \approx \boxed{.975}$

} 3-dec.

Sep 18-2:36 PM