Statistics
Summer 2023
Lecture 5

Feb 19-8:47 AM

Class QZ 5

Consider the Sample below

12	18	15	15	10
10	8	20	16	9

✓ $\bar{x} = 13.3$ } Round to 3-decimal

✓ $S = 4.084$ } Round to 3-decimal

✓ $S^2 = \frac{1501}{90}$ } Reduced Fraction

Use your TI calc to find

$n = 10$

Clear all lists

Reset all lists

STAT Edit

1:Edit

L1
12
18
15
⋮
9

STAT → CALC

1-Var Stats

L1 Enter

List: L1

FreqList: clear

Calculate

VARS 5: Statistics 3: S_x x^2 Math 1: ▶Frac Enter

Jun 15-11:40 AM

SG 9

Complete the chart below

x	y	x^2	y^2	$x \cdot y$
1	3	1	9	3
2	4	4	16	8
2	5	4	25	10
4	8	16	64	32

1) $n = 4$

2) $\Sigma x = 9$ $\quad \Sigma x^2 = 25$

3) $\Sigma y = 20$ $\quad \Sigma y^2 = 114$

4) $\Sigma xy = 53$

Clear all lists

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

L1	L2
1	3
2	4
2	5
4	8

STAT → CALC

2: 2-Var Stats

No Menu

L1, L2 enter

Enter

$\Sigma x = 9$ $\quad \Sigma y = 20$

$\Sigma x^2 = 25$ $\quad \Sigma y^2 = 114$

$n = 4$ $\quad \Sigma xy = 53$

With Menu

x List: L1

y list: L2

FreqList: clear

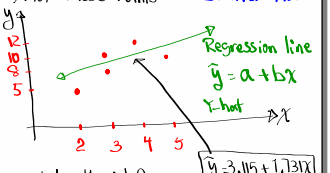
Calculate

Jun 20-7:39 AM

Working with ordered Pairs: (x, y)

x	y
2	5
3	8
3	10
4	12
5	10

1) Plot these Points Scatter Plot



$\hat{y} = 3.115 + 1.731x$

Clear all lists, $x \to L1$, $y \to L2$

Use 2-Var Stats with L1 & L2 to find

$\Sigma x = 17$ $\quad \Sigma y = 45$ $\quad \hat{y} \approx 3 + 2x$

$\Sigma x^2 = 63$ $\quad \Sigma y^2 = 433$

$n = 5$ $\quad \Sigma xy = 162$

STAT → CALC

8: LinReg(a+bx)

No Menu

L1, L2 enter

Enter

With Menu:

x list: L1

Y list: L2

Clear

Calculate

$y = a + bx$

$a = 3.115$

$b = 1.731$ } 3 decimal Places

If missing { $r^2 = .556$ / $r = .746$

2nd 0 ↓↓↓---↓ ▸Diagnostics On Enter Enter

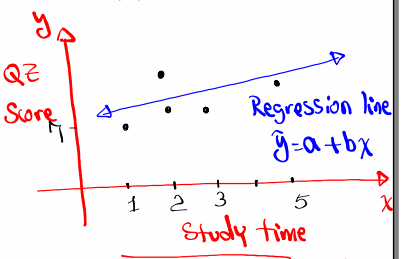
Now redo the last step to view r & r^2.

Jun 20-7:50 AM

Consider the chart below

Study time	QZ Score
1	7
2	8
2	10
3	8
5	9

Scatter Plot



Study time → x → L1

QZ Score → y → L2

Use [2-Var Stats] with L1 & L2 to find

$\Sigma x = 13$ $\Sigma y = 42$

$\Sigma x^2 = 43$ $\Sigma y^2 = 358$

$n = 5$ $\Sigma xy = 112$

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

$a = 7.6$ ✓

$b = .3$ ✓

} 1-decimal → $\hat{y} = a + bx$ → $\hat{y} \approx 7.6 + .3x$

$r^2 = 16\%$ } whole %

$r = .405$ } 3-decimal

Jun 20-8:05 AM

How to find a & b using formula:

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

$n = 5$, $\Sigma x = 13$, $\Sigma x^2 = 43$, $\Sigma y = 42$, $\Sigma y^2 = 358$, $\Sigma xy = 112$

$$a = \frac{42 \cdot 43 - 13 \cdot 112}{5 \cdot 43 - 13^2} = \frac{350}{46} = 7.609 \approx 7.6$$

$$b = \frac{n \Sigma xy - \Sigma x \Sigma y}{n \Sigma x^2 - (\Sigma x)^2} = \frac{5 \cdot 112 - 13 \cdot 42}{5 \cdot 43 - 13^2} = \frac{14}{46}$$

$$= .304 \approx .3$$

Jun 20-8:17 AM

Given $n=10$, $\Sigma x=64$, $\Sigma x^2=464$,

$\Sigma y=782$, $\Sigma y^2=62632$, $\Sigma xy=5277$

Find eqn of the regression line

$$\hat{y}=a+bx$$

$$a=\frac{\Sigma y\,\Sigma x^2-\Sigma x\,\Sigma xy}{n\Sigma x^2-(\Sigma x)^2}=\frac{782\cdot 464-64\cdot 5277}{10\cdot 464-64^2}=\frac{25120}{544}\approx \boxed{46}$$

$$b=\frac{n\Sigma xy-\Sigma x\,\Sigma y}{n\Sigma x^2-(\Sigma x)^2}=\frac{10\cdot 5277-64\cdot 782}{10\cdot 464-64^2}=\frac{2722}{544}\approx \boxed{5}$$

$$\hat{y}\approx a+bx$$

$$\boxed{\hat{y}\approx 46+5x}$$

Jun 20-8:25 AM

Regression line $\hat{y}=a+bx$

Linear Correlation Coefficient r

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

x	y
2	5
3	8
4	10
5	10

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

use [2-Var Stats]

$\Sigma x=14$ $\Sigma y=33$

$\Sigma x^2=54$ $\Sigma y^2=289$

$n=4$ $\Sigma xy=124$

$$r=\frac{n\Sigma xy-\Sigma x\,\Sigma y}{\sqrt{n\Sigma x^2-(\Sigma x)^2}\sqrt{n\Sigma y^2-(\Sigma y)^2}}=\frac{4\cdot 124-14\cdot 33}{\sqrt{4\cdot 54-14^2}\sqrt{4\cdot 289-33^2}}=\frac{34}{\sqrt{20}\sqrt{67}}$$

$$=\frac{34}{\sqrt{1340}}=\boxed{.929}$$

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

$a=2.3$ $\Rightarrow \hat{y}=2.3+1.7x$

$b=1.7$

$r^2\approx 86\%$

$r=.929$ 3-decimal Places

Jun 20-8:50 AM

Linear Correlation Coefficient r

It is a numerical value that measures the strength of linear correlation between all ordered-pairs (x, y).

$-1 \le r \le 1$

when r is close to 0

⇒ Linear Correlation is not Significant.

when r is close to 1 or -1

⇒ Linear Correlation is Significant.

what is r^2? It is the Coef. of determination.

Always express as whole %.

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

From Last example → $r^2 \approx 86\%$.

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

14% were unexplained.

Jun 20-9:03 AM

Suppose $r = .959$ for correlation between x: study time and y: exam results.

1) Since r is close to 1,

Linear correlation is Significant.

2) find r^2 in %. $r^2 = (.959)^2 \approx 92\%$.

So 92% of exam scores are explained by study time. 8% are unexplained.

Jun 20-9:11 AM

I randomly selected 8 people, chart below is walking time & Blood Sugar level the next day.

Time	BS level
30	115
35	110
40	100
20	130
20	125
10	145
25	130
40	105

Time → x → L1

BS level → y → L2

Use LinReg(a+bx)

a = 157.8

b = -1.4 } 1-decimal

$\hat{y} \approx 157.8 - 1.4x$

$r^2 = 95\%$ } whole %

$r = -.972$ } 3-decimals

95% of BS level are explained by walking time

5% unexplained

Since r is close to -1, linear Correlation is Significant.

Jun 20-9:16 AM

How to make predictions:

r

1) is Significant

⇒ use $\hat{y} = a + bx$

Plug in x-value, find y-value

2) is not Significant

⇒ use $\bar{y}$

Jun 20-9:26 AM

Given $\hat{y} = 55.6 + 2.8x$ & $\bar{y} = 71.5$

Predict y when $x = 5.5$.

1) Assume r is Significant

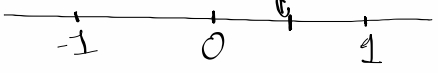
$y = 55.6 + 2.8(5.5)$

$= 71$

2) Assume r is not Significant.

use $\bar{y}$ $\quad y = 71.5$

Suppose $r = .5$



Is r close to 1 or 0? we test this much later in this class.

Jun 20-9:29 AM

Suppose $\hat{y} = 73.5 - 4.5x$ & $\bar{y} = 58.5$

Predict y when $x = 4.8$.

1) Assume r is Significant

$y \approx 73.5 - 4.5(4.8) \approx 51.9$

2) Assume r is not Significant

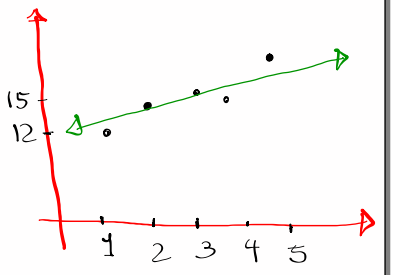
use $\bar{y}$

$y = 58.5$

Jun 20-9:35 AM

Study time	Qz Score
2	15
3	18
4	16
5	20
1	12

Scatter Plot



Study time → X → L2

Qz Score → Y → L2

Use [Lin Reg(a+bx)]

Regression line

$\hat{y} \approx 11 + 2x$

79% of Scores are explained by Study time.

21% unexplained.

$a = 11.1 \approx 11$

$b = 1.7 \approx 2$

$r^2 \approx 79\%$

$r = .886$

Jun 20-9:40 AM

Predict Qz Score for Someone who Studied 4.5 hrs.

1) Assume r is Significant.

$y \approx 11 + 2(4.5) = \boxed{20}$

2) Assume r is not Significant.

use $\bar{y}$

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

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

$\bar{y} = 16.2$ $\boxed{y \approx 16}$

SG 9 ✓

Jun 20-9:47 AM

Intro. To Probabilities: SG 10

E → Desired event (outcome)

P(E) → Prob. that E happens

$$P(E) = \frac{\text{Total \# of all desired outcome}}{\text{Total \# of all outcomes}}$$

Ex: There are 20 students,
12 Females & 8 Males

If we randomly select one student,

find the prob. of selecting one female.

$$P(\text{Female}) = \frac{12}{20} = .6 = \frac{3}{5}$$

Jun 20-10:32 AM

In a standard deck of playing cards,
there are 52 cards, 26 Red, 12 face cards,
and 4 aces.

If we randomly select one card,

1) $P(\text{Red}) = \frac{26}{52} = \frac{1}{2} = .5$

12 ÷ 52 Enter
MATH 1:▶Frac Enter

2) $P(\text{face}) = \frac{12}{52} = .231 = \frac{3}{13}$

3) $P(\text{ace}) = \frac{4}{52} = .077 = \frac{1}{13}$

Acceptable Form of answer:
1) Reduced fraction
2) Round to 3-decimal places
3) Scientific Notation

Jun 20-10:37 AM

I Surveyed 80 people I asked them if they were in favor of abortion law.

	Yes	No	Total
Males	25	5	30
Females	10	40	50
Total	35	45	80

If we randomly Select one of these people,

1) $P(\text{Male}) = \frac{30}{80} = .375$

2) $P(\text{Yes}) = \frac{35}{80} = \frac{7}{16}$

3) $P(\text{Male and Yes}) = \frac{25}{80} = \frac{5}{16} = .313$

4) $P(\text{Male or Yes}) = \frac{40}{80} = \frac{1}{2} = .5$

Jun 20-10:44 AM

A 20-Sided fair object has numbers 1 to 20 on the Sides. 1, 2, 3, 4, - - - - , 20.

If we roll this object, find the prob. of getting

1) 4 $\frac{1}{20}$

2) less than 4 (1, 2, 3) $\frac{3}{20}$

3) at most 4 (1, 2, 3, 4) $\frac{4}{20} = \frac{1}{5} = .2$

4) at least 15. (15, 16, 17, 18, 19, 20) $\frac{6}{20} = \frac{3}{10} = .3$

5) at most 4 or at least 15 (1, 2, 3, 4 / 15, 16, 17, 18, 19, 20) $\frac{10}{20} = .5$

5) At most 4 and at least 15 $\frac{0}{20}$

Impossible

Jun 20-10:54 AM

Class QZ 6

Use the chart below

x	y
3	2
5	5
6	4
6	3
8	2

Find

1) $a = 3.45$ } Round to 2-decimals

2) $b = -.05$ } Round to 2-decimals

3) $r^2 = 0\%$ } whole %.

4) $r = -.063$ } 3-decimals

Jun 20-11:03 AM