Statistics

Lecture 4

Feb 19-8:47 AM

Consider the chart below

Class limits	Class F
10 — 19	4
20 — 29	6
30 — 39	10
40 — 49	7
50 — 59	5

1) 5 classes

2) Class width = 20 - 10 = 10
$\qquad$ = 30 - 20 = 10
$\qquad$ = 50 - 40 = 10

3) $n = \sum f = 4 + 6 + 10 + 7 + 5 = 32$

4) Find $\bar{x}$ & S

Class MP	Class F
14.5	4
24.5	6
34.5	10
44.5	7
54.5	5

Class M.P. = $\dfrac{+ \text{ Class limits}}{2} = \dfrac{10 + 19}{2}$
$\qquad = \dfrac{29}{2}$
$\qquad = 14.5$

Store class MP → L1,
$\qquad$ Class F → L2

Stat → CALC

1:1-Var Stats

use L1 & L2

$\bar{x} = 35.4375$
$S = S_x = 12.536$
$n = 32$

σ → Sigma
← Population
standard deviation

Find S^2 in reduced fraction.

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

$\dfrac{38975}{248}$

Jan 14-4:31 PM

Work with ordered-pairs
(x, y)

x	y
1	3
2	5
2	8
3	8
4	10

1) $n = 5$ # of ordered-pairs

2) Plot these points

Regression Line
$y = ax + b$

$\boxed{y = a + bx}$

$y \approx 1.5 + 2.2x$

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

x	y	x^2	y^2	xy
1	3	1	9	3
2	5	4	25	10
2	8	4	64	16
3	8	9	64	24
4	10	16	100	40

$\sum x = 12$ $\sum y = 34$

$\sum x^2 = 34$ $\sum y^2 = 262$

$n = 5$ $\sum xy = 93$

$\boxed{STAT}$ $\boxed{\to}$ CALC

2:2-Var Stats

Menu	No Menu
Xlist: L1	L1, L2
Ylist: L2	Enter
Freqlist: Clear	
Calculate	

↓ How to find $y = ax + b$ using TI

$a = 1.538$

$b = 2.192$

$\boxed{STAT}$ $\boxed{\to}$ CALC

xlist: L1

Ylist: L2

4:

8: LinReg$(a + bx)$

Clear

Calculate

Jan 14-4:46 PM

I randomly selected 6 students, I have study time in hrs & Quiz scores (out of 10 pts)

Study time	Qz Score
2	6
3	8
3	7
4	8
1	5
5	10

1) $n = 6$

2) Regression line
$y = a + bx$

$\boxed{y = 3.7 + 1.2x}$

Qz Score

Study time

$\boxed{STAT}$ $\boxed{\to}$ CALC

2:2-Var Stats

L1 & L2

study time $\to x \to L1$

Qz Score $\to y \to L2$

use

$\boxed{STAT}$ $\boxed{\to}$ CALC

8: LinReg$(a + bx)$

with L1 & L2

$\sum x = 18$ $\sum y = 44$

$\sum x^2 = 64$ $\sum y^2 = 338$

$n = 6$ $\sum xy = 144$

$a = 3.7\overline{3} \approx 3.7$

$b = 1.2 = 1.2$

$\boxed{r^2 = .939}$

$\boxed{r = .969}$

$\boxed{2nd}$ $\boxed{0}$ ↓↓↓ … ↓ ▶DiagnosticOn

$\boxed{Enter}$ $\boxed{Enter}$

Jan 14-5:02 PM

2

$r \rightarrow$ Linear Correlation Coefficient

$-1 \leq r \leq 1$

when r is close to ± 1, then

Linear Correlation is Significant.

when r is close to 0, then

Linear Correlation is not Significant

what about r^2? Always display as whole %.

Coefficient of determination.

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

Last example:

$r = .969 \rightarrow r$ is close to $1 \rightarrow$ Linear Correlation
between study time & QZ Scores
Seem to be Significant.

$r^2 = .939$

$\rightarrow r^2 \approx 94\% \rightarrow 94\%$ of QZ Scores are explained by study time.

Jan 14-5:14 PM

Two branches in Statistics:

1) Descriptive

2) Inferential

Let's make Predictions using linear Model:

when r is Significant

$\Rightarrow$ Use regression line eqn to make Prediction.

when r is not Significant

$\Rightarrow$ Use $\bar{y}$

$\bar{y} = \dfrac{\sum y}{n}$ or [VARS] [5: Statistics]

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

Jan 14-5:22 PM

I randomly selected 5 students, I have Qz Scores and exam Scores

Qz Score	Exam Score
8	92
7	85
9	90
6	75
10	98

Qz Score $\to x \to L1$

Exam Score $\to y \to L2$

$\sum x = 40$ $\sum y = 440$

$\sum x^2 = 330$ $\sum y^2 = 39078$

$n = 5$ $\sum xy = 3571$

find   Use LinReg(a + bx)

$a = 47.2 \approx 47$ $\Rightarrow$ $\boxed{y \approx 47 + 5x}$

$b = 5.1 \approx 5$

$r^2 = .873 \Rightarrow r^2 \approx 87\%$ Coef. of determination

$r = .934$

r is linear Correlation Coef.

r is close to 1

$\Rightarrow$ Linear Correlation between Qz Scores and exam Scores Seem to be Significant.

87% of exam Scores are explained by Qz Scores.

Jan 14-5:28 PM

William got 8 on the Quiz.
Predict his exam Score.

1) Assume r is Significant

Use regression line

$y = 47 + 5x$

$y \approx 47 + 5(8) = 47 + 40 \approx \boxed{87}$

2) Assume r is not Significant

Use $\bar{y}$

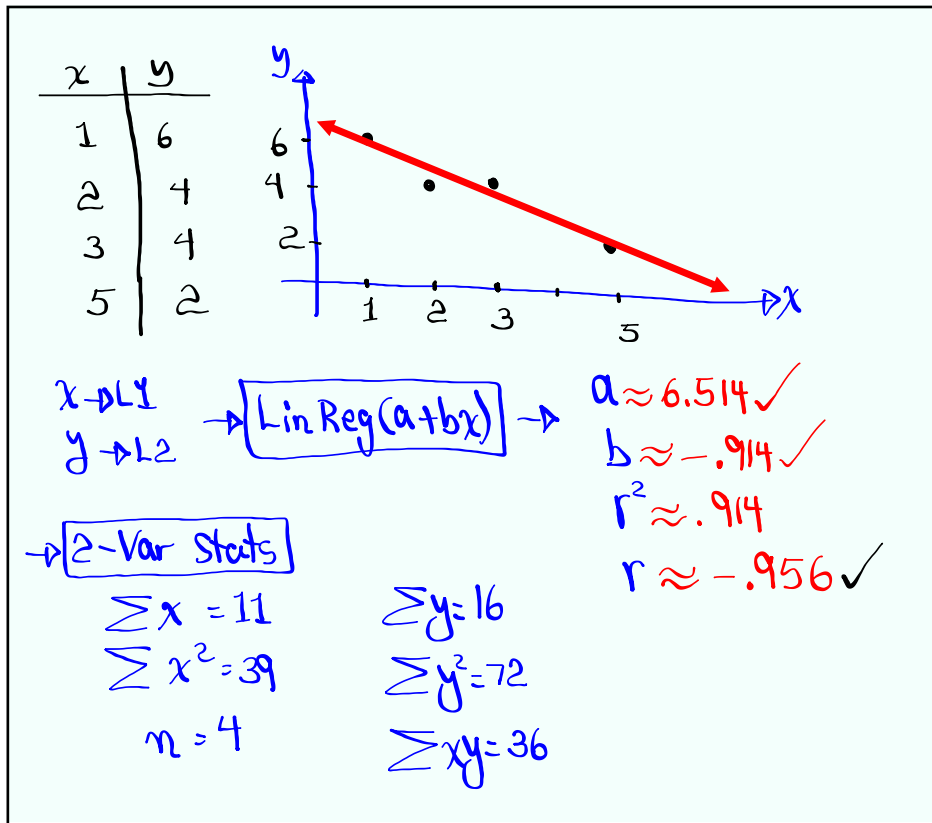$\bar{y} = \frac{\sum y}{n} = \frac{440}{5} = \boxed{88}$

$\boxed{\text{VARS}}$

$\boxed{5: \text{Statistics}}$

$\boxed{5: \bar{y}}$ $\boxed{\text{Enter}}$

$\boxed{88}$

Jan 14-5:38 PM

x	y
1	6
2	4
3	4
5	2

$x \to L1$
$y \to L2$ $\to \boxed{\text{Lin Reg}(a+bx)} \to$

$a \approx 6.514$ ✓
$b \approx -.914$ ✓
$r^2 \approx .914$
$r \approx -.956$ ✓

$\to \boxed{\text{2-Var Stats}}$

$\sum x = 11$ $\sum y = 16$
$\sum x^2 = 39$ $\sum y^2 = 72$
$n = 4$ $\sum xy = 36$

Jan 14-6:02 PM

$\sum x = 11$ $\sum y = 16$ Regression line
$\sum x^2 = 39$ $\sum y^2 = 72$ $y = a + bx$
$n = 4$ $\sum xy = 36$

$$a = \frac{\sum y \sum x^2 - \sum x \sum xy}{n \sum x^2 - (\sum x)^2} = \frac{16 \cdot 39 - 11 \cdot 36}{4 \cdot 39 - 11^2} = \frac{228}{35}$$

$$\approx \boxed{6.514}$$

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = \frac{4 \cdot 36 - 11 \cdot 16}{4 \cdot 39 - 11^2} = \frac{-32}{35} \approx \boxed{-.914}$$

Jan 14-6:08 PM

formula for r:

$\sum x = 11$ $\sum y = 16$

$\sum x^2 = 39$ $\sum y^2 = 72$

$n = 4$ $\sum xy = 36$

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} \; \sqrt{n\sum y^2 - (\sum y)^2}} = \frac{4 \cdot 36 - 11 \cdot 16}{\sqrt{4 \cdot 39 - 11^2} \cdot \sqrt{4 \cdot 72 - 16^2}}$$

$$= \frac{-32}{\sqrt{35} \; \sqrt{32}} = \frac{-32}{\sqrt{1120}} \approx \boxed{-.956}$$

$$r^2 \approx (-.956)^2 \approx \boxed{.914} \qquad \text{SG 9} \; \checkmark$$

Jan 14-6:14 PM

Given $\quad y \approx 32 - 2.5x$, $\boxed{\bar{y} = 18}$

Predict $\quad y$ if $\boxed{x = 4}$

1) Assume Linear Correlation is Significant

use regression line

$$y = 32 - 2.5(4) = 32 - 10 = \boxed{22}$$

2) Assume linear Correlation is not Significant.

use $\bar{y}$ $\boxed{\bar{y} = 18}$

Jan 14-6:21 PM

Use the Chart below

class MP	class F
18	3
25	7
32	12
39	18
46	10

1) # of classes = 5

2) class width = 7

3) $n = \sum f = 50$

4) $\bar{x} = 35.5$

5) $S = 8.031$

6) S^2 in reduced frac.

$$= \frac{129}{2}$$

class MP → L1
class F → L2
1 - Var Stats with L1 & L2

Jan 14-6:25 PM

$\bar{x} \approx 36 \qquad S \approx 8$

Usual Range $\qquad \bar{x} \pm 2S = 36 \pm 2(8)$

$$= 36 \pm 16 \rightarrow \boxed{20 \text{ to } 52}$$

find Z-Score for data element 45.

$$Z = \frac{x - \bar{x}}{S} = \frac{45 - 36}{8} = \frac{9}{8} = \boxed{1.125} \quad \text{usual element}$$

find the data element with Z = 1.75.

$$Z = \frac{x - \bar{x}}{S} \qquad 1.75 = \frac{x - 36}{8}$$

$$x - 36 = 8(1.75)$$

$$x = 36 + 8(1.75) = \boxed{50}$$

Jan 14-6:31 PM

I randomly selected 35 nurses. Here are their ages.

42	35	50	28	30
40	32	48	52	55
25	24	60	58	43
49	60	29	35	35
43	57	51	48	44
44	42	29	30	39
48	33	24	53	38

Store in a list

Sort that list

View that list

Make Stem Plot

```
2 | 445899
3 | 002355589
4 | 0223344889
5 | 0123578
6 | 00
```

find $\bar{x}$, S.

$\bar{x} = 41.514$

$S = 10.763$

$n = 35$

Min. $= 24$

$Q_1 = 32$

Med $= 42$

$Q_3 = 50$

Max $= 60$

$S^2 = \dfrac{68928}{595}$

5-Number Summary

24   32 42 50   60

$IQR = Q_3 - Q_1 = 18$

upper fence $= Q_3 + 1.5(IQR) = 77$

Lower fence $= Q_1 - 1.5(IQR) = 5$

No outliers

Jan 14-6:37 PM

```
2 | 445899
3 | 002355589
4 | 0223344889
5 | 0123578
6 | 00
```

find P_{30}

$L = \dfrac{30}{100} \cdot 35 = 10.5 \Rightarrow L = 11$

$P_{30} = 11$th element

[2nd] [1] (11) [Enter]   = 35
$\underbrace{\qquad}_{L1}$

find P_{50} = Median

$L = \dfrac{50}{100} \cdot 35 = 17.5 \quad L = 18 \qquad \boxed{42}$

[2nd] [Enter] use arrow → change 11 to 18
                                    [Enter]

find P_{80} $L = \dfrac{80}{100} \cdot 35 = 28$

$P_{80} = \dfrac{28\text{th element} + \text{Next one}}{2} = \dfrac{51 + 52}{2}$

$= \boxed{51.5}$

Jan 14-6:53 PM

```
2 | 4 4 5 8 9 9
3 | 0 0 2 3 5 5 5 8 9
4 | 0 2 2 3 3 4 4 8 8 8 9
5 | 0 1 2 3 5 7 8
6 | 0 0
```

find the percentile ranking of 50.

find K such that $P_K = 50$

Below

$$K = \frac{B}{n} \cdot 100$$

$$= \frac{26}{35} \cdot 100 = 74.2857$$

$$\approx 74$$

$$\boxed{P_{74} = 50}$$

Jan 14-6:59 PM

Intro. to Probabilities:

$E \longrightarrow$ Event

$P(E) \longrightarrow$ Prob. that E happens

$$P(E) = \frac{\# \text{ of outcomes of } E}{\# \text{ of total outcomes}}$$

In a class of 18 students, there were 10 females & 8 males.

If we select one person,

$$P(\text{Select a female}) = \frac{10}{18} = \boxed{\frac{5}{9}}$$

Jan 14-7:04 PM

A standard deck of playing Cards.

  52 Cards , 26 Red, 12 face Cards

One Card is randomly taken,

1) $P(\text{Red Color}) = \dfrac{26}{52} = \boxed{\dfrac{1}{2}}$

2) $P(\text{face Card}) = \dfrac{12}{52} = \dfrac{3}{13}$

   12 ÷ 52 [Math] [1:►frac] [Enter]

Jan 14-7:08 PM