Statistics

Lecture 5

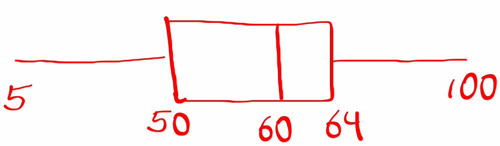
Feb 19-8:47 AM

Class QZ 6

Consider the 5-Number Summary give below

Min	Q_1	Med.	Q_3	Max
5	50	60	64	100

1) Draw Box Plot



2) $IQR = Q_3 - Q_1 = 64 - 50 = \boxed{14}$

3) Upper fence

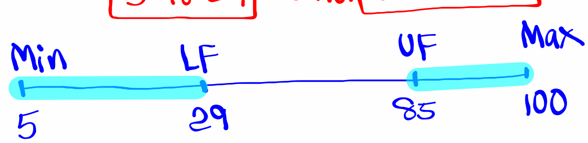
$= Q_3 + 1.5(IQR) = 64 + 1.5(14) = \boxed{85}$

4) Lower fence

$= Q_1 - 1.5(IQR) = 50 - 1.5(14) = \boxed{29}$

5) Discuss range for possible outliers.

$\boxed{5 \text{ to } 29}$ and $\boxed{85 \text{ to } 100}$



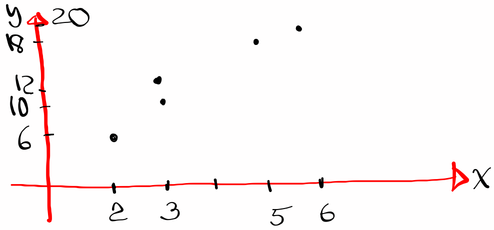
Sep 19-9:03 PM

Given (2,6), (3,10), (3,12), (5,18), (6,20)

SG 9

x	y
2	6
3	10
3	12
5	18
6	20

1) Draw Scatter Plot



Clear all lists

Reset all lists

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

STAT → CALC

2:2-Var stats

Menu

x list: L1

Y list: L2

FreqList: Blank

Calculate

No Menu

L1, L2

enter

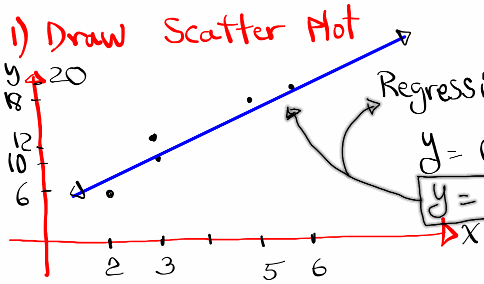
$\sum x = 19$ $\sum y = 66$

$\sum x^2 = 83$ $\sum y^2 = 1004$

$n = 5$ $\sum xy = 288$

Sep 26-6:52 PM

1) Draw Scatter Plot



Regression line

$y = a + bx$

$\boxed{y = .\bar{1} + 3.\bar{4}x}$

STAT → CALC

8: LinReg(a+bx)

Menu

x list: L1

Ylist: L2

Blank

Calculate

No Menu

L1, L2

Enter

$a = .\bar{1}$

$b = 3.\bar{4}$

$r^2 = .965$

$r = .982$

2nd 0 ↓↓↓ - - - ↓ ▸ DiagnosticOn

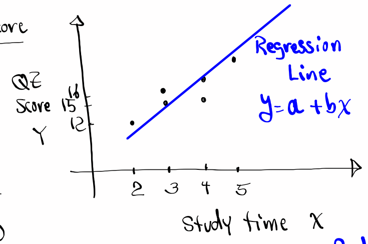
Enter Enter

If r & r^2 missing.

Sep 26-7:00 PM

Consider the chart below

x Study time	y Quiz Score
2	12
3	15
3	16
4	18
4	15
5	20



Clear all lists

Study time $\to x \to L1$

QZ Score $\to y \to L2$

Use 2-VAR Stats to find

$\sum x = 21$ $\quad \sum y = 96$

$\sum x^2 = 79$ $\quad \sum y^2 = 1574$

$n = 6$ $\quad \sum xy = 349$

Use LinReg(a+bx) to find

$a = 7.\overline{72} \approx 8$

$b = 2.\overline{36} \approx 2$ $\quad \Rightarrow \boxed{y \approx 8 + 2x}$

$r^2 = .809$ $\quad r^2(\%) = 81\%$

$r = .899$

Coef. of determination

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

81% of QZ Score are explained by study time

19% are unexplained

Sep 26-7:06 PM

$r \to$ Linear Correlation Coef.

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

When r is close to ± 1,
Linear Correlation is Significant

When r is close to 0,
Linear Correlation is not Significant.

Sep 26-7:17 PM

Consider the chart below

x	y
1	5
2	12
3	10
3	15
5	18

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

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

$\checkmark a = 3.727 \approx 4$

$\checkmark b = 2.955 \approx 3$

$\Rightarrow \boxed{y = 4 + 3x}$

Coef. of determination

$r^2 \approx 78\%$

$r^2 = .784 \approx 78\%$

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

$\checkmark r = .885$

Linear Correlation Coef.

It is close to 1 therefore, Linear Correlation is Significant.

Sep 26-7:21 PM

How to find a & b using formula:

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

$$= \frac{60 \cdot 48 - 14 \cdot 194}{5 \cdot 48 - 14^2}$$

$$= \frac{164}{44} = 3.\overline{72} \approx \boxed{4}$$

$n = 5$

$\Sigma x = 14$

$\Sigma x^2 = 48$

$\Sigma y = 60$

$\Sigma y^2 = 818$

$\Sigma xy = 194$

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

$$\frac{5 \cdot 194 - 14 \cdot 60}{5 \cdot 48 - 14^2} = \frac{130}{44}$$

$$= 2.955 \approx \boxed{3}$$

Sep 26-7:28 PM

How to find r using formula:

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

$n = 5$
$\sum x = 14$
$\sum x^2 = 48$
$\sum y = 60$
$\sum y^2 = 818$
$\sum xy = 194$

$$= \frac{5 \cdot 194 - 14 \cdot 60}{\sqrt{5 \cdot 48 - 14^2}\ \sqrt{5 \cdot 818 - 60^2}}$$

$$= \frac{130}{\sqrt{44}\ \sqrt{490}} = \frac{130}{\sqrt{21560}} \approx \boxed{.885}$$

For $r^2 = (.885)^2 = .783 \approx \boxed{78\%}$

Sep 26-7:34 PM

I randomly selected 8 days, below are my walking time, and next day blood sugar level.

Walk time	B.S. level
10	140
15	125
20	110
10	135
20	115
30	100
35	105
40	100

Walk time → x → L1
B.S. level → y → L2
Lin Reg (a+bx) with L1 & L2
$a = 144.375 \approx 144$
$b = -1.25 \approx -1$
$r^2 = .833 \approx 83\%$
$r = -.913$

Regression line
$y = a + bx$
$\boxed{y \approx 144 - x}$

Coef. of determination
$r^2 = 83\%$
83% of my B.S. level are explained by walking time.

Linear Correlation Coef.
$r = -.913$
It is close to -1, so it is considered significant.

Sep 26-7:42 PM

Two branches in Statistics:

1) Descriptive

2) Inferential

How To make predictions:

If r is significant ⇒ use the regression line
Plug in x-value To
Predict Y-value.

If r is not significant ⇒ use $\bar{y}$

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

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

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

For last example

$\bar{y} = 116.25$

≈ 116

Sep 26-7:51 PM

Study time	exam Score
8	75
7	70
9	84
10	95
5	65
12	100

Study time → x → L1

exam Score → y → L2

LinReg(a + bx) with
L1 & L2

$a = 34.102 \approx 34$

$b = 5.576 \approx 6$

$r^2 = .938 \approx 94\%$

$r = .969$

Regression line

$\boxed{y \approx 34 + 6x}$

94% of exam Scores are explained by Study time

r is close To 1,
Linear Correlation is significant.

Sep 26-8:12 PM

Predict exam Score for Someone that studied $\boxed{10 \text{ hrs}}$ for the exam

1) Assume r is Significant

use Regression line $y = 34 + 6x$

$y = 34 + 6(10) = 34 + 60 = \boxed{94}$

2) Assume r is not Significant

SG 9

use $\bar{y}$ for prediction value

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

$\bar{y} = 81.5$ $\boxed{\bar{y} \approx 82}$

Sep 26-8:18 PM

Introduction To Probabilities:

SG 10-13

$E \to$ Desired event (outcome)

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

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

There are 25 students, 10 males and 15 females. Let's randomly select one person

$$P(\text{Select a female}) = \frac{\text{Total \# of all females}}{\text{Total \# of people}}$$

$$= \frac{15}{25} = \boxed{\frac{3}{5}}$$

Sep 26-8:22 PM

A standard deck of playing cards has 52 cards, 4 Aces.

Let's randomly draw one card,

$$P(\text{draw an ace}) = \frac{\text{Total \# of aces}}{\text{Total \# of cards}}$$

$$= \frac{4}{52} = \boxed{\frac{1}{13}}$$

4 [÷] 52 [MATH] [1:▶Frac] [Enter]

If we randomly select one person, find the Prob. that he/she had a birthday this month

$$\frac{1 \text{ Month}}{12 \text{ months}} = \boxed{\frac{1}{12}}$$

Sep 26-8:27 PM

Do You Support tougher Gun Laws?

	Yes	No	Total
Males	15	45	60
Females	25	15	40
Total	40	60	100

If we randomly select one of these people,

$P(\text{Male}) = \frac{60}{100} = \boxed{.6}$ $P(\text{Yes}) = \frac{40}{100} = \boxed{.4}$

$P(\text{Male and Yes}) = \frac{15}{100} = \boxed{.15}$ $P(\text{Male or Yes}) = \frac{85}{100} = \boxed{.85}$

Acceptable Form of answers:

1) Reduced fraction
2) Round to 3-decimal places
3) Scientific Notation

Sep 26-8:31 PM

Some rules & Terminologies

1) $0 \leq P(E) \leq 1$

2) Sum of all prob. is always 1.

3) $P(E) = 1 \iff$ Sure event

4) $P(E) = 0 \iff$ Impossible event

5) $0 < P(E) \leq .05 \iff$ Rare Event

Sep 26-8:39 PM

$E \to$ Desired event

$\bar{E} \to$ E-bar, E-Complement, Not E

$P(E) + P(\bar{E}) = 1$

Complement Rule
$P(\bar{E}) = 1 - P(E)$

A deck of playing cards has 52 cards, and 12 are face cards. Let's draw one card

$P(\text{Face}) = \frac{12}{52} = \frac{3}{13}$

$P(\overline{\text{Face}}) = \frac{40}{52} = \frac{10}{13}$

$\to \frac{3}{13} + \frac{10}{13} = \frac{13}{13} = 1.$

$P(\overline{\text{Face}}) = 1 - P(\text{face}) = 1 - \frac{3}{13} = \frac{10}{13}$

1 [−] 3 [÷] 13 [Math] [1:▶Frac] [Enter]

Sep 26-8:42 PM

Addition Rule

Keyword: OR   Single-Action Event

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

$P(A) = .4$, $P(B) = .7$, $P(A \text{ and } B) = .25$

$P(\bar{A}) = 1 - P(A) = 1 - .4 = \boxed{.6}$

$P(\bar{B}) = 1 - P(B) = 1 - .7 = \boxed{.3}$

$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

$= .4 + .7 - .25$

$= \boxed{.85}$

$P(\overline{A \text{ or } B}) = 1 - P(A \text{ or } B)$

$= 1 - .85 = \boxed{.15}$

Sep 26-8:48 PM

$P(HB) = .45$

$P(FF) = .35$

$P(HB \text{ and } FF) = .25$

$P(\overline{HB}) = 1 - .45 = \boxed{.55}$ $P(\overline{FF}) = 1 - .35 = \boxed{.65}$

$P(HB \text{ or } FF) = P(HB) + P(FF) - P(HB \text{ and } FF)$

$= .45 + .35 - .25 = \boxed{.55}$

Sep 26-8:55 PM

Mutually Exclusive Events (Disjointed Events)

If A and B are M.E.E., then
$P(A \text{ and } B) = 0.$

Suppose $P(A)=.2$, $P(B)=.7$, A & B are M.E.E.

1) $P(\overline{A}) = 1-P(A) = \boxed{.8}$ 2) $P(\overline{B}) = 1-P(B) = \boxed{.3}$

3) $P(A \text{ and } B) = \boxed{0}$ Since A & B are M.E.E.

4) $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$
$= .2 + .7 - 0$
$= \boxed{.9}$

Make sure to watch all videos on the right-hand side of SG 10-13.

Sep 26-9:00 PM

Class QZ 7

Use the chart below

x	y
2	8
4	12
5	15
5	18
8	20

Find

$a = 4.775 \approx \boxed{5}$
$b = 2.053 \approx \boxed{2}$
} Round to whole #

$r^2 = .869 \approx \boxed{87\%}$ } Round to whole %

$r = \boxed{.932}$ } Round to 3-decimal places

Sep 26-9:09 PM