Notes 9.1 & 9.2 Correlation & Linear Regression

Correlation

- A correlation is a relation ship between "x" and "y"
- The data can be represented by (x, y) where x is the explanatoly or ________________ variable, and y is the response, or ________________ variable.
- One way to determine whether a linear correlation exists between two variables is to use a Regression Line.

"Line of best fit" →

$y = ax + b$

r ⇒ from the calculator

⇒ correlation coefficient

Example 1:

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following Monday are given below:

Hours spent online, x	0	1	2	3	3	5	5	5	6	7	7	10
Test score, y	96	85	82	74	95	68	76	84	58	65	75	50

a. Sketch a graph of the distribution and then describe the distribution.

100 90 80 70 60 50

x 0 1 2 3 4 5 6 7 8 9 10

b. Find the regression line.

c. Find the correlation coefficient.

d. Use the regression line to predict the test scores given the time online:

x = 4 hours x = 9 hours x = 15 hours

Notes 9.1 & 9.2 Correlation & Linear Regression

Correlation

- A correlation is a relationship between "x" and "y".
- The data can be represented by (x, y) where x is the explanatory variable, and y is the response variable.
- One way to determine whether a linear correlation exists between two variables is to use a Regression $\boxed{\text{Line}}$

Correlation coefficient: "r" tells the strength of the regression line.

"Line of best fit" $y = ax + b$ ← this is a line

Example 1:

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following Monday are given below:

Hours spent online, *x*	0	1	2	3	3	5	5	5	6	7	7	10
Test score, *y*	96	85	82	74	95	68	76	84	58	65	75	50

a. Sketch a graph of the distribution and then describe the distribution.

b. Find the regression line.

c. Find the correlation coefficient.

d. Use the regression line to predict the test scores given the time online:

x = 4 hours x = 9 hours x = 15 hours

- One way to determine whether a linear correlation exists between two variables is to use a Regression Line

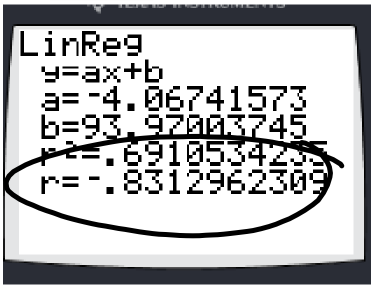
Correlation coefficient : "r" tells the strength of the regression line.

"Line of best fit"

$y = ax + b$ ← this is a line

↖ slope

Example 1:



$\boxed{r = .83}$

$$\frac{\text{Rise}}{\text{Run}} \Rightarrow \frac{y_2 - y_1}{x_2 - x_1} = \frac{\text{Response}}{\text{Explanatory}}$$

- One way to determine whether a linear correlation exists between two variables is to use a Regression Line.

"Line of best fit" →

$y = ax + b$

r ⇒ from the calculator

⇒ correlation coefficient

Example 1:

a ⇒ slope $\frac{y_2 - y_1}{x_2 - x_1}$ ← this means?

← Science

Rate of Change $\frac{\text{Rise}}{\text{Run}}$

Notes 9.1 & 9.2 Correlation & Linear Regression

Example 2:

The budgets & worldwide grosses of 15 of the most expensive 20th Century Fox Movies are shown.

Budget, x (millions)	200	150	125	125	115	115	115	110	110	110	105	102	100	100	100
Gross, y (millions)	1835.4	459.4	406.4	542.7	924.3	656.7	848.5	571.1	211.4	150.5	348.8	358.8	365.3	359.1	249.0

a. Sketch a graph of the distribution and then describe the distribution.

b. Find the regression line.

c. Find the correlation coefficient.

d. Use the regression line to predict the gross amount of money for the given budget:

$120,000,000 *$93,000,000*

Notes 9.1 & 9.2 Correlation & Linear Regression

Correlation

- A correlation is a relationship between "x" and "y"
- ✱ The data can be represented by (x, y) where x is the explanatory variable, and y is the response variable.
- One way to determine whether a linear correlation exists between two variables is to use a Regression Line

• correlation coefficient "r" ✱ calculator

"Line of Best Fit"

$\boxed{y = ax + b}$

Line $y = mx + b$

positive 0 to 1

negative 0 to −1

Example 1:

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following Monday are given below:

Hours spent online, *x*	0	1	2	3	3	5	5	5	6	7	7	10
Test score, *y*	96	85	82	74	95	68	76	84	58	65	75	50

a. Sketch a graph of the distribution and then describe the distribution.

b. Find the regression line.

c. Find the correlation coefficient.

d. Use the regression line to predict the test scores given the time online:

x = 4 hours x = 9 hours x = 15 hours

and y is the ______________ variable.

- One way to determine whether a linear correlation exists between two variables is to use a Regression Line

• correlation coefficient "r"

"Line of Best Fit"

positive 0 to 1

negative 0 to −1

calculator

$\boxed{y = ax + b}$

Line

$y = mx + b$

slope

Example 1:

$$\frac{\text{Rise}}{\text{Run}} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\text{Response}}{\text{Explanatory}}$$

Notes 9.1 & 9.2 Correlation & Linear Regression

Correlation

- A correlation is a relationship between "x" and "y"
- The data can be represented by (x, y) where *x* is the explanatory variable, and *y* is the response variable.
- One way to determine whether a linear correlation exists between two variables is to use a Regression Line "Line of Best Fit"

- correlation coefficient: "r" → calculator

$\boxed{y = ax + b}$ $y = mx+b$

Line

- positive 0 to 1
- negative 0 to -1

Example 1:

The number of hours 12 students spent online during the weekend and the scores of each student who took a test the following Monday are given below:

Hours spent online, x	0	1	2	3	3	5	5	5	6	7	7	10
Test score, y	96	85	82	74	95	68	76	84	58	65	75	50

a. Sketch a graph of the distribution and then describe the distribution.

b. Find the regression line.

c. Find the correlation coefficient.

d. Use the regression line to predict the test scores given the time online:

x = 4 hours x = 9 hours x = 15 hours

- One way to determine whether a linear correlation exists between two variables is to use a Regression Line

- correlation coefficient : "r" calculator
  - positive 0 to 1
  - negative 0 to -1

"Line of Best Fit"

$\boxed{y = ax + b}$ $y = mx + b$

Line

Slope

$\frac{\text{Rise}}{\text{Run}} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\text{Response}}{\text{Explanatory}}$