Two - Variable Statistics

Statistical studies often involve more than one variable. We are interested in knowing if there is a relationship between two characteristics for the same subject.

Example: Temperature and ice cream sales

When the data is quantitative, the variables can be written as an ordered pair (x, y).

Correlation is the study and description of the relationship (if any) that exists between the variables.

A) *Qualitative Interpretation of Correlation*

Data will be organised and displayed in a scatterplot (Cartesian plane) or a contingency table.

By *inspection*, we will describe the type, the direction, and the intensity (or strength) of the relation between the variables.

Type: Refers to the function that most resembles the relation between the variables (Linear for us).

Direction: Refers to the variation.
If both variables move in the *same direction* (both increase together or decrease together), the direction is positive.
If both variables move in *opposite directions*, (one increases and the other decreases) the direction is negative.

Intensity or Strength: May be categorised as...

Zero, weak, moderate, strong or perfect.

Example: In a gym class, 20 students were required to do push-ups (x) and sit-ups (y). Each student's achievements was recorded as an ordered pair.

(27, 30), (26, 28), (38, 45), (52, 55), (35, 36),
(40, 54), (40, 50), (52, 46), (42, 55), (61, 62),
(35, 38), (45, 53), (38, 42), (63, 55), (55, 54),
(46, 46), (34, 36), (45, 45), (30, 34), (68, 62)

Contingency Table

- Listed down the left side is one variable (x - value); listed across the top is the other variable (y - value).

- The values in the cells (where the row and column intersect) are the frequencies for each unique combination of the two variables.

- Each row and column may be added and the total displayed; the bottom right cell shows the total frequency (number of subjects) of the distribution.

(27, 30), (26, 28), (38, 45), (52, 55), (35, 36),
(40, 54), (40, 50), (52, 46), (42, 55), (61, 62),
(35, 38), (45, 53), (38, 42), (63, 55), (55, 54),
(46, 46), (34, 36), (45, 45), (30, 34), (68, 62)

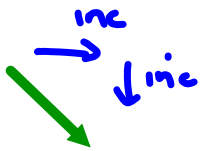
inc

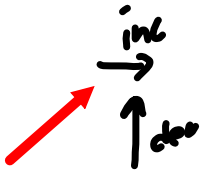
Sit-ups / Push-ups	[20,30[	[30,40[	[40,50[	[50,60[	[60,70[	
[20,30[	1	1	0	0	0	
[30,40[	0	4	2	0	0	
[40,50[	0	0	2	4	0	
[50,60[	0	0	1	2	0	
[60,70[	0	0	0	1	2	

inc ↑ dec ↓

linear strong positive

When the majority of the values fall into a diagonal of a contingency table, and the corners contain mostly zeros, the correlation is said to be linear and strong.

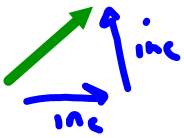
Direction is positive if the diagonal is 

Direction is negative if the diagonal is 

Scatterplot - A Cartesian graph
- Each ordered pair is represented as a point

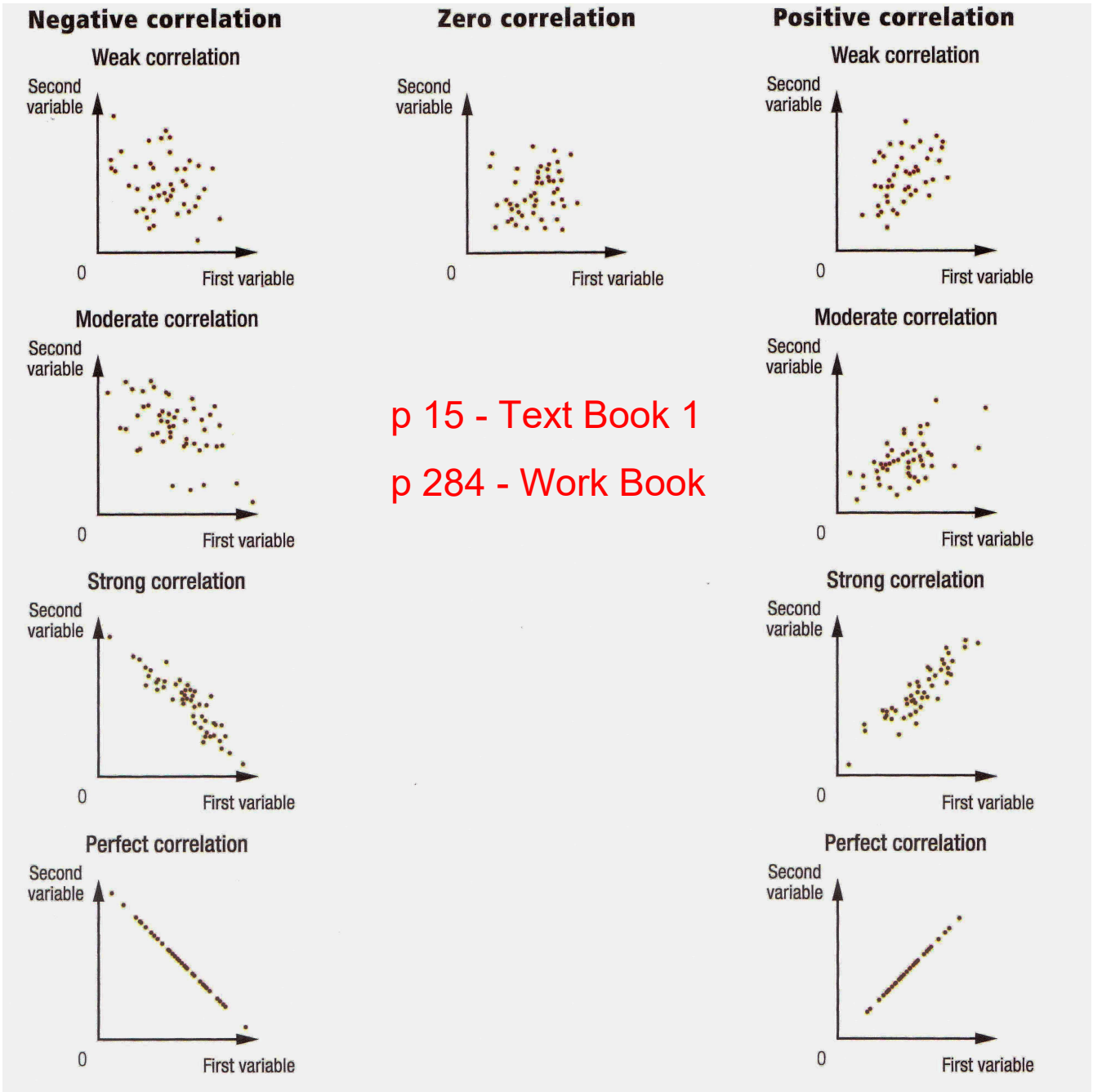
(27, 30) (26, 28), (38, 45), (52, 55), (35, 36), (40, 54), (40, 50), (52, 46), (42, 55), (61, 62), (35, 38), (45, 53), (38, 42), (63, 55), (55, 54), (46, 46), (34, 36), (45, 45), (30, 34), (68, 62)

- The type of correlation is linear; when we look at the scatterplot we see if the arrangement of the points resembles a straight line.
- The intensity describes how closely the points come together to form a straight line.
- To determine the direction:

Positive 

Negative 

(think of slope)



p 15 - Text Book 1

p 284 - Work Book

The following are some characteristics of an exponential correlation: