3.2 Measures of Variation

GOALS:

I. Understand and measure variation in data sets
   1. Range
   2. Variance
   3. Standard Deviation

II. Understand differences in computation for sample and population variations

Study Ch. 3.2 # 59 - 69 (work out),
73, ~77, 81 (calculator)

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

Compute:

a) mean of each set
b) How are the sets different?
c) Which has least variation?
d) range for each
e) Standard dev, using Defining Formula
f) Which better distinguishes the spread of the data, Range or Stand dev?

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

Compute: a) mean of each set $\bar{x} = \frac{\Sigma x}{n}$

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

Compute: a) mean of each set $\bar{x} = \frac{\Sigma x}{n}$

$\frac{50}{10} = 5.0$ $\frac{50}{10} = 5.0$ 5.0 $\bar{x} = 5.0$

All 4 sets have the same mean = 5.0

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

Compute: a) mean of each set $\bar{x} = \frac{\Sigma x}{n}$

$\frac{50}{10} = 5.0$ $\frac{50}{10} = 5.0$ 5.0 $\bar{x} = 5.0$

All 4 sets have the same mean = 5.0

b) How are the sets different?

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

Compute: a) mean of each set $\bar{x} = \frac{\Sigma x}{n}$

$\frac{50}{10} = 5.0$ $\frac{50}{10} = 5.0$ 5.0 $\bar{x} = 5.0$

All 4 sets have the same mean = 5.0

b) How are the sets different?

I has repeats of 5 numbers: 1,2,5,8,9

II has repeats of 2 numbers: 1,9

III has repeats of 1 number: 5

IV has repeats of 3 numbers: 2,4,10

Variation

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

c) Which has **least** variation?

d) range for each

RANGE = Max - Min
> Describes bounds
> Easiest measure of variation

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

c) Which has least variation? III

d) range for each

RANGE = Max - Min
> Describes bounds
> Easiest measure of variation

range 9-1=8 9-1=8 5-5=0 10-2=8

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

c) Which has least variation? d) range for each

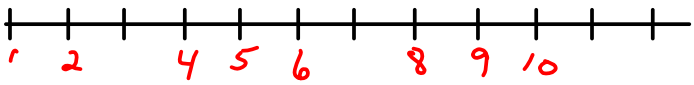
range 9-1=8 9-1=8 5-5=0 10-2=8

RANGE = Max - Min

> Describes bounds

> Easiest measure of variation

> Need more info to distinguish variation within the bounds



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

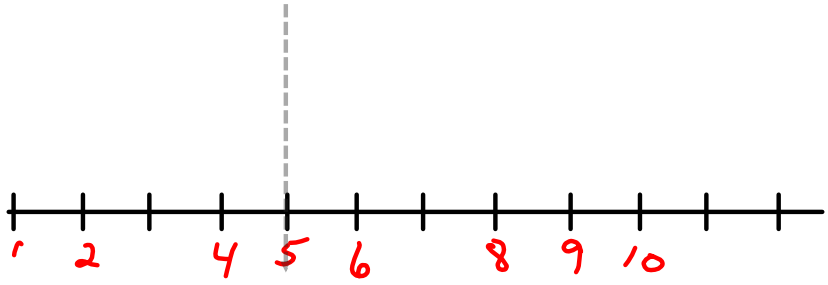
Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

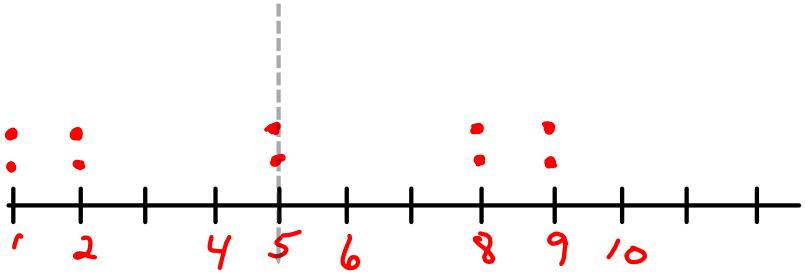
Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

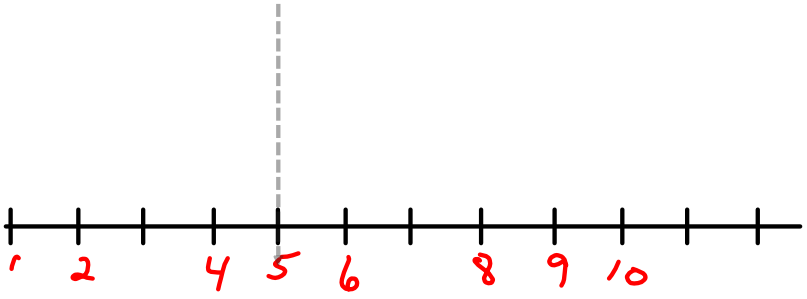
Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

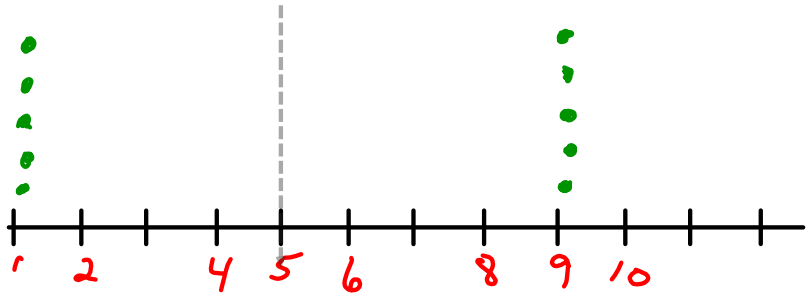
3.2 Measures of Variation

Given the following:

DATA_I	DATA_II	DATA_III	DATA_IV
1	1	5	2
5	9	5	4
1	1	5	4
8	9	5	4
2	1	5	4
$\bar{x}=5$	$\bar{x}=5$	$\bar{x}=5$	$\bar{x}=5$
			4
			10
			4
			10

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

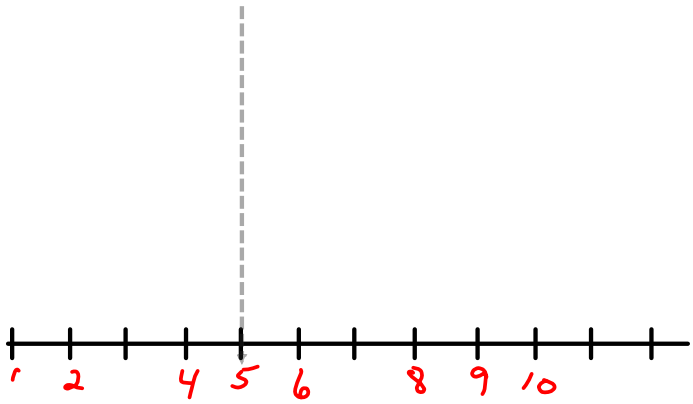
3.2 Measures of Variation

Given the following:

DATA_I	DATA_II	DATA_III	DATA_IV
1	1	5	2
5	9	5	4
1	1	5	4
8	9	5	4
2	1	5	4
$\bar{x}=5$	$\bar{x}=5$	$\bar{x}=5$	$\bar{x}=5$
			4
			10
			4
			10

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

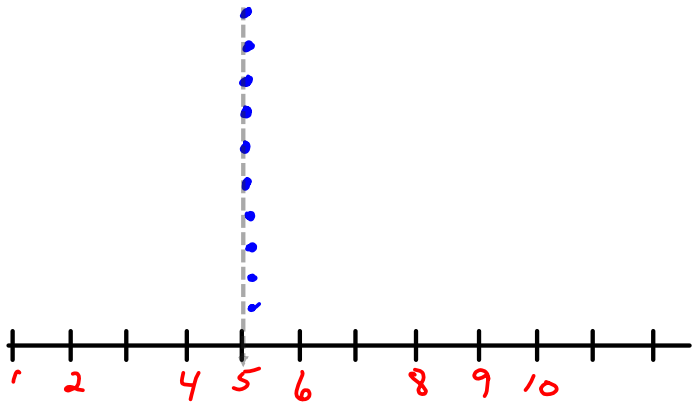
Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$ s

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

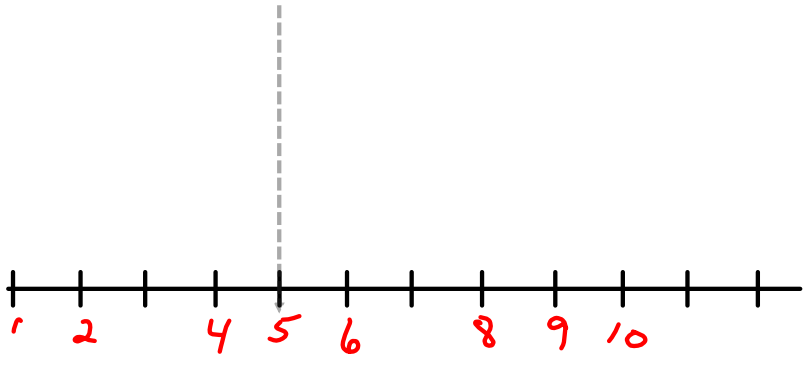
Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$ s

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

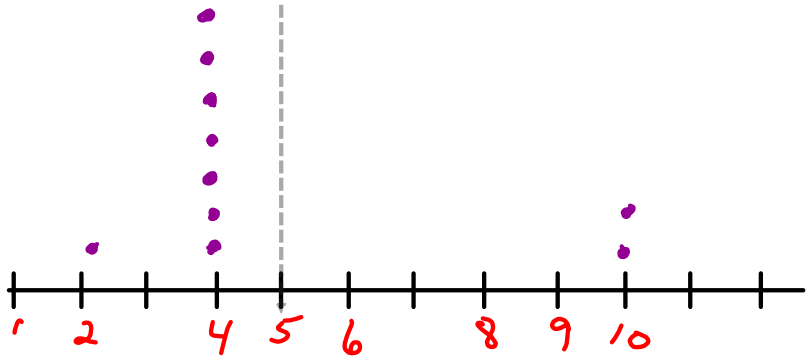
3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$ $\bar{x} = 5.0$ all

How far away from the mean is each data item?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Given the following:

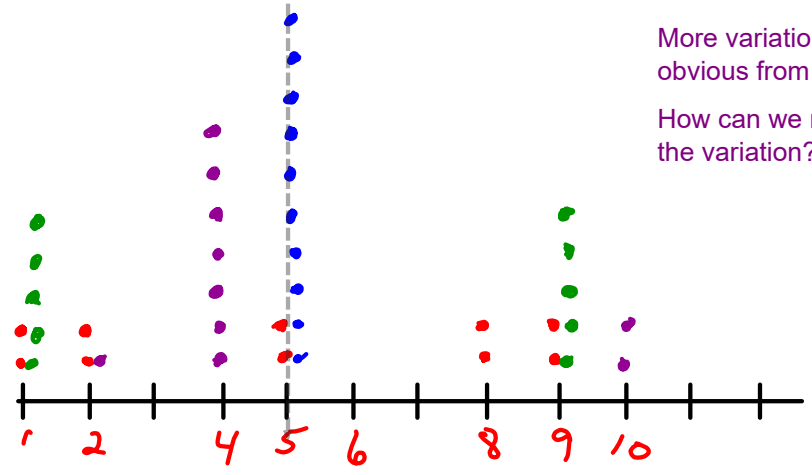
DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5$ $\bar{x} = 5.0$ all

range **9-1=8** **9-1=8** **5-5=0** **10-2=8**

More variation than is obvious from Range.

How can we measure the variation?



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

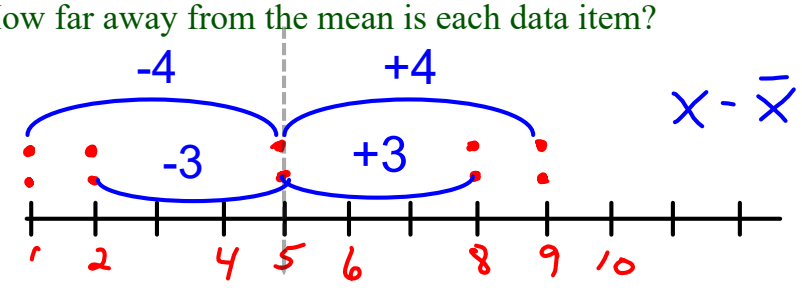
3.2 Measures of Variation

Given the following:

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

$\bar{x} = 5.0$ all

How far away from the mean is each data item?



If we add the differences, we get $\Sigma x = 0$.

+4 +(-4) +4 +(-4) +3 +(-3) +3 +(-3) = 0

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Want a measure that describes the distance away from the mean of each data item.

Data I	$\bar{x} = 5.0$	
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1		
5		
1		
8		
2		
8		
2		
9		
5		
9		

Looking for mathematical measure: Average distance away from the mean.

>Data set with small distances from the mean has little variation.

>Data set with larger distances from the mean has large variation.

Start with $x - \bar{x}$

Since $\bar{x}$ is designed to be the fulcrum of distribution, if we add distance from $\bar{x}$ we get 0.

Therefore, square the differences

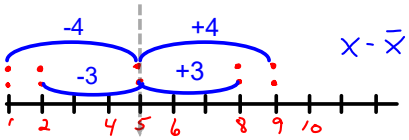
Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Want a measure that describes the distance away from the mean of each data item.

Data I	$\bar{x} = 5.0$	
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	-4	16
5	0	0
1	-4	16
8	3	9
2	-3	9
8	3	9
2	-3	9
9	4	16
5	0	0
9	4	16
	0	100



- 14 + 14 = 0

Looking for mathematical measure: Average distance away from the mean.
>Data set with small distances from the mean has little variation.
>Data set with larger distances from the mean has large variation.

Start with $x - \bar{x}$

Since $\bar{x}$ is designed to be the fulcrum of distribution, if we add distance from $\bar{x}$ we get 0.

Therefore, square the differences

Average squared difference

$\frac{100}{10} = 10$

Find square root to return to order of magnitude of the original data.

$\sqrt{10} \approx 3.16$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Want a measure that describes the distance away from the mean of each data item.

Data I	$\bar{x} = 5.0$	
x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1		
5		
1		
8		
2		
8		
2		
9		
5		
9		

Want: Average distance away from the mean of all data.
So find average squared distance, then find the square root.

$$\frac{\Sigma(x - \bar{x})^2}{n} \qquad \sigma = \sqrt{\frac{\Sigma(x - \mu)^2}{n}}$$

$$s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}$$

VARIABLES	Statistic Sample	Parameter Population
Mean	$\bar{x}$	μ
Standard Dev.	s	σ
	statistics	parameters

Why n-1 ?

results in better estimate of the population σ

(see student heights xls)

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework μ $\bar{x}$ Σ

3.2 Measures of Variation

Want a measure that describes the distance away from the mean of each data item.
But, also want an **average distance away from the mean** - more meaningful

Defining Formula for the **Sample** Standard Deviation

$$s = \sqrt{\frac{\Sigma (x_i - \overline{x})^2}{n-1}}$$

3.2 Measures of Variation

#50. Given the following:

$\bar{x} = 5$

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

e) Standard dev, using Defining Formula
f) Which better distinguishes the spread of the data, Range or Stand dev?

Data II $\bar{x} = 5.0$

x_i	$x - \bar{x}$	$(x-\bar{x})^2$
1		
9		
1		
9		
1		
9		
1		
9		
1		
9		

$$s = \sqrt{\frac{\Sigma (x_i - \overline{x})^2}{n-1}}$$

3.2 Measures of Variation

#50. Given the following: $\bar{x} = 5$

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10

e) Standard dev, using Defining Formula

f) Which better distinguishes the spread of the data, Range or Stand dev?

Data II

$\bar{x} = 5.0$

x_i	$x - \bar{x}$	$(x-\bar{x})^2$
1	-4	16
9	4	16
1	-4	16
9	4	
1	-4	
9	4	
1	-4	
9	4	
1	-4	
9	4	16
		Σ 160

$$s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{160}{9}} = 4.22$$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

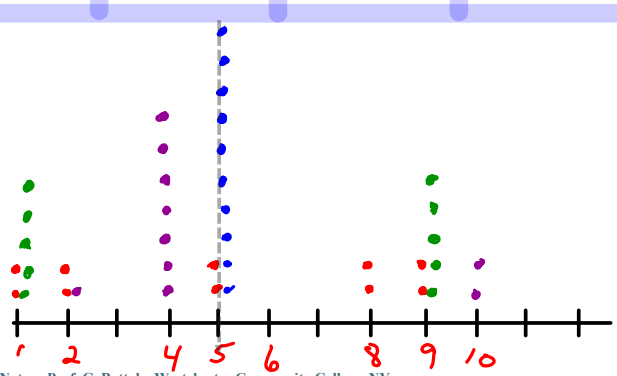
Statistics Home Page Class Notes Homework

3.2 Measures of Variation

f) Which better distinguishes the spread
Range or Stand dev?

Given the following: $\bar{x} = 5$

DATA_I		DATA_II		DATA_III		DATA_IV	
1	8	1	9	5	5	2	4
5	2	9	1	5	5	4	4
1	9	1	9	5	5	4	10
8	5	9	1	5	5	4	4
2	9	1	9	5	5	4	10
s	**3.33**		**4.22**		**0**		**2.71**
range	**9-1=8**		**9-1=8**		**5-5=0**		**10-2=8**



Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Computing Formula for the Standard Deviation

$$s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$$

Does not require finding the mean and doing subtraction before squaring. Easier to do arithemetic w/o calculator (historical). Now use calculator.

Data I	x_i^2	
x_i		
1		
5		
1		
8		
2		
8		
2		
9		
5		
9		

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Computing Formula for the Standard Deviation

$$s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}} = \sqrt{\frac{350 - (50)^2/10}{9}}$$

$= \sqrt{\frac{350 - \frac{2500}{10}}{9}}$

$= \sqrt{\frac{350 - 250}{9}} \quad \frac{100}{9}$

Data I	x_i^2	$\bar{x}=5$	
x_i		$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	1	-4	16
5	25	0	0
1	1	-4	16
8	64	3	9
2	4	-3	9
8	64	3	9
2	4	-3	9
9	81	4	16
5	25	0	0
9	81	4	16
50	350	$\Sigma = 0$	100

Required for defining formula. **Not required** for computing formula

$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$

$= \sqrt{\frac{100}{9}}$

$= \frac{10}{3} = 3.33$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Computing Formula for the Standard Deviation

$$s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$$

Data I		
x_i	x_i^2	
1		
9		
1		
9		
1		
9		
1		
9		
1		
9		

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Computing Formula for the Standard Deviation

$$s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$$

Data I		
x_i	x_i^2	
1	1	
9	81	
1	1	
9	81	
1	1	
9	81	
1	1	
9	81	
1	1	
9	81	
50	410	

$$s = \sqrt{\frac{410 - (50)^2/10}{9}}$$

s = 4.2

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

3.2 Measures of Variation

Know formulas:

mean, $\bar{x} = \frac{\Sigma x}{n}$

Defining Formula Std Dev, $s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$

Computing Formula Std Dev $s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

$s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$

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

$\mu = \frac{\Sigma x}{N}$

$s = \sqrt{\frac{\Sigma x_i^2 - (\Sigma x_i)^2/n}{n-1}}$

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework $\bar{x}$

Attachments

studentHeight_rev.xlsx