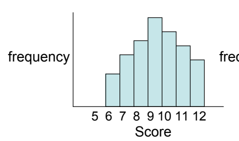
Measures of Dispersion (variability)

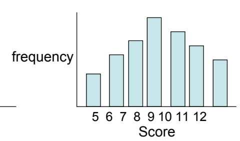
I. Introduction
II. Range
III. Average Deviation
IV. SS
V. Variance and Standard Deviation

I. Introduction: Why Measure Dispersion?

A. Describing a distribution's "spread"

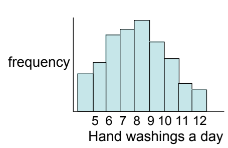
Two distribution with the same mean and different "spread"





I. Introduction: Why Measure Dispersion?

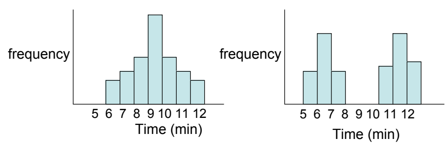
B. To compare a given score to the rest of the distribution



I. Introduction: Why Measure Dispersion?

C. To make informed decisions:

Two routes to campus - same average time



Measures of Dispersion

Range: describes the spread by giving contrasting the two most extreme scores

Three ways to report it

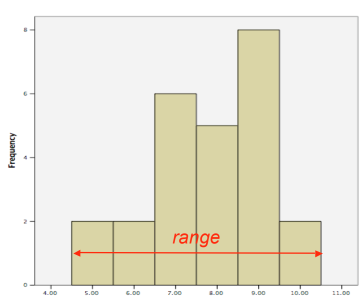
1. Give both scores

"The grades ranged between 56 and 97."

2. The highest score minus the lowest score:

97-56 = 41

Note, to be accurate, you should use the upper real limit of the high number and the lower real limit of the low number.



10.5 - 4.5 = 6

Measures of Dispersion

Range (continued)

3. High score - Low score + 1

If the scores are integers, adding 1 adds in the upper and lower real limits.

Examples:

10-5+1 = 6

if 97.5 is the URL, and 55.5 the LRL

then 97.5-55.5 = 42

and 97-56+1 = 42

NOTE: This is the convention we will use in this course!

Measures of Dispersion

Problems with the range:

1) Completely determined by extreme scores.

2) Not very useful for describing the shape of the distribution or drawing conclusions.

Measures of Dispersion

Average deviation:

Average distance each score is from the mean of the distribution.

	x	x-*M*		|x-*M*|
	8	8-3 =	5	5
	1	1-3 =	-2	2
	3	3-3 =	0	0
	0	0-3 =	-3	3
Σ =	12		0	10

M = 12/4 = 3.0 Av. Dev. = 10/4 = 2.5

Measures of Dispersion

Average Deviation (continued)

Conceptually very useful (closely related to the standard deviation), but it is unstable under sampling.

Measures of Dispersion

SS (sum of the squared deviations)

	x	x-*M*		(x-*M*)^2
	8	8-3 =	5	25
	1	1-3 =	-2	4
	3	3-3 =	0	0
	0	0-3 =	-3	9
Σ=	12		0	38

M = 12/4 = 3.0

$SS = \Sigma(x-M)^2$

$= 38$

Measures of Dispersion

Variance (σ^2)

Population Variance is the average of the squared deviations from the mean.

$$\sigma^2 = \frac{\Sigma(x-M)^2}{N} = SS/N$$ (N = population size)

example 38/4 = 9.5

Measures of Dispersion

Standard Deviation (σ) for a population

Square root of variance

$\sigma = \sqrt{\sigma^2}$

$= \sqrt{\frac{\Sigma(x-M)^2}{N}}$

$= \sqrt{SS/N}$

Example: $\sqrt{9.5} = 3.08$

Measures of Dispersion

Summary

Range = high-low+1

SS = $\Sigma(x-M)^2$

	Population
Variance:	$\sigma^2 = SS/N$
Standard Deviation:	$\sigma = \sqrt{SS/N}$

Measures of Dispersion

Example problem: Calculate the range, SS, variance, and standard deviation for the following populations of scores:

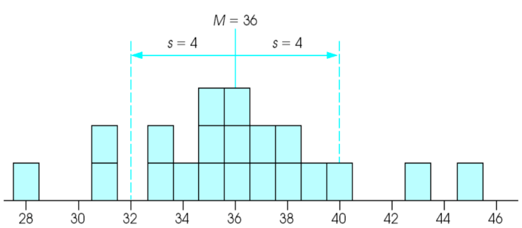
	Group 1	Group 2
	1	6
	9	9
	5	5
	8	8
	7	7
Range		
Mean		
SS		
Var		
Std		

Measures of Dispersion

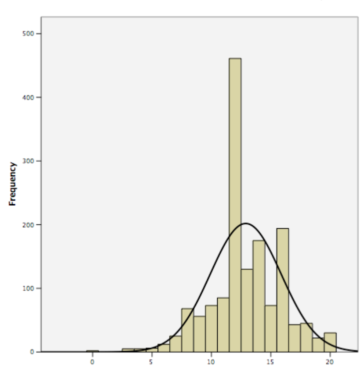
Visualizing what the standard deviation measures.

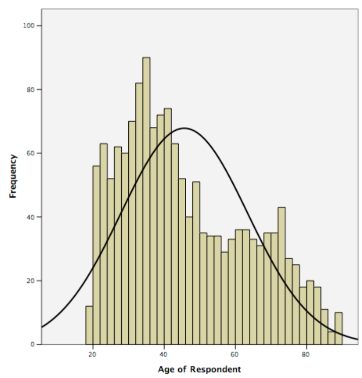
- about halfway to the range
- the point of inflection on a frequency graph
- most of the scores (68 %) should fall within 1

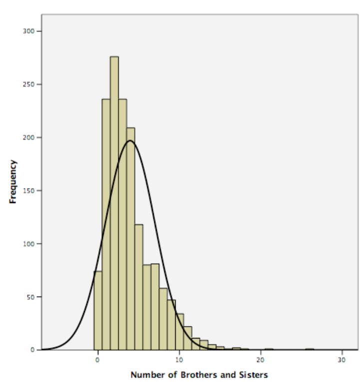
standard deviation of the mean.



(data from 1991 U.S. General Social Survey courtesy of SPSS)







Measures of Dispersion

SS and the computational formula

Definition: $SS = \Sigma(x-M)^2$

Computational: $SS = \Sigma x^2 - (\Sigma x)^2/N$

Examples:

Measures of Dispersion

Degrees of freedom and populations estimates

Definition: number of pieces of information that are free to vary.

Example: cafeteria line

Measures of Dispersion

Degrees of freedom and deviations scores

x	x-M	
2	-2	
4	0	df = 3
6	2	
Σ = 10		
M = 4		

Measures of Dispersion

Because deviation scores are based on the mean, once I have calculated the mean I have reduced my degrees of freedom by one. Thus, in sample estimates of variance, the degrees of freedom are n-1.

In general *df* = n - 1

Measures of Dispersion

Define:

s^2 = sample variance = SS/n-1

s = sample standard deviation = $\sqrt{SS/n-1}$

Measures of Dispersion

Summary

Range = high-low+1

SS = $\Sigma(x-M)^2$

	Population	Sample
Variance:	σ^2=SS/N	s^2=SS/n-1
Standard Deviation:	$\sigma=\sqrt{SS/N}$	s= $\sqrt{SS/n-1}$

Measures of Dispersion

Sample Problems:

Measures of Dispersion

Factors Affecting Variability

1. Sample Size
   a) Range increases with sample size
   b) Variance decreases with sample size
2. Extreme scores - greatly influence most measures, especially range