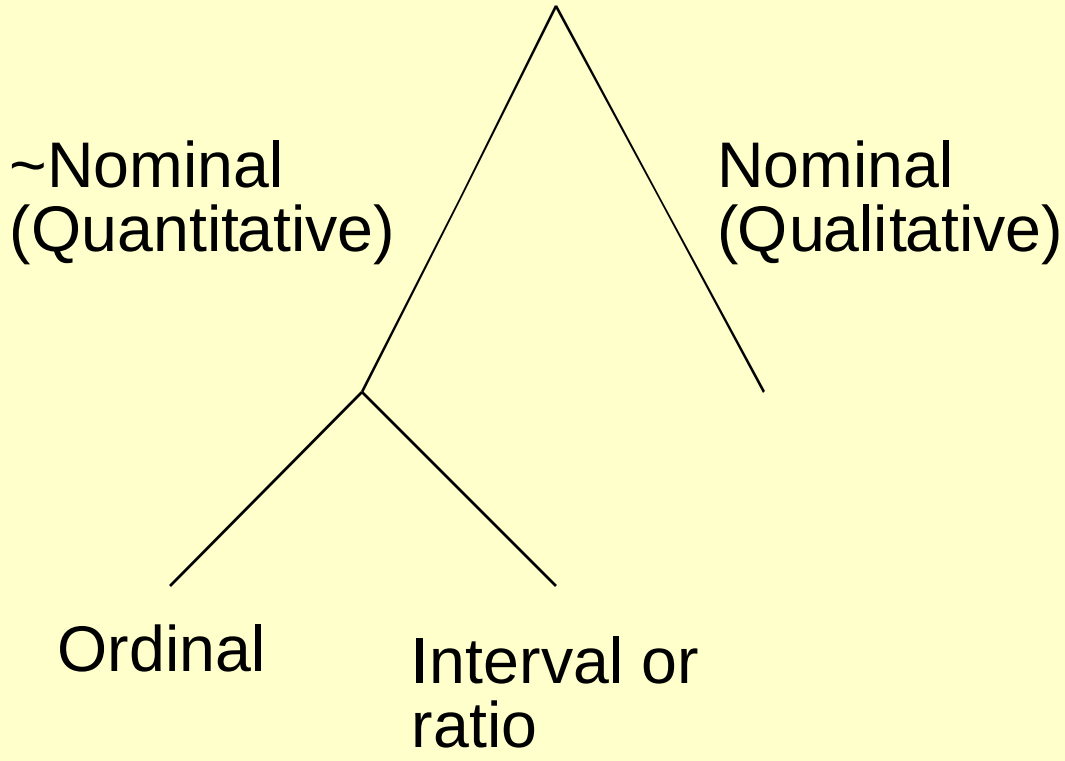
Introduction to Descriptive Statistics

2/21/02

Population vs. Sample Notation

Population	*Vs*	Sample
Greeks		Romans
:, F, $		s, b

Types of Variables



Describing data

	Moment	Non-mean based measure
Center	Mean	Mode, median
Spread	Variance (standard deviation)	Range, Interquartile range
Skew	Skewness	--
Peaked	Kurtosis	--

Mean

$$\frac{\sum_{i=1}^{n} x_i}{n} \equiv m \equiv \overline{X}$$

Variance, Standard Deviation

$$\sum_{x=1}^{n} \frac{(x_i - m)^2}{n} \equiv s^2,$$

$$\sqrt{\sum_{x=1}^{n} \frac{(x_i - m)^2}{n}} \equiv s$$

Variance, S.D. of a Sample

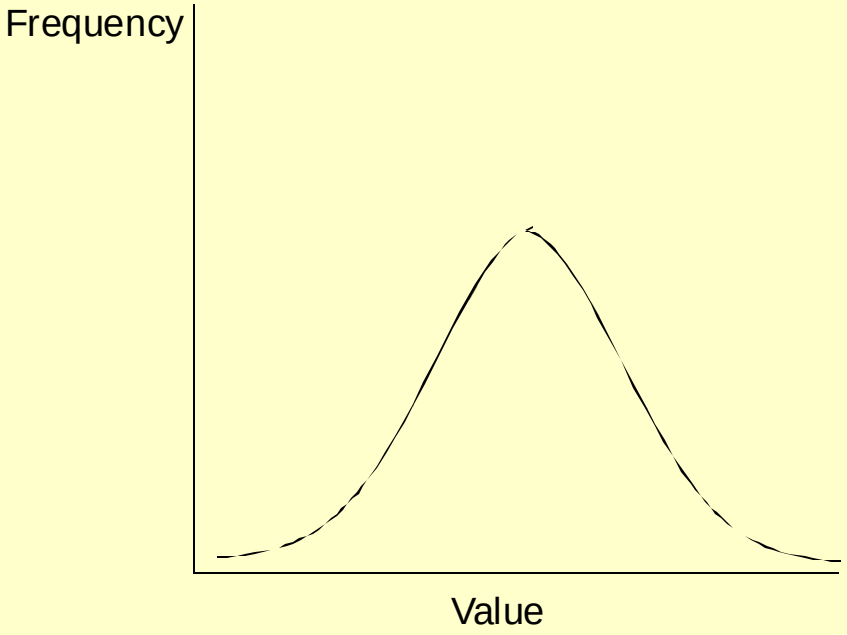
$$\sum_{x=1}^{n} \frac{(x_i - m)^2}{n-1} \equiv s^2,$$

$$\sqrt{\sum_{x=1}^{n} \frac{(x_i - m)^2}{n-1}} \equiv s$$

Coefficient of variation

$$c.v. = \frac{s}{m} \times 100$$

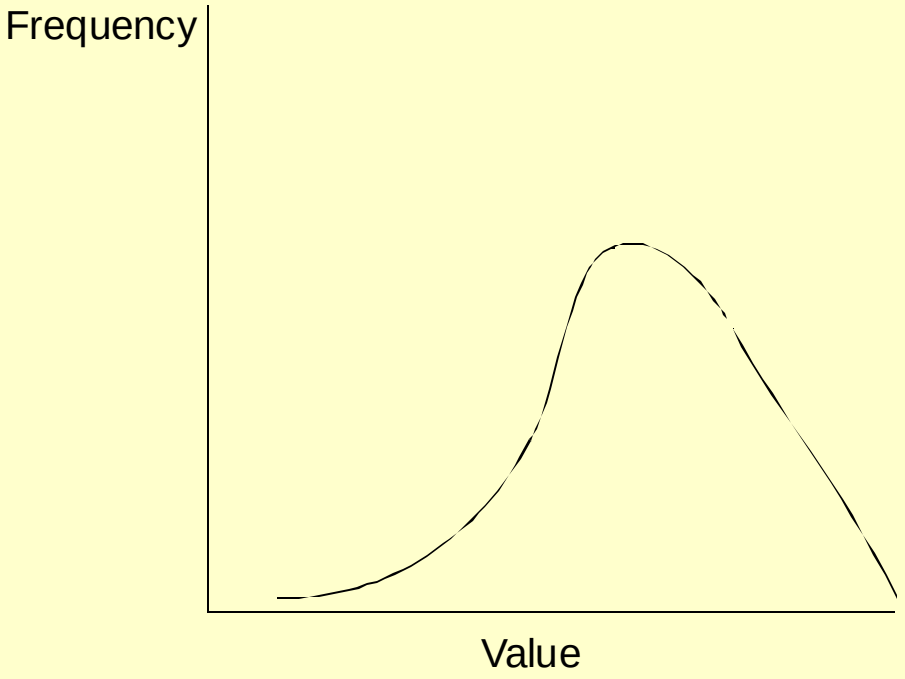
Skewness
Symmetrical distribution



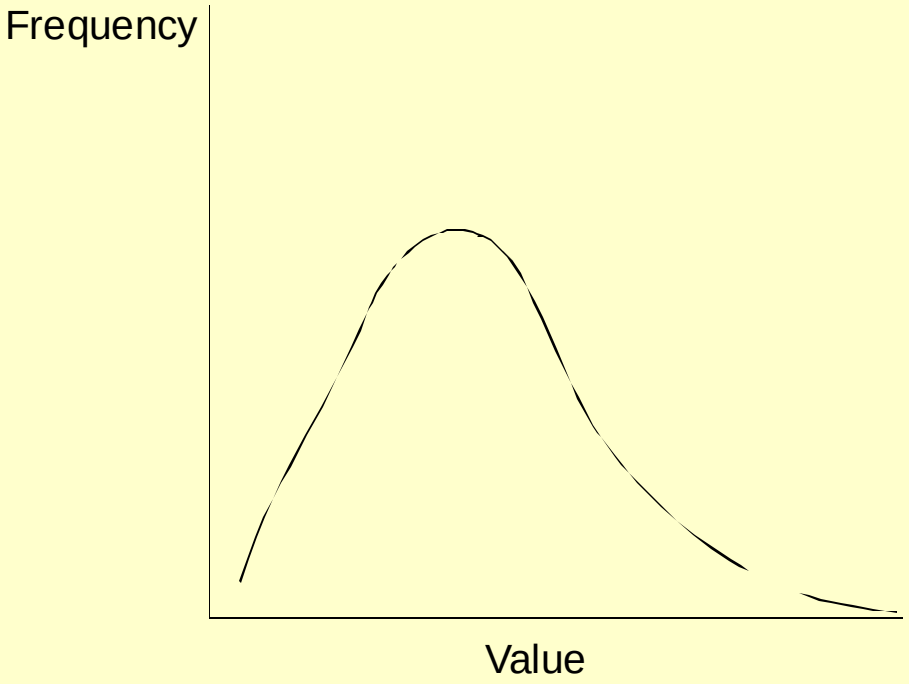
- IQ
- SAT

Skewness
Asymmetrical distribution

- GPA of MIT students



Skewness (Asymmetrical distribution)



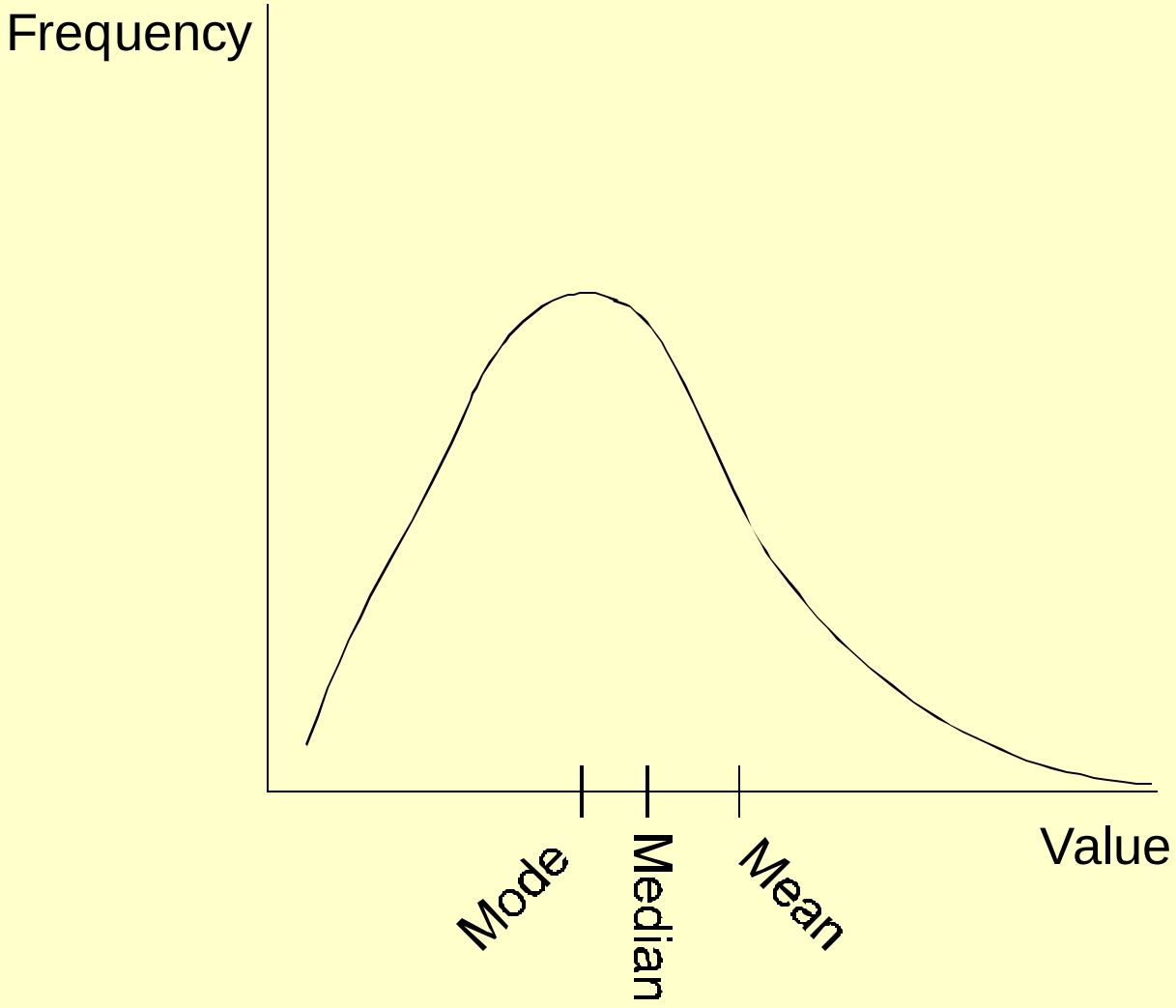
- Income
- Contribution to candidates
- Populations of countries
- “Residual vote” rates

Skewness

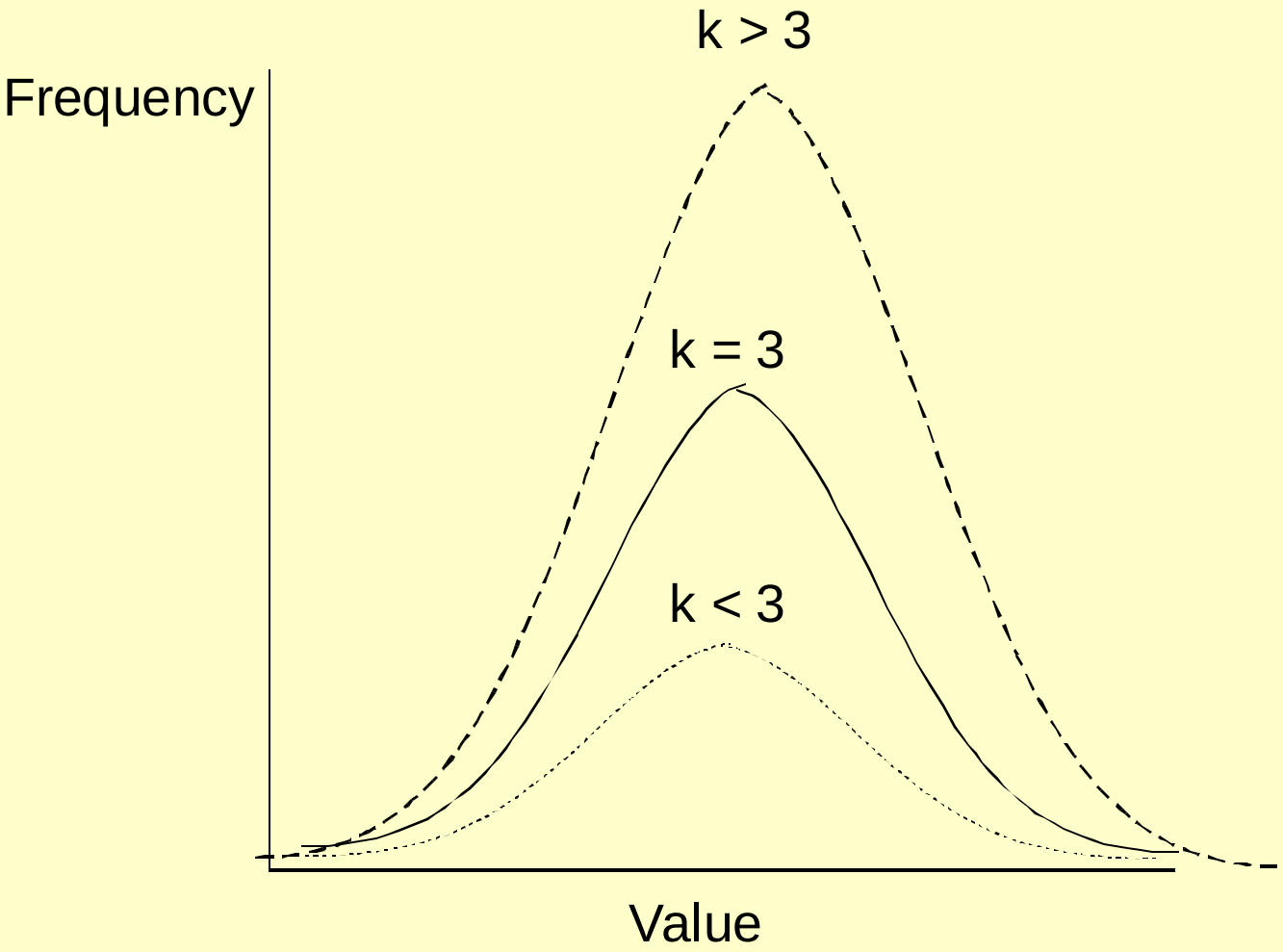
$$\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^3}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^{3/2}}\times\frac{(n-1)^{3/2}}{n-2}\approx$$

$$(mean-\mathrm{mod}\,e)/s\approx 3\times(mean-median)/s$$

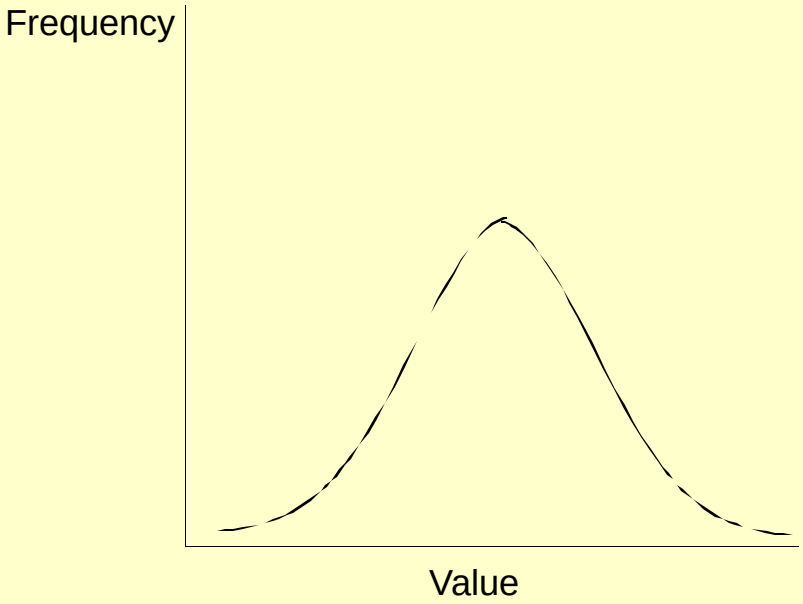
Skewness



Kurtosis



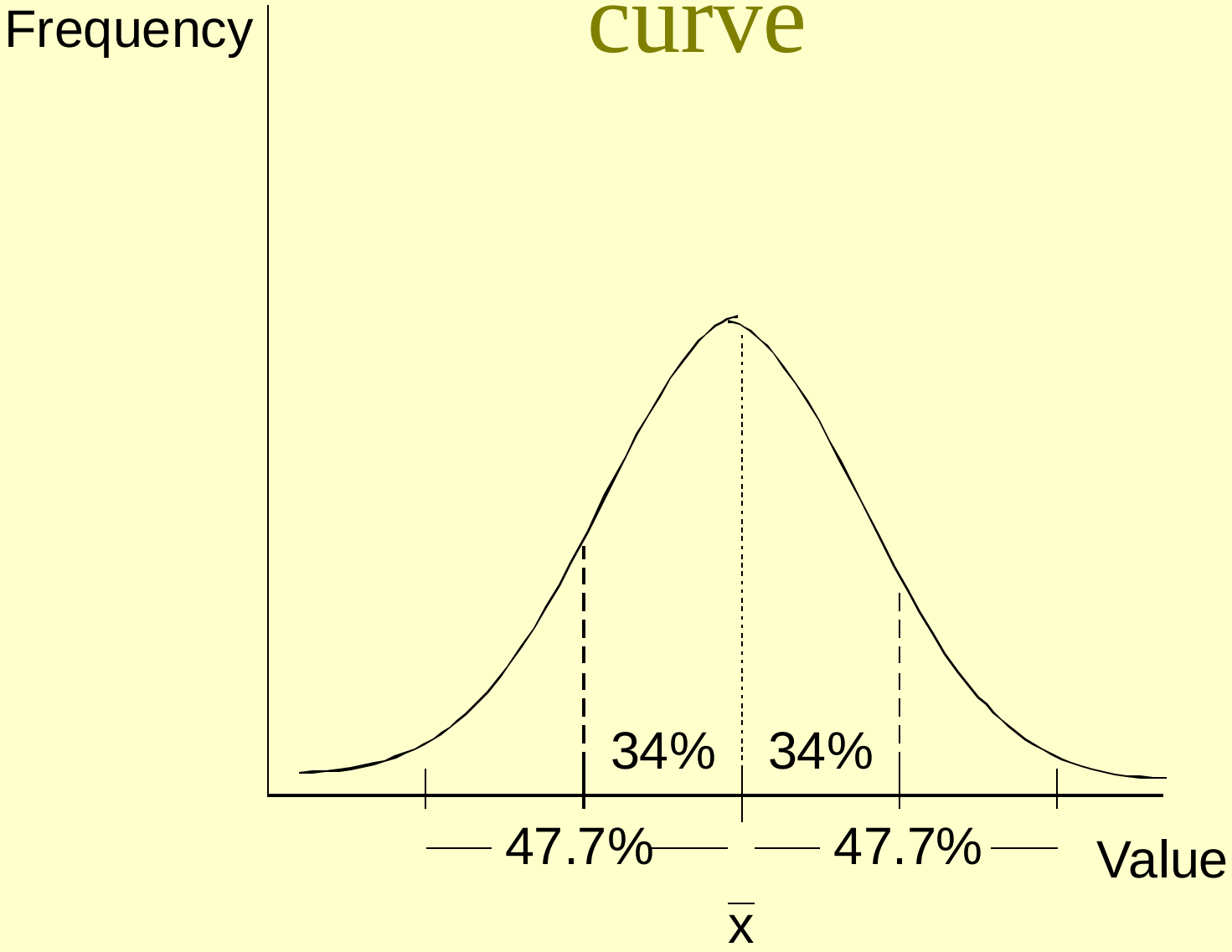
A few words about the normal curve



- Skewness = 0
- Kurtosis = 3

$$f(x) = \frac{1}{s\sqrt{2p}} e^{-(x-m)/2s^2}$$

More words about the normal curve



CEG example

The instructor and/or section leader:					
	Mean	s.d.	Skew	Kurt	Graph
Gives well-prepared, relevant presentations	5.9	0.75	-1.5	5.8	
Explains clearly and answers questions well	5.8	0.78	-1.1	4.1	
Uses the blackboard well	5.5	0.93	-0.61	3.0	
Speaks clearly	6.2	0.59	-1.1	6.4	
Is available outside of class	5.5	0.80	-0.55	2.8	
Is friendly and supportive	6.1	0.71	-1.0	3.8	
Overall rating	5.8	0.79	-1.1	4.1	

Binary data

$$\bar{X} = prob(X) = 1 = \text{proportion of time } x = 1$$

$$s_x^2 = \bar{x}(1-\bar{x}) \Rightarrow s_x = \sqrt{\bar{x}(1-\bar{x})}$$

Commands in STAT for getting univariate statistics

- summarize
- summarize, detail
- graph, bin() normal
- graph, box
- tabulate [NB: compare to table]