

this histogram;
 time: mean, 50
 next
 time: normal curve

read: JJ
 ch. 1, 3, 3

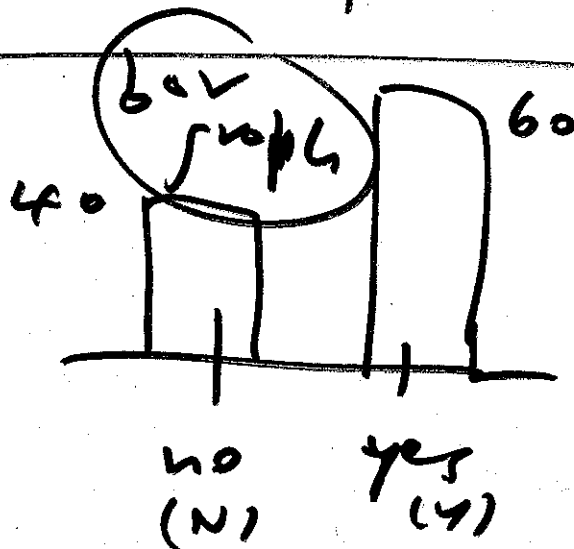
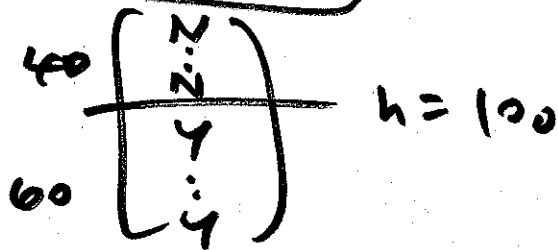
AMS7
 13 Jan
 08

course
 packet

①

(690 pages: 300 page book, 300 page
 lecture notes, 100 page reader) now
 available at back of room: \$60
 cash
 productive
 cost

work 1 due Tue 20 Jan | disc. sec
 week; no labs this week; labs
 & disc. sec. next week &
 preposter



butterfly wing length (cm)

n = 24

- 4.4
- 3.6
- .
- .
- 3.9.

sort →

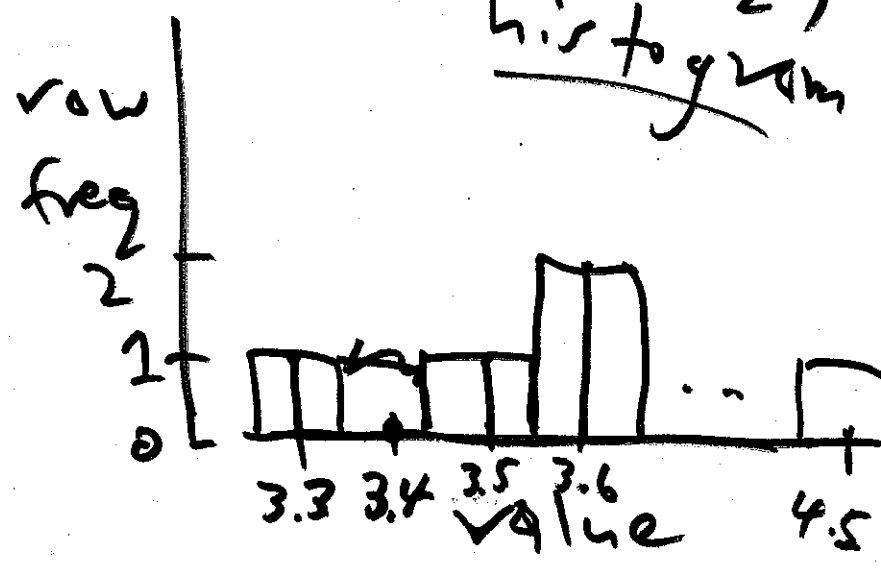
- 3.3
- 3.5
- 3.6
- 3.6
- .
- 4.5

value (row frequency)
(count)

3.3	1
3.4	0
3.5	1
3.6	2
.	.
.	1
4.5	1

n = 24

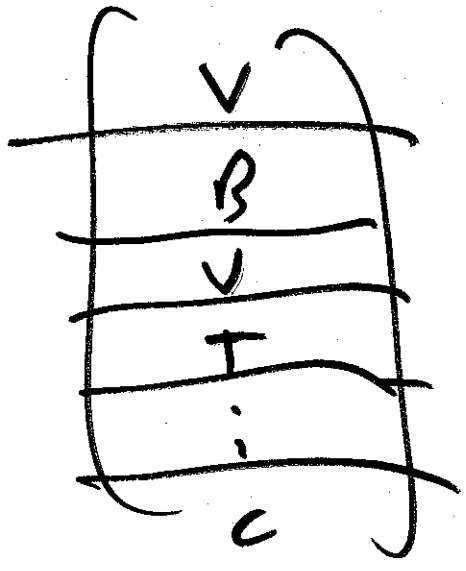
(row freq.)
histogram



(JMP does this backwards as its default)

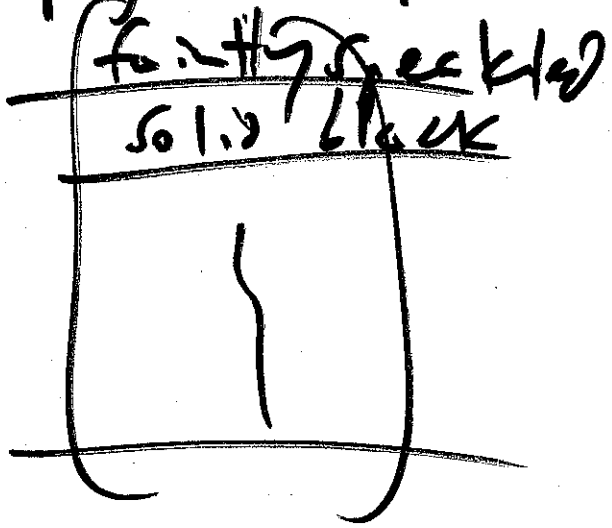
one secret to good statistics work: visualize the raw data.

location



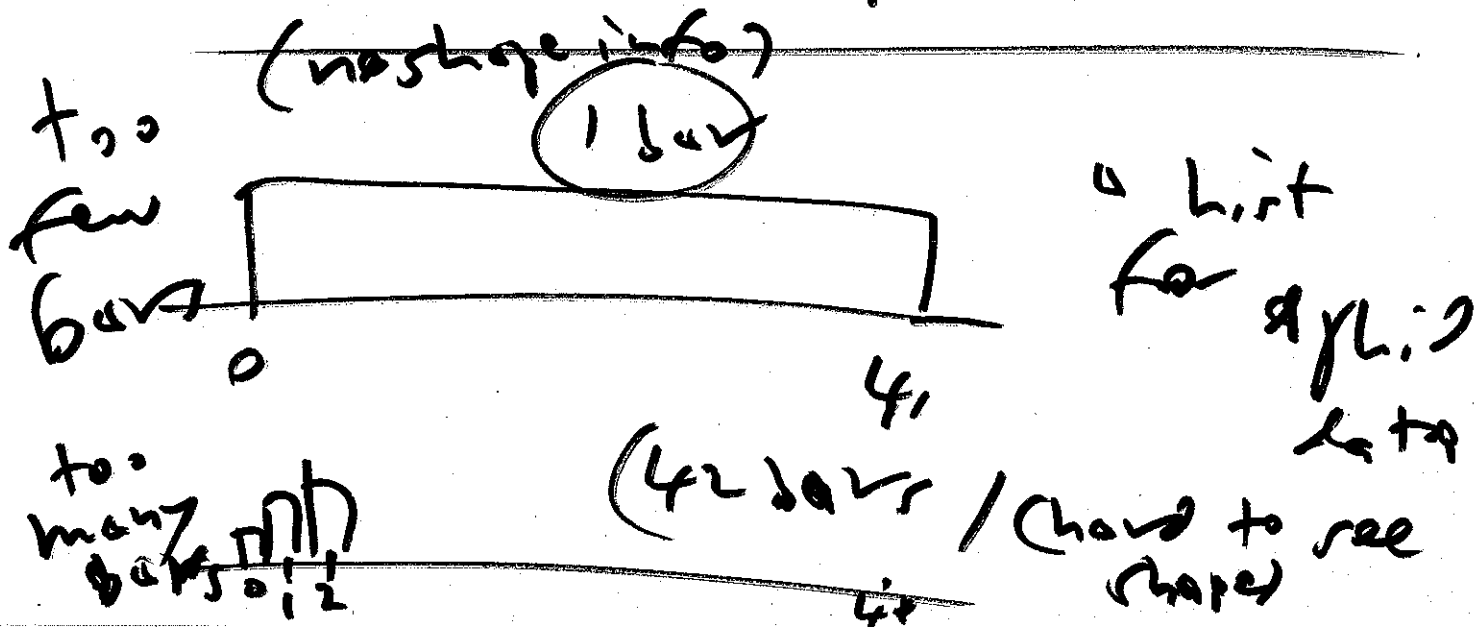
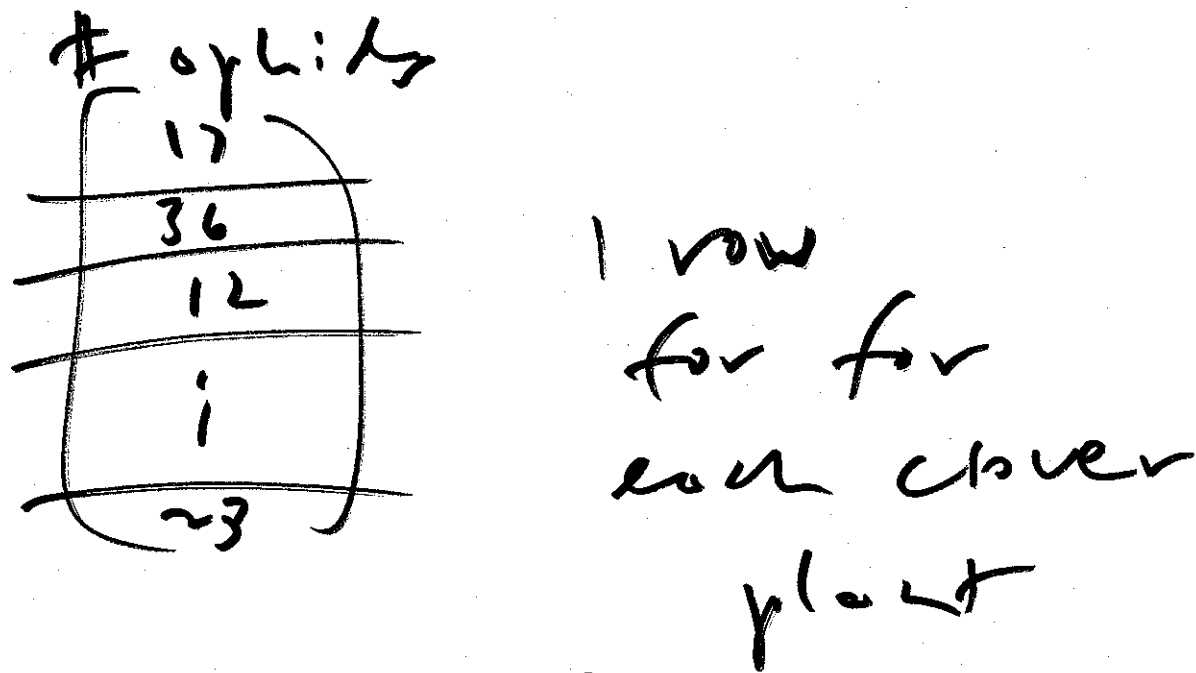
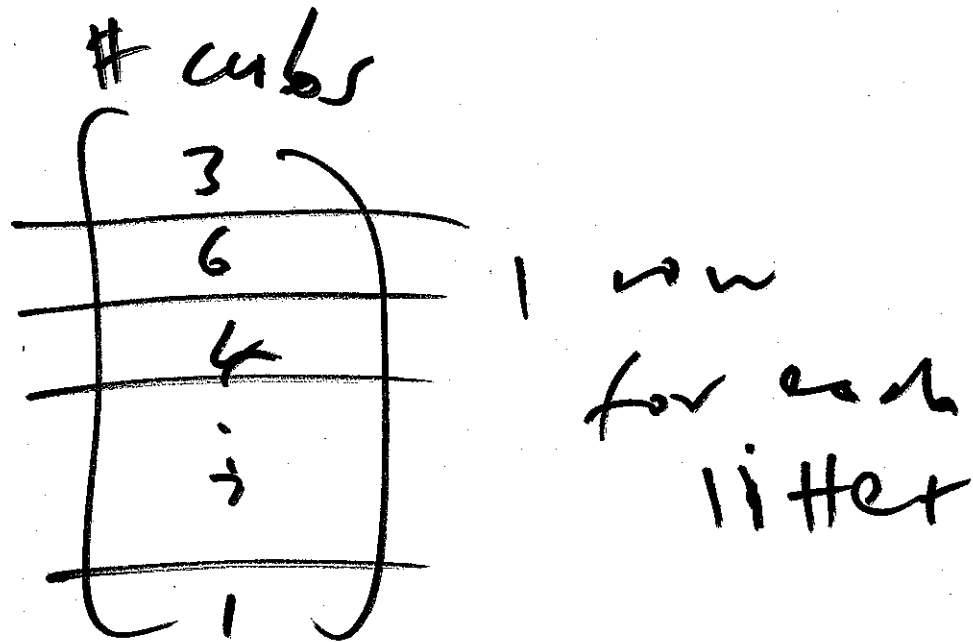
$n = 251$

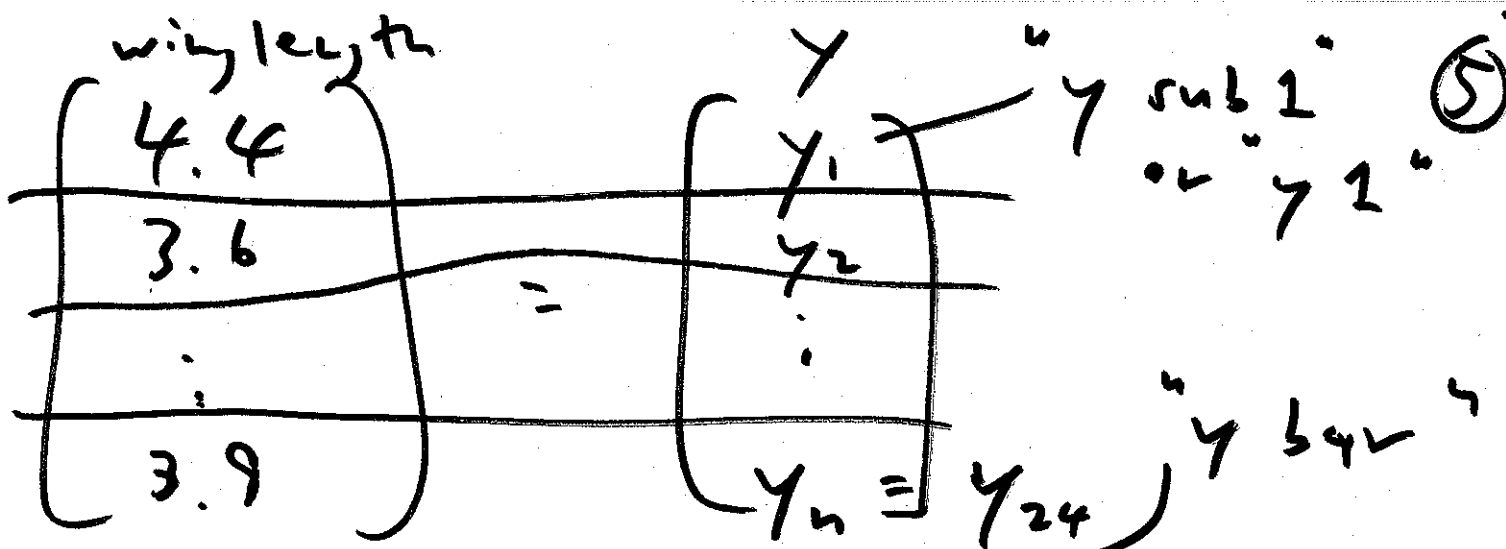
pigmentation



1 row for each sunfish

$n = 154$





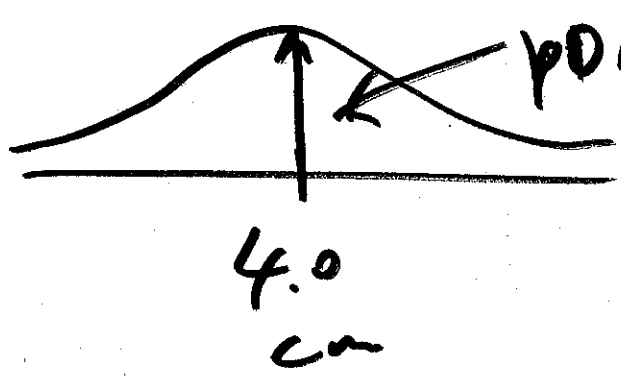
mean $\frac{95}{24} = 3.96$ cm

mean $\bar{y} = \frac{y_1 + y_2 + \dots + y_n}{n}$

capital sigma

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

↑
index of summation



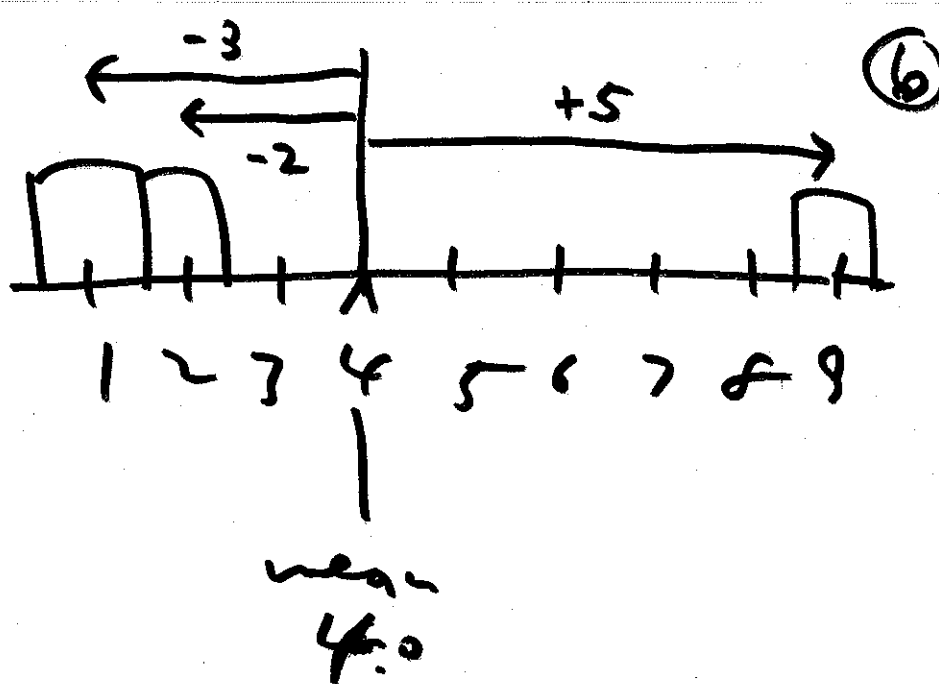
point of highest frequency = mode

$\begin{pmatrix} 1 \\ 2 \\ 9 \end{pmatrix}$

$n=3$

mean $\bar{y} = 4$

mean = balance point of dataset



another measure of center: median

$\begin{pmatrix} 3.3 \\ 4.5 \end{pmatrix}$

middle of data
sorted from smallest
to largest

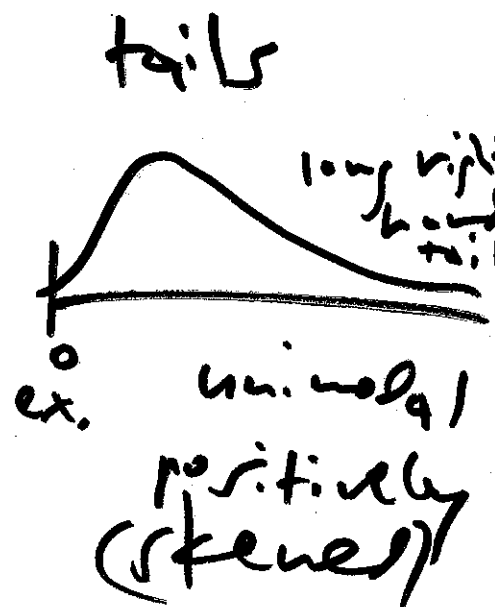
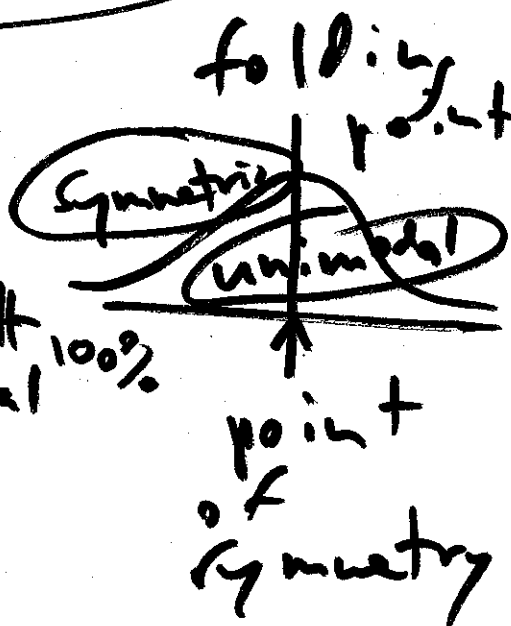
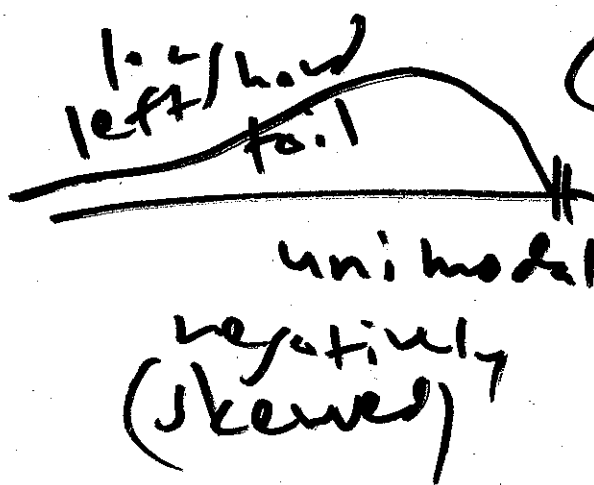
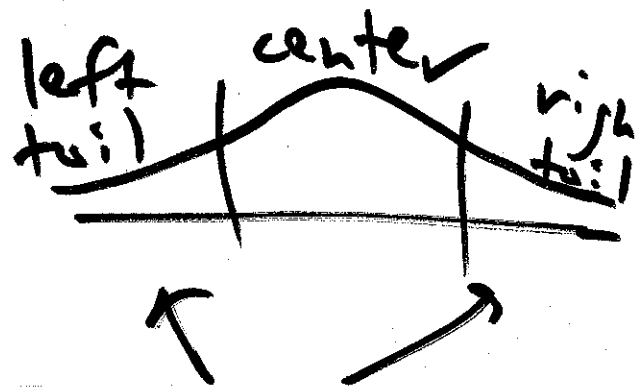
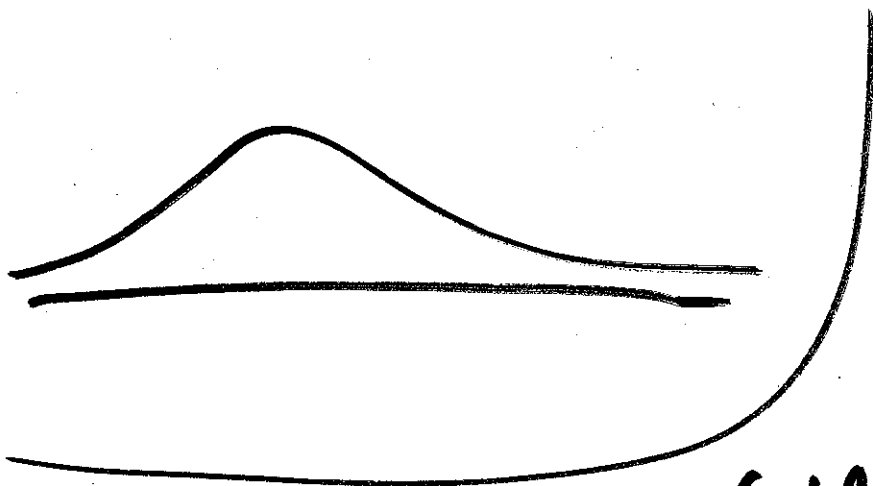
$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 9 \end{pmatrix}$

median

$2.5 = \tilde{y}$ ← "y tilde"

median $\begin{pmatrix} 1 \\ 2 \\ 9 \end{pmatrix}$

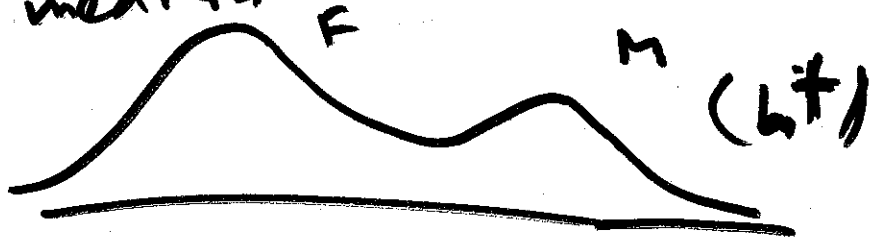
Graphical interpretation of median ②



= mean
 = mode
 = median

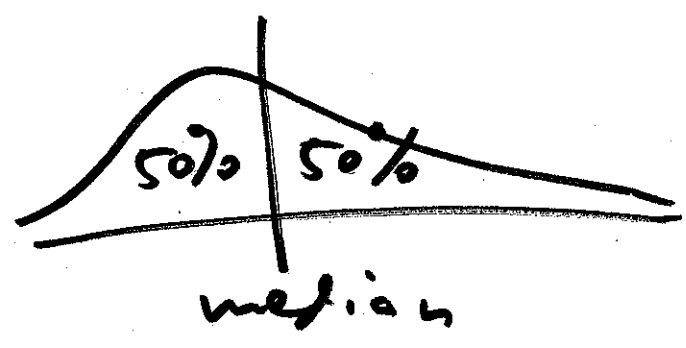


unimodal



bimodal

dist. # modes
 (multimodal) 2



raw
freq

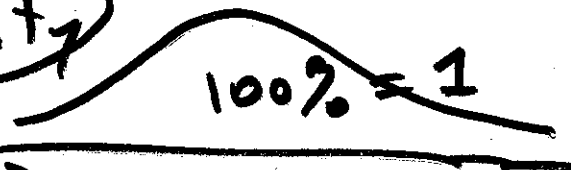
value

default:
sketch hist
are always on
density scale

relative
freq

value

density



on density scale
relative

value

freq into = area
under curve
approximating hist

density
scale