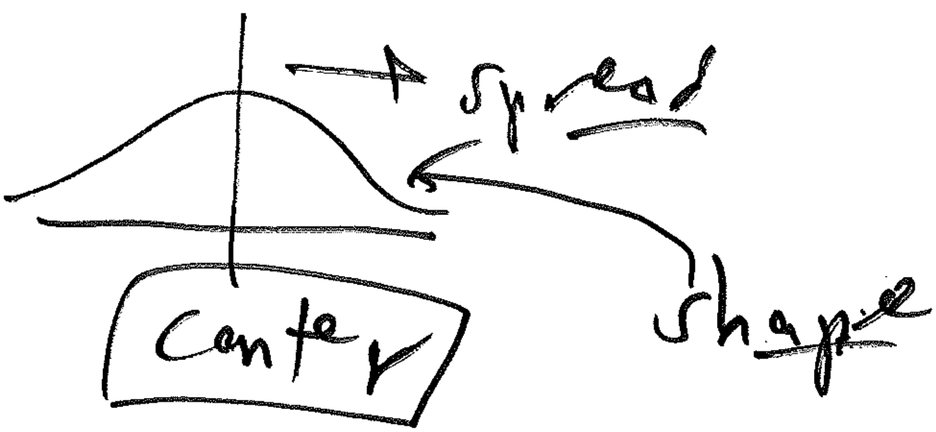
this time: measures of center & spread; normal curve

read: AMS 7 6 Oct (1)
DD (A) Ch. 1-3
(B) Ch. 1-6
LN pp. L-1-85

next time: experimental design

can turn homework 1 in on Fri or over weekend, in box outside Baskin 357C

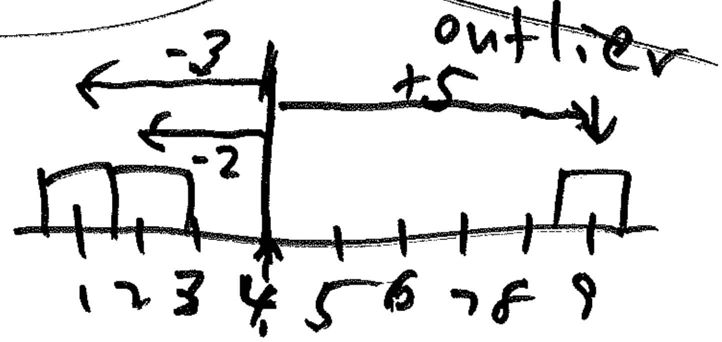
get course materials at Literary Guillotine: tomorrow 10 AM → $45



graphical interp. of mean

$$\begin{array}{c} y_1 \\ \vdots \\ y_n \end{array} \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \quad n = 3 \quad \xrightarrow{\text{subtract } 4} \quad \begin{bmatrix} -3 \\ -2 \\ +5 \end{bmatrix}$$

mean $\bar{y} = 4$ mean 0

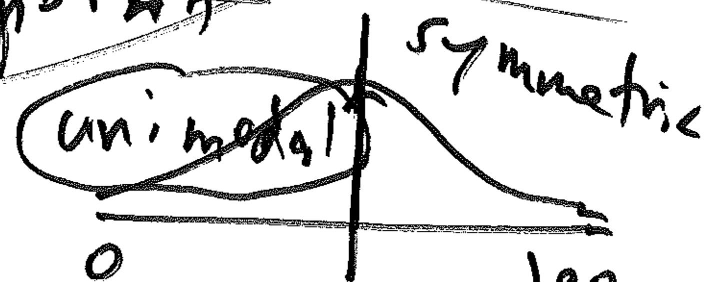


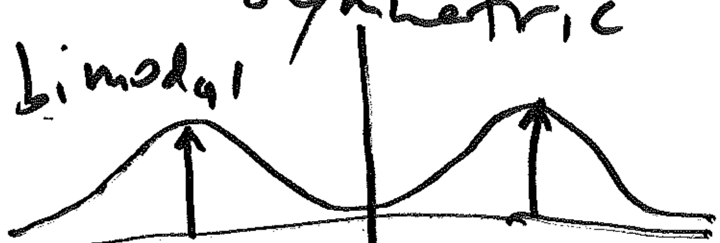
$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \xrightarrow[\bar{y}]{\text{subtract}} \begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{bmatrix} n$$

mean $\bar{y}$ — mean 0 ✓

deviations from mean

mean = balance point

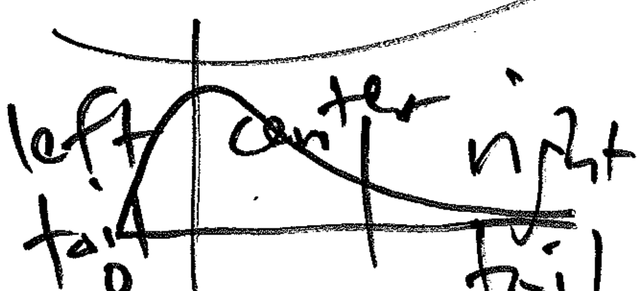


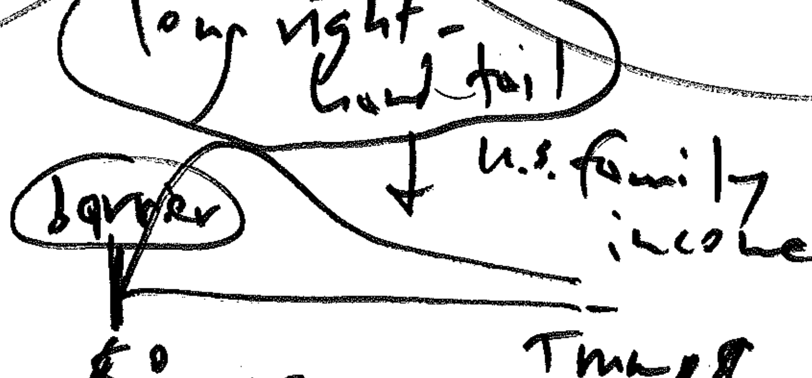


ex. of <u>multimodal</u>

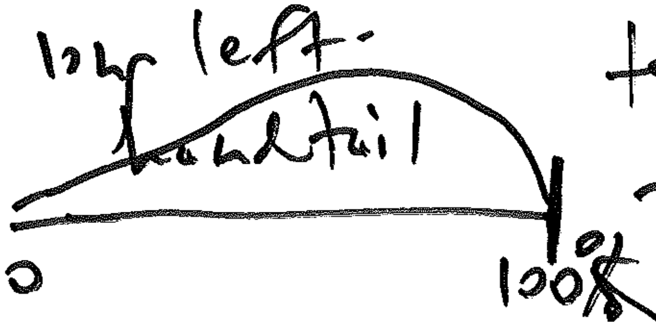
point of symmetry = mean if hist is symmetric = mode

<u>mode</u> = highest point on hist (in freq)



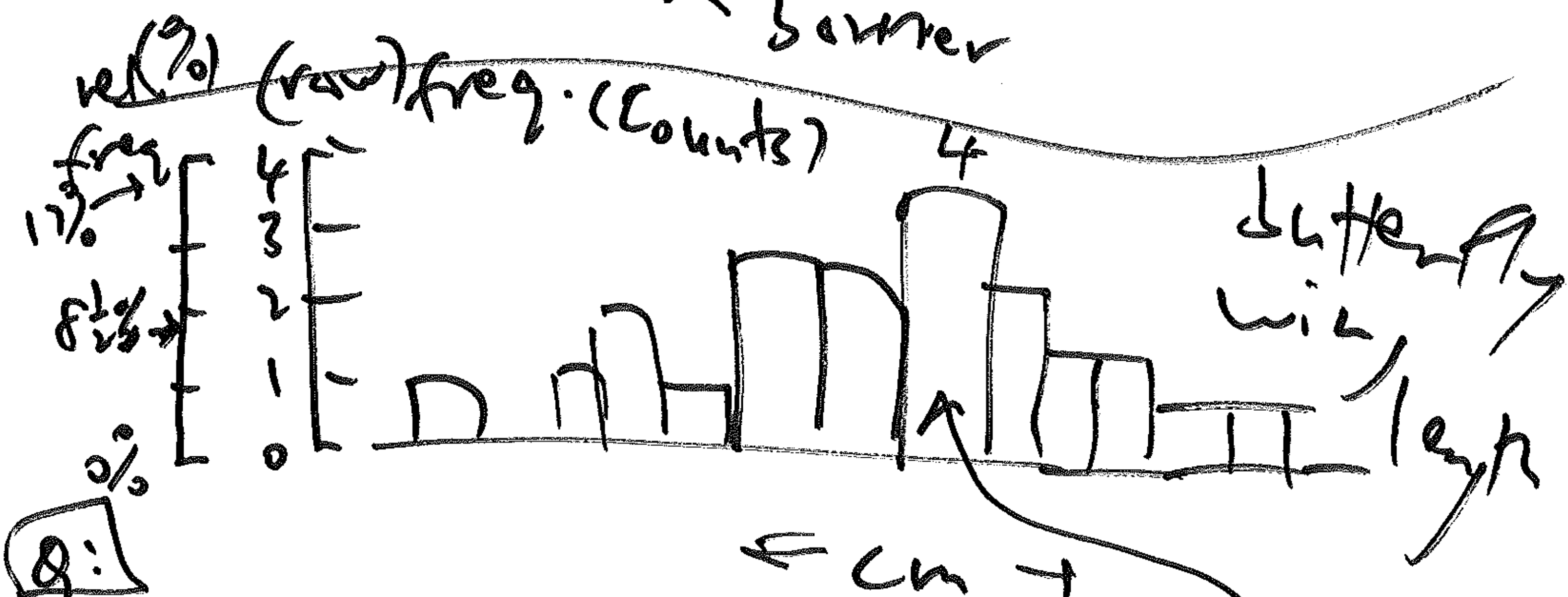


skewed ≡ asymmetric



test scores from easy test

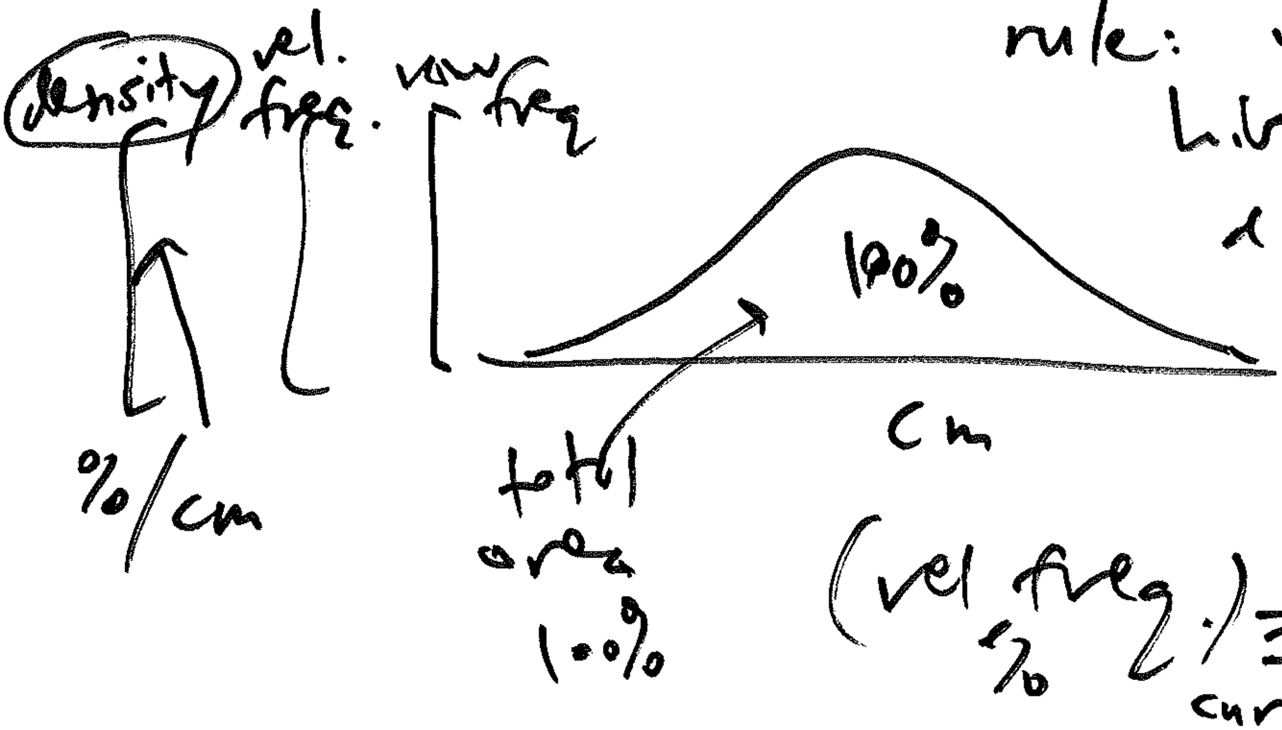
barrier



Q: what relative freq. is in highest bar?

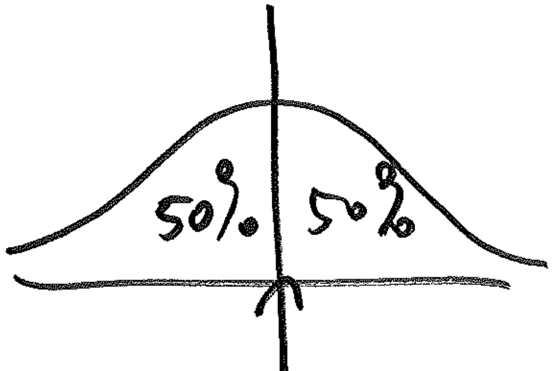
A: $\frac{4}{24} = \frac{1}{6} \cdot 100\% \doteq 17\%$ (16.67..)

$n = 24$



rule: when hist are drawn on density scale

(rel freq.)$_{\%}$ $\equiv$ area under curve (hist.)



pt. of sym.
= mean
= mode
= median

convention:

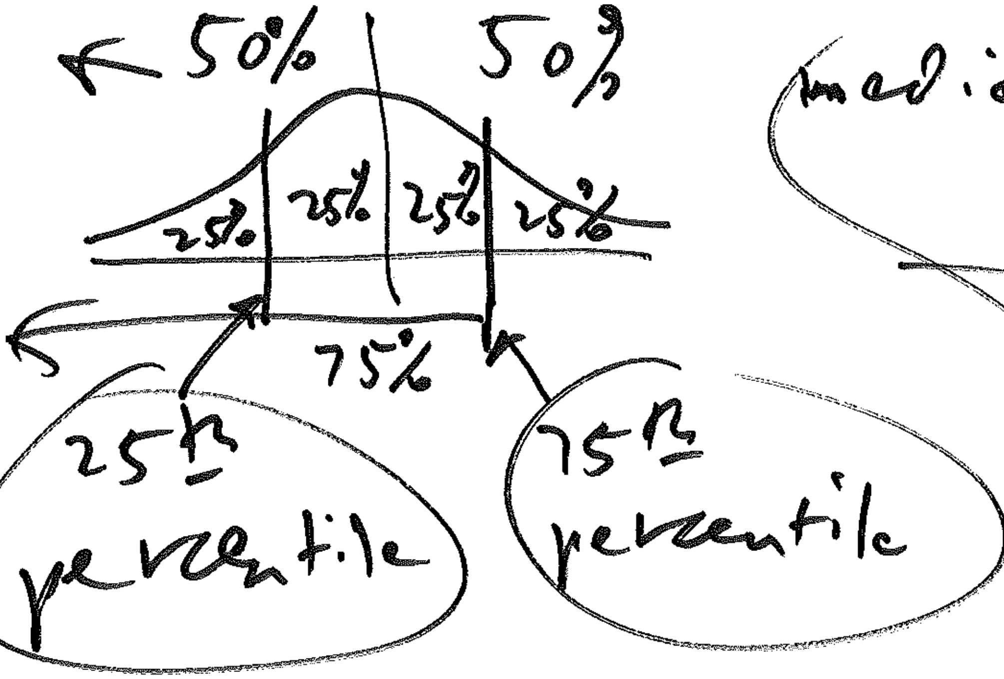
all hist in this class from now on will be on density scale

median = 50/50 pt. in hist. in rel. freq terms

$$\begin{bmatrix} 1 \\ 9 \\ 2 \end{bmatrix} \xrightarrow{\text{sort}} \begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \leftarrow \text{median}$$

$$\begin{bmatrix} 9 \\ 1 \\ 3 \\ 2 \end{bmatrix} \xrightarrow{\text{sort}} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 9 \end{bmatrix} \leftarrow \text{mean of } \frac{2+3}{2} = 2.5$$

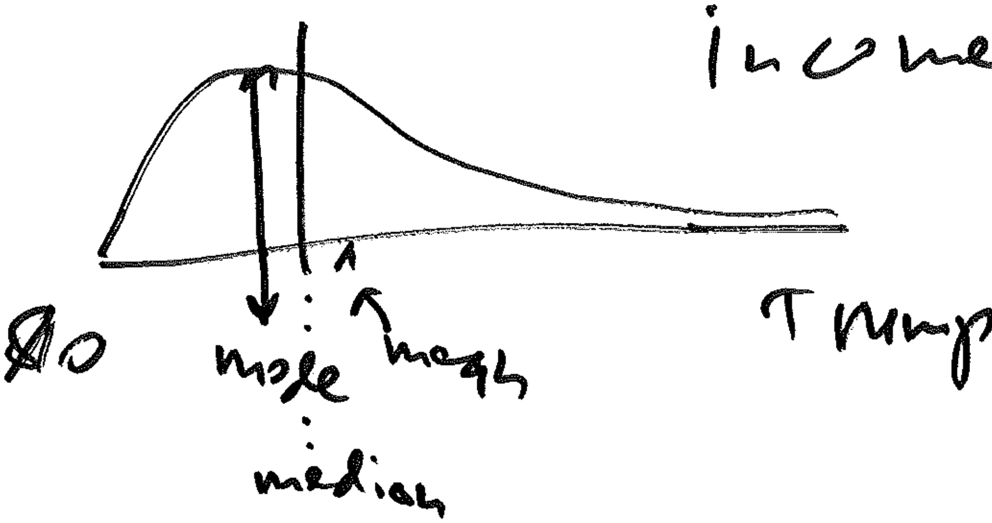
median = 2.5

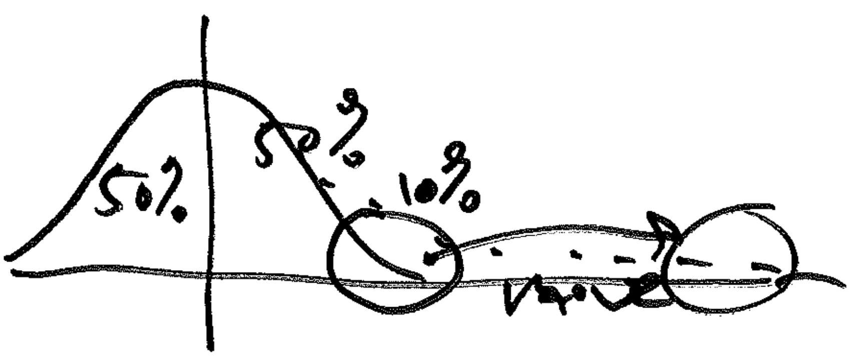


median = 50th percentile

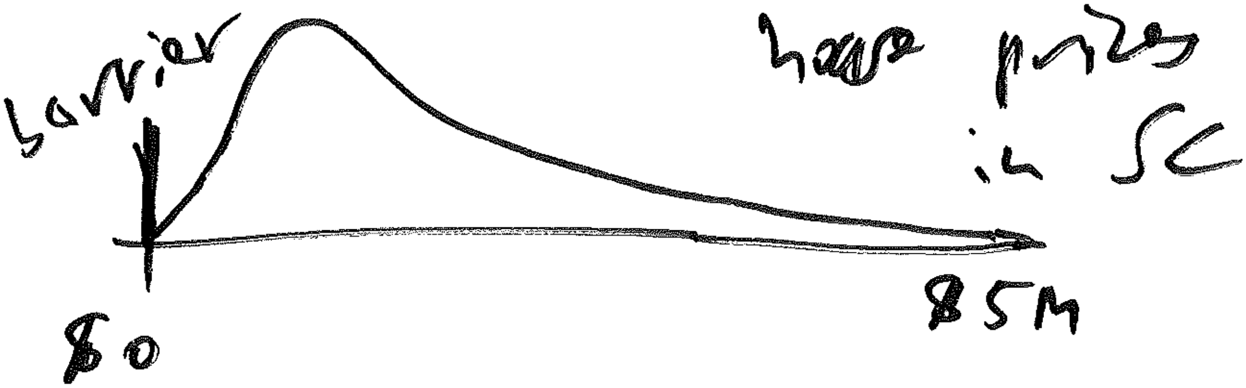
skewed dist?

income





median stays same
mean ↑



idea 1

$$\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix} \xrightarrow[4]{\text{substract}} \begin{pmatrix} -3 \\ -2 \\ +5 \end{pmatrix} \xrightarrow[|\cdot|]{\text{take}} \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$$

absolute value mean $\frac{10}{3} \approx 3.3$

mean absolute deviation (from mean) (MAD)

not used much

idea 2

square

$$\begin{pmatrix} \$1 \\ \$2 \\ \$9 \end{pmatrix} \xrightarrow[\$4]{\text{subtract}} \begin{pmatrix} \$-3 \\ \$-2 \\ \$+5 \end{pmatrix} \xrightarrow{\text{square}} \begin{array}{l} (-3)^2 = +9\,\$^2 \\ (-2^2) = +4\,\$^2 \\ (+5)^2 = +25\,\$^2 \end{array}$$

mean $\bar{y} = \$4$

mean $\frac{38}{3} = 12.7\ \$^2$

final step: take $\sqrt{\ }$ $\rightarrow \sqrt{\$^2\, 12.7} \approx \$3.6$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \xrightarrow[\bar{y}]{\text{subtract}} \begin{bmatrix} y_1 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{bmatrix} \xrightarrow{\text{square}} \begin{bmatrix} (y_1 - \bar{y})^2 \\ \vdots \\ (y_n - \bar{y})^2 \end{bmatrix}$$

mean $\bar{y}$

mean

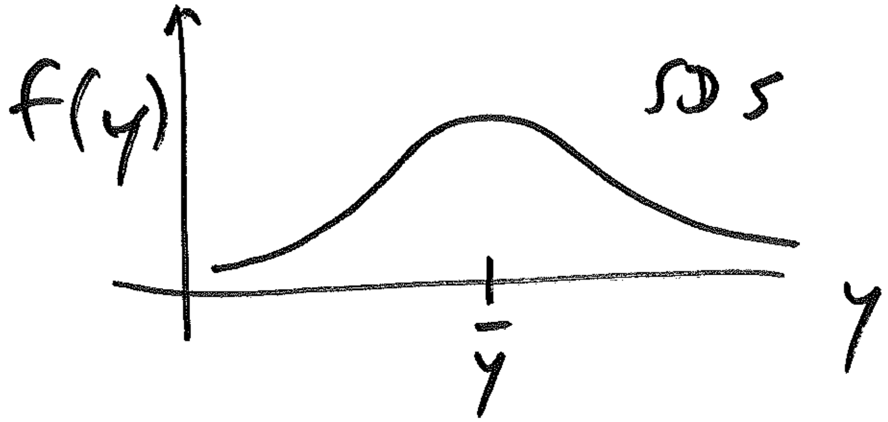
$$\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2} \overset{\text{def}}{=} \text{(sample) standard deviation (SD)} = s$$

trick

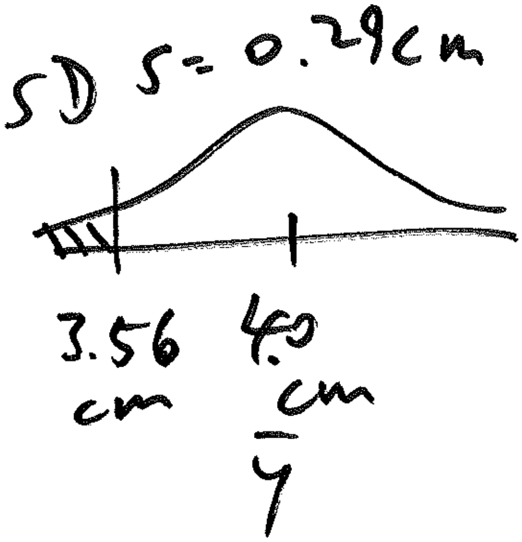
square of SD $= \dfrac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$

= (sample) <u>variance</u>

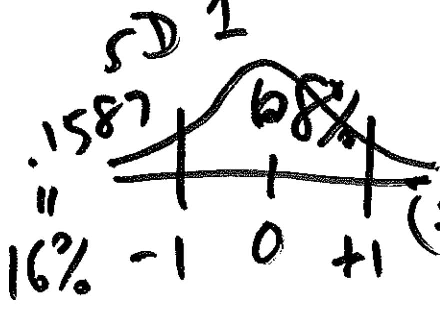
$= s^2$



$$f(y) = \frac{1}{s\sqrt{2\pi}} \exp\left(-\frac{(y-\bar{y})^2}{2s^2}\right)$$



<u>fact</u>: <u>all</u> normal curves satisfy empirical rule <u>exactly</u>



standard normal curve