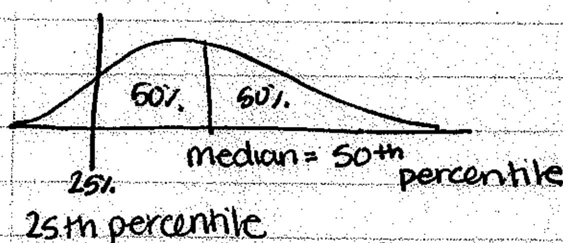
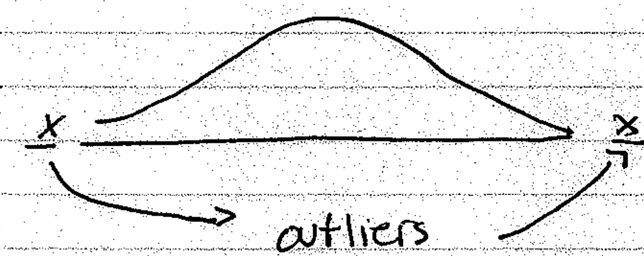
Caitlin Doughty

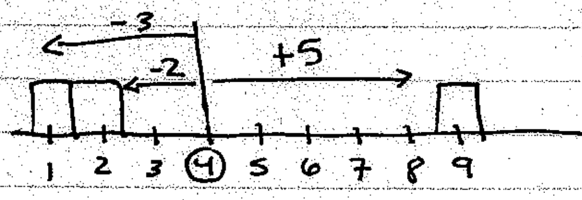
1 Lecture #4 Spread, Normal Curve

next time: experimental design





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



mean = 4

$$\begin{bmatrix}1\\2\\9\end{bmatrix} \xrightarrow{\text{subtract } \bar{y}} \begin{bmatrix}1-4=-3\\2-4=-2\\9-4=5\end{bmatrix} \leftarrow \text{deviation from the mean}$$

mean $\bar{y} = 4$

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

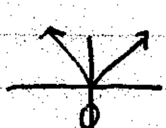
get rid of cancellation of ⊕ & ⊖ signs:

idea1: use absolute values

$$\begin{bmatrix}|y_1-\bar{y}|\\\vdots\\|y_n-\bar{y}|\end{bmatrix} = \begin{bmatrix}|-3|\\|-2|\\|+5|\\\$\end{bmatrix} \begin{matrix}=+3\\=+2\\=+5\\ \end{matrix}$$

mean absolute deviation (MAD)

meane $\frac{10}{3} = 3.3$ \$



MAD not used much.

Caitlin Daughty

AMS7
1/15/0

$\boxed{2}$ Lecture # 4

idea 2 : square the deviations

$$\begin{bmatrix} (y_1-\bar{y})^2 \\ (y_2-\bar{y})^2 \\ \vdots \\ (y_n-\bar{y})^2 \end{bmatrix} \$ = \begin{bmatrix} (-3)^2 \\ (-2)^2 \\ (5^2) \end{bmatrix} \$^2 = \begin{matrix} 9 \\ 4 \\ 25 \end{matrix}$$

$38 \div 3 = 12.7$ mean

idea so far

$$s = \sqrt{(y_1-\bar{y})^2 + (y_2-\bar{y})^2 + \ldots + (y_n+\bar{y})^2 / n-1}$$

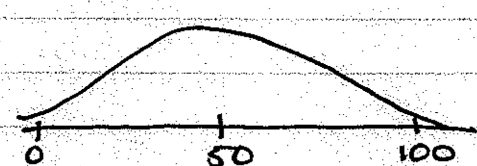
↳ standard deviation (SD) ← sample

$$s^2 = \frac{(y_1-\bar{y})^2 + \ldots + (y_n-\bar{y})^2}{n-1} \quad \text{sample variance}$$

mean 4 $\begin{bmatrix} 1 \\ 2 \\ 9 \end{bmatrix}$ mean 4 $\begin{bmatrix} \checkmark \\ \checkmark \\ \times \end{bmatrix}$

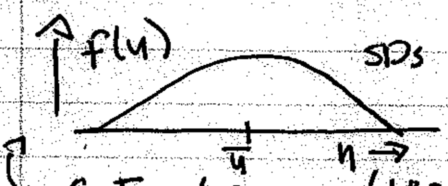
A data set with n observations has only (n-1) degrees of freedom for measuring spread.

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i-\bar{y})^2}$$



SD about 20

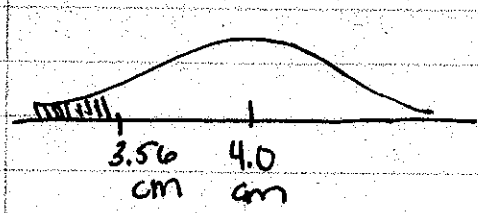
each # in this data set is around 50, give or take around 20



C.F. Gauss (1800) Gaussian Dist

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

density scale



SD .29 cm

← Butterfly Wing Length

Caitlin Doughty

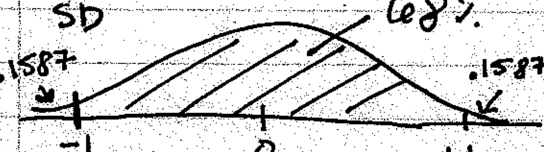
AMS 7
1/15/09

3 Lecture #4

84%

SD

68%



.1587 .1587 standard normal curve

-1 0 +1

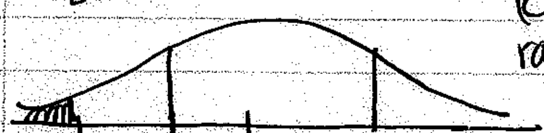
$\longrightarrow z$

$$\begin{bmatrix} 5 \\ 5 \\ \vdots \\ 5 \end{bmatrix} \quad \text{mean: } 5, \ SD: 0 \qquad \begin{bmatrix} c \\ \vdots \\ c \end{bmatrix} \quad \text{m... SD...}$$

- Facts about normal curve
  1. symmetric
  2. total area underneath it is 1 or 100%.

SD .29 cm

(cm) raw units



3.56 cm 4.0 cm 4.29 cm

3.71 cm

.0643 from table
= 6.4%

actual answer: 8.3%
normal approximation here:
good not great

-1 0 +1

standard units (su)

$$\frac{3.56-4.0}{0.29} \leftarrow \text{converting to standard units}$$

$$= \frac{-.44\ \text{cm}}{0.29\ \text{cm}} = -1.52$$

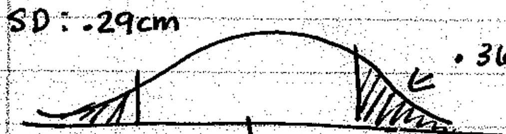
to convert to su:

$$su = \frac{\# - \text{mean}}{SD}$$

$$z = \frac{y-\bar{y}}{\sigma}$$

ex: what % of data is above 4.1 cm?

SD: .29 cm



.3669 4.0 cm 4.1 cm

.3669 = 36.7%
= 37%

↳ right answer $\frac{9}{24} = 38\%$.

$$\frac{4.1\ \text{cm} - 4.0\ \text{cm}}{.29\ \text{cm}} = \frac{.1}{.29} = +.34$$

normal curve does pretty well

-.34 +.34