

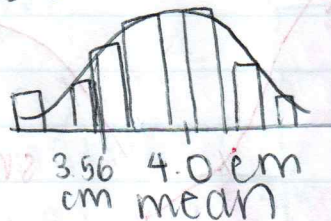
10/11/18

Wave Moretto

Stats Lecture #5: Normal Curve + Experimental Design

• HW #1 due by 11:59 pm next wed 10/17/18

SD: 0.29 cm

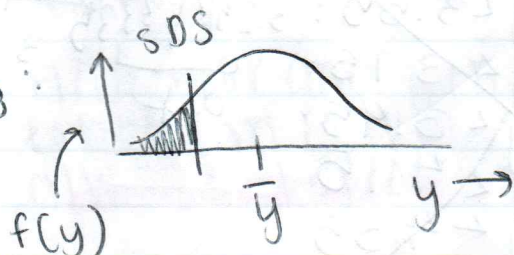


• Question: what % of butterflies in data set had wing length ≤ 3.56 cm?

• Answer₁: $2/24 = 8.3\%$
(exact)

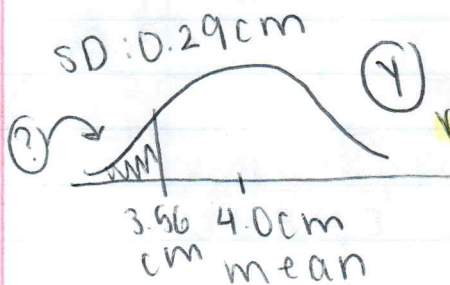
• Answer₂: area of relevant hist. bars = 8.3 %
(exact)

• Answer₃:
(approx)



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

• ~~Analytic~~ \rightarrow numerical integration

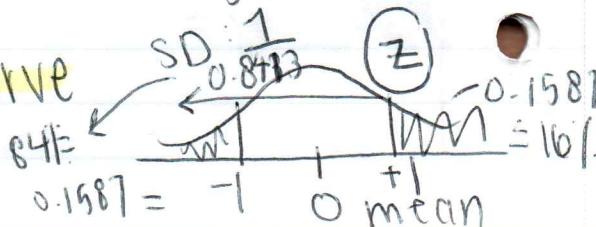


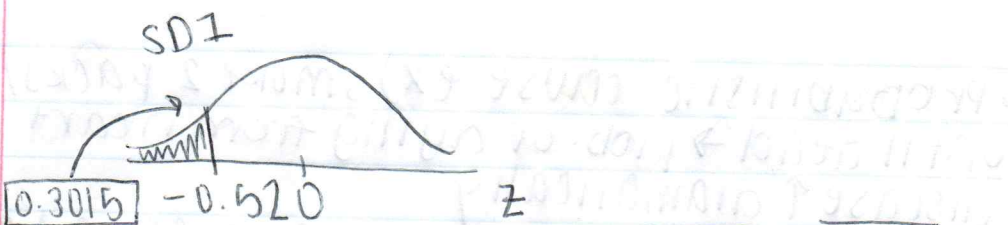
raw scale

$\begin{bmatrix} 42 \\ 42 \\ \vdots \\ 42 \end{bmatrix}$	any # C	$\begin{bmatrix} C \\ \vdots \\ C \end{bmatrix}$
mean 42		mean C
SD 0		SD 0

• Fact: all normal curves satisfy the empirical rule exactly

• Standard normal curve





0.3015 → 30.15% = 0.3 = 30%

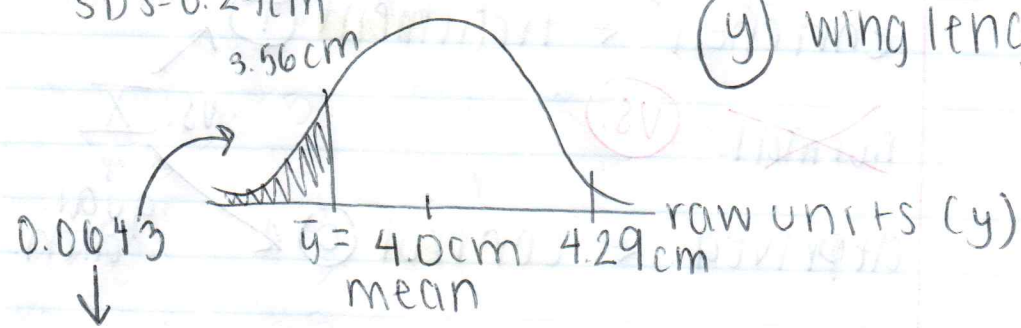
2 basic facts about normal curves:

- ① They're symmetric
- ② Total area = 1 = 100%

* this is the normal approximation to a hist.

SDs = 0.29 cm

(y) wing length



0.0643
↓
= 6.43%

$$\frac{3.56 \text{ cm} - 4.0 \text{ cm}}{0.29 \text{ cm}} = \frac{y - \bar{y}}{s} = z$$

$$= \frac{-0.44}{0.29} = -1.52 = \frac{\# - \text{mean}}{\text{SD}}$$

standard units (z)
↑ pure # w/o units
← converting to standard units (su)

$y = \bar{y} + zS$ (have z and want y)

LN pp. 36 → 66 additional notes (document camera notes)

Experimental Design

- most valuable form of causality

• CAUSE → effect
if (let hammer go) (it falls) always,

deterministic causality

• Probabilistic cause (ex) smoke 2 packs/day until dead \rightarrow prob. of dying from heart disease \uparrow dramatically

• Def: 2 variables X + Y are associated if when one goes \uparrow , the other tends to go \uparrow or \downarrow on average

- this is an association between X and Y
- association \neq causation

enriched \leftarrow treatment (T)

~~normal~~ (vs.)

deprived \leftarrow control (C)

