

# MATH:1260 Pokémath

## The Mathematics of Pokémon Go<sup>®</sup>

Week 10 Wednesday, Spring 24

Popular curve:

Slugma-like curve



---

## Plan for Today

- Module 2: Gotta Catch 'em All®
  - Statistics and Module 2 Wrap-up

---

## Class Reminders

- Project 2 work day on Thursday.
  - Earn up to 3 points extra credit.
- Project 2 Stage 2: Data due tonight at Midnight.
- Project 2 Stage 3: Report due Wednesday April 3 at Midnight.

---

# Statistics

Statistics is the study of data.

Two main types: Descriptive and Inferential.

**Descriptive:** Organize, Display, Describe a data set.

**Inferential:** Draw conclusions about an underlying population using a data set.

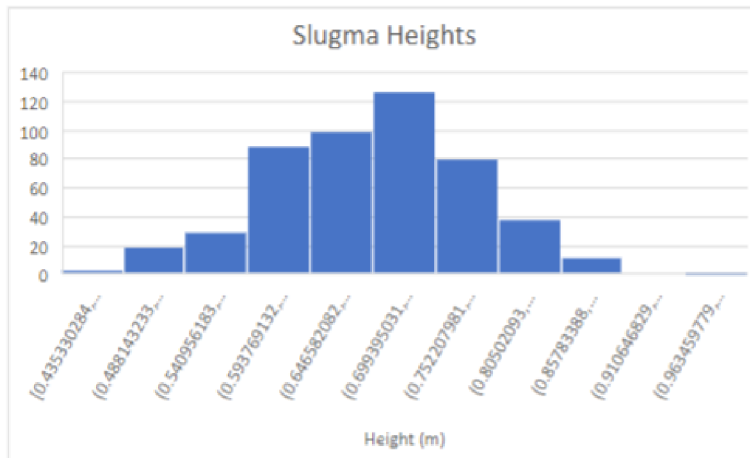
Recall data we recorded for Pikipek for GW.

We could compute the mean of the number of balls to catch. That would be descriptive.

We could use the data to estimate the underlying catch probability. That would be inferential.

## Slugma

Here is a data set of the heights of many, many Slugma, collected by TAs over the years.



Mean for this data set of 500 Slugma is 0.699m.

What do you notice about the shape of the curve? The median and mode?

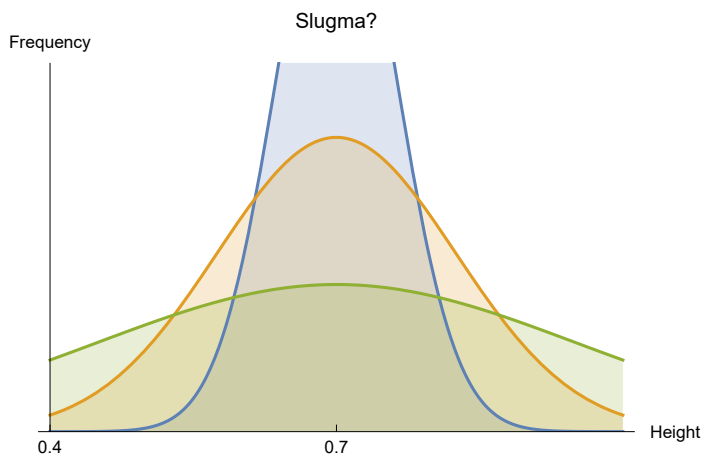
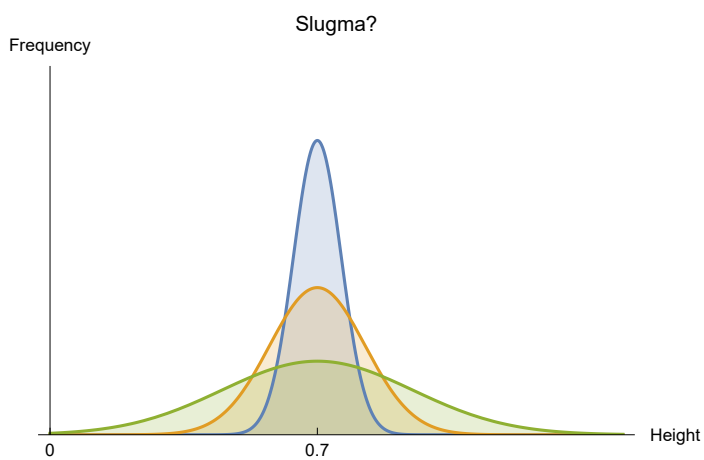
## Normal Distribution

The height of each Pokemon is chosen from a Normal Distribution.

The means are given in this chart by species

<https://pokemondb.net/pokedex/stats/height-weight>

	218	Slugma	<div>FIRE</div>	2'04"	0.7	77.2	35.0	71.4
---	-----	--------	-----------------	-------	-----	------	------	------



The normal curve is symmetric about the \_\_\_\_\_

## A more manageable set of Slugma

Slugma	Height (m)
1	0.66
2	0.69
3	0.78
4	0.63
5	0.66

TopHat: What is the average height of these Slugma?

## We can measure the “spread” of the data

We use “deviation from mean”. How did I calculate these?

Slugma	Height (m)	Deviation from mean
1	0.66	-.024
2	0.69	.006
3	0.78	.096
4	0.63	-.054
5	0.66	

## Negatives can be hard to work with, so we square the deviation

This makes the “distance from the mean” always positive

Slugma	Height (m)	Deviation from Mean	Square of Deviation
1	0.66	-.024	.000576
2	0.69	.006	.000036
3	0.78	.096	.009216
4	0.63	-.054	.002916
5	0.66	-.024	.000576

## Variance

**Population Variance** is the average of the square of deviation.

Notation  $\sigma^2$



---

## Full Formulas

### Population Variance

of data  $x_1, x_2, x_3 \dots x_n$  with mean  $\bar{x}$  is

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

---

## Formula for Standard Deviation

Population Standard Deviation (STDEV in excel)

of data  $x_1, x_2, x_3 \dots x_n$  with mean  $\bar{x}$  is

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}}$$

Notice how we square all the deviations from the mean, and then take the huge square root at the end. Do these cancel out?

## Formula for Correlation

Correlation (CORREL in excel)

of data  $x_1, x_2, x_3 \dots x_n$  with mean  $\bar{x}$  and data  $y_1, y_2, y_3 \dots y_n$  with mean  $\bar{y}$

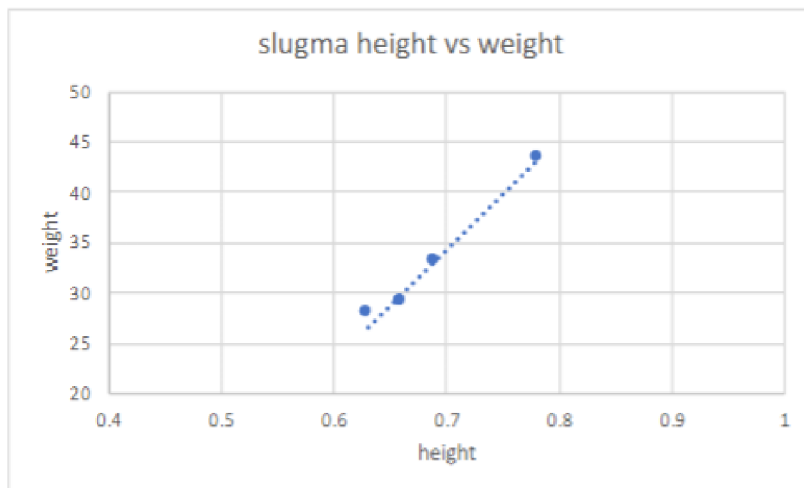
$$r = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + (x_3 - \bar{x})(y_3 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})}{n \sigma_x \sigma_y}$$

## Formula for Correlation

Now let's compute a correlation for height ( $h_1, h_2, \dots, h_5$  with mean  $\bar{h}$  and stdev  $s_h$ ) and weight ( $w_1, w_2, \dots, w_5$  with mean  $\bar{w}$  and stdev  $s_w$ )

$$r = \frac{(h_1 - \bar{h})(w_1 - \bar{w}) + (h_2 - \bar{h})(w_2 - \bar{w}) + (h_3 - \bar{h})(w_3 - \bar{w}) + (h_4 - \bar{h})(w_4 - \bar{w}) + (h_5 - \bar{h})(w_5 - \bar{w})}{(n-1) s_h s_w}$$

Slugma	Height (m)	Weight (kg)
1	0.66	29.71
2	0.69	33.12
3	0.78	43.42
4	0.63	28.10
5	0.66	29.16



$r = .991$

---

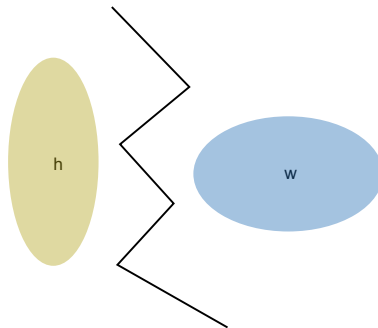
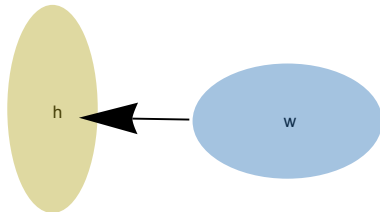
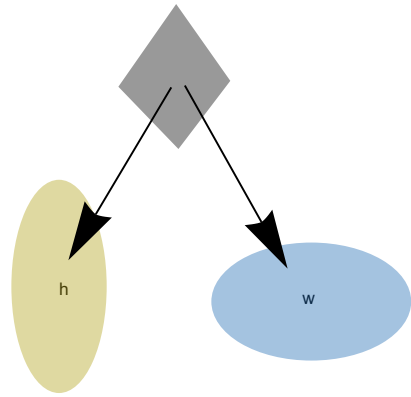
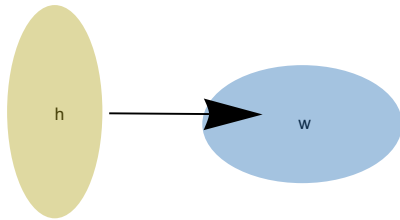
## Correlation is NOT Causation

“Permitting statistical treatment and the hypnotic presence of numbers and decimal points to befog causal relationships is little better than superstition. And it is often more seriously misleading.”

--Darrell Huff. *How to Lie with Statistics*.

Be careful!

Correlation can not distinguish between these cases...



Some funny examples of spurious correlations  
<http://tylervigen.com/spurious-correlations>