

One way which you can describe a variable is by detailing its values and the number of times those values occur — the distribution.

Histogram — represents a distribution, by describing the frequency of a variable.

Mode — is the most common value.

Normal/Gaussian Distrib. — bell shape distrib.

↳ Tail prejudice on either side of the distrib.

Uniform Distribution — all values occur with the same frequency.

### Outliers

Data which disproportionately falls outside an acceptable context.

### Histograms

- Make the most frequent values obvious.
- Not optimal for comparing distributions
  - ↳ does not account for differences in the sample size.
- a complete description of the distribution sample
  - ↳ can reconstruct the values in the sample (no order)
- details of a distrib. can be found in a hist.,
  - ↳ however it is oft better to use `desc.stats.`

- central tendency - vals. cluster around a point?
- mode - is there more than one cluster?
- spread - how much spread is there in the values?
- tails - how quickly do probabilities drop off as we move from the modes?
- outliers - are their extreme values far from the modes?

These type of statistics are called summary statistics.

mean - describe central tendency of distrl.  
mean calculation:

$$\bar{x} = \frac{1}{n} \sum_i x_i$$

average - differs in. one of several summary statistics that can be chosen to describe a central tendency

### Variance

In absence of a central tendency, can describe instead using-

- mean
- variance - describing the variability + spread

Squaring the Deviation of a distrl. The dispersion has these advantages,  $s^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$

ensures the deviations

can be used constructively

The squares remove negativity

known as the

deviation from the

and prevent cancellation from mean.  
competing terms.

[2] The exponent is appropriate deviation.  
for non-linear relationship  
of a variables value; where Standard Deviation is  $S^2$  -  
a deviation's scale is proportionate the variable.  
to its significance.

[3] Reconciling the squaring Variance is not a good  
is manageable → more so, than summary stat., but  
other styles of weighting and it is useful in some  
enforced constructionism. cases.

### Effect Size

Often the differences between groups doesn't

have much context; there isn't a popular scale which to root the delta, like University's letter system.

Instead by leveraging the amount of variation in both segments range of scores the diff. can be contextualized.

Formula, + semblance of a scale can be found.

Sample SD

Pooled =

$$\sqrt{\frac{(S_1^2 + S_2^2)}{2}}$$

- an evaluation of how much a variable is influenced between two groups.  
It is a summary statistic, calculating the size of effect.

A common choice is to calculate the difference between the means of two segments for a particular variable.

An alternative method is to use Cohen's Effect Size (Cohen's d).

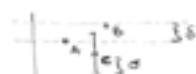
$d = 0.2 \rightarrow$  small

$d = 0.5 \rightarrow$  medium

$d = 0.8 \rightarrow$  large

$d = 1 \rightarrow (1) \text{ standard dev. in diff.}$

To compute the difference in effect relative to the typical dispersion.



\* The idea is to use or see the width of the distribution for change.

$$d = \frac{M_E - M_C}{\text{Sample SD Pooled}}$$

$$\times \left( \frac{N-3}{N-2.25} \right) \times \sqrt{\frac{N-2}{N}}$$

Connection factor for

$n_1 = \text{unsample}$

$n_2 = \text{vrsample}$

$n_1 + n_2 = \text{total sample}$

If the scores of one segment are entirely different from the other, thus

\* Can also use Hedge's g, which employs a less potent connection factor.

\* Sample SD Pooled can be replaced with the Standard Deviation of the population.

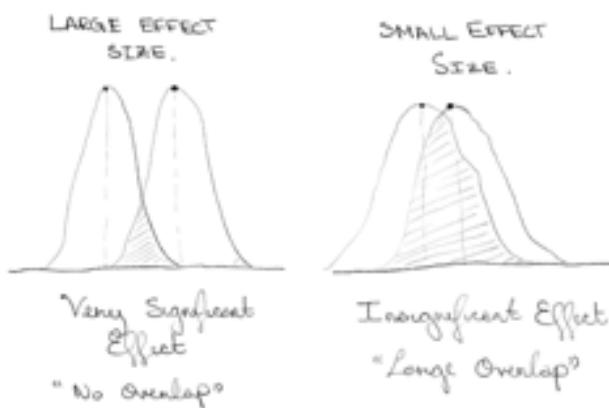
would suggest a high

$$\text{Z-score} = \frac{x_i - \bar{x}}{\sigma}$$

diff from the  
mean in units of  
( $\sigma$ ).

### I nterpretation of Effect Size - Cohen's d.

- Measures the difference in means in terms of the standard deviation.
- Effect size is equivalent to the z-score of a standard normal distribution.
- The effect size can be used to express the z-score, or the experiment's mean as a percentile relative to the control group. That percentile can be converted to rank-order describing the placement of the experiment mean amongst the control group.
- Can also convert to a probability of successfully guessing a datapoint's membership to the control or experiment group.



Effect can be gauged by assessing the variation in common—"the overlap", and using the commonality to contextualize the change in effect.

- ①  $\epsilon$ -based Binomial Effect Size Display (BESD)
  - 'understandability', and somewhat interpretable.
- ② Common Language Effect Size (CLES)
  - ↳ the probability of a score sampled from one distribution will be greater than the score sampled from another.

### Raw Cohen's "d" Value

- ↪ 0.2 → small  $\Delta$ height between 15 to 16 yr. old girls
- 0.5 → med.  $\Delta$ height 14 to 18 yr. old girls
- 0.8 → large  $\Delta$ height 13 to 18 yr. old girls.

### Relationship Between ES & Significance:

- ES lends significance to the amount of difference.
  - ↪ statistical significance - the likelihood that difference is <sup>not</sup> coincidental.
- effect size does not account for this.

### Margin for Error in estimating Effect Size:

- more data/larger sample  $\rightarrow$  more accuracy / stat. sig.
- can quantify margin of error w/ confidence interval:  
$$\text{margin of error} = \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$