Practical Significance

Practical significance refers to the magnitude of the difference, which is known as the **effect size**. Results are practically significant when the difference is large enough to be meaningful in real life. What is meaningful may be subjective and may depend on the context.

Cohen's d: For some tests there are commonly used measures of effect size. For example, when comparing the difference in two means we often compute Cohen's d

which is the difference between the two observed sample means in standard deviation units:

$$d = (\bar{x}_1 - \bar{x}_2)/s_p$$

Where: $s_p = \sqrt{((n_1-1)s_1^2 + (n_2-1)s_2^2)/(n_1 + n_2 - 2)}$

Cohen's d	**Interpretation**
0 - 0.2	Little or no effect
0.2 - 0.5	Small effect size
0.5 - 0.8	Medium effect size
0.8 or more	Large effect size

For a single mean, you can compute the difference between the observed mean and hypothesized mean in standard deviation units:

$$d = (\bar{x}_1 - \mu)/s$$

Clearly needs metric data

Common language effect size: the probability that a score sampled at random from one distribution will be greater than a score sampled from some other distribution. (McGraw and Wong, 1992)

Let

	Mean	SD
Group A	μ_A	s_A
Group B	μ_B	s_B

pnorm(0, $\mu_A - \mu_B$, $\sqrt{s_A^2 + s_B^2}$, lower.tail = FALSE) = x

- ⇨ x*100% chance that if you put an observation from Group A and one from Group B together at random, the one from Group A will be greater.

Clearly assumes Normal distributed, continuous metric data, but CLT will help out

A quicker *and* more exhaustive brute-force method for computing common-language effect sizes (Guillaume Rousselet).

Randomly sample one member of group A and one member of group B and see if the member of group A has a higher score. Count ties as half-wins (0.5) Do this 'run' number of times. Useful for non-normal noncontinuous data or data with at least a preference relationship defined

Equivalent approach useful with {0,1} data.

$$\text{CommonLanguageSizeEffect or CLES} = \frac{\text{(data in favor -data against)}}{\text{(total data)}} = \frac{\text{number of data values for and against}}{} = r$$

Correlation Methods

|r|>0.5, strong
0.2< |r|<=0.5,moderate
0<|r|<=0.2 weak

Pearson r correlation (McLeod, S. A. (2019, July 10). What does effect size tell you? Simply psychology: https://www.simplypsychology.org/effect-size.html)

This parameter of effect size summarises the strength of the bivariate relationship. The value of the effect size of Pearson r correlation varies between -1 (a perfect negative correlation) to +1 (a perfect positive correlation).

According to Cohen (1988, 1992), the effect size is low if the value of r varies around 0.1, medium if r varies around 0.3, and large if r varies more than 0.5.

Coefficient, r

Strength of Association	Positive	Negative
Small	0.1 to 0.3	-0.1 to -0.3
Medium	0.3 to 0.5	-0.3 to -0.5
Large	0.5 to 1.0	-0.5 to -1.0

Spearman rho

A correlation between two rank random variables by testing whether or not the population Spearman's rho = 0 (the null hypothesis).