Consulting Analytical Chemists and Geochemists

MAKING SENSE OF SIGNIFICANT DIFFERENCE USING COHEN'S *d*

Prepared by Allan Fraser

June 17th, 2020

Application Note: 34

1. INTRODUCTION

A *p*-value outcome in a statistical test of inference is a measure of the strength of evidence against a null hypothesis. Statistical significance validates that the findings of an experiment are not due to chance, but it does not provide a measure of how much (or how large) a difference there is. In statistical analysis of analytical results, we often want to know not just whether a result is significant, but how large the effect is. Effect size is simply the difference between the sample means being compared. However, effect size needs to be expressed as a standardised mean difference so that any comparisons are based on the same baseline comparison. Essentially, Cohen's *d* (Cohen,1988) effect size represents the degree to which two data sets (samples of a population) do not overlap. Cohen's *d* calculates a standardised mean difference between two groups (McLeod, 2019). The less that two samples overlap, the larger the effect size. Cohen developed a system of comparison between two groups as being either "small", "medium", or "large". Cohen's *d* is only calculated after rejecting a null hypothesis. If the null is not rejected, effect size has no meaning.

Both effect size and statistical significance are essential for a comprehensive understanding of the results of a statistical experiment and should be reported when quoting an experimental outcome.

There are two Cohen's *d* formulas, equation [1] gives the formula for groups with equal numbers of values and equation [3] is sued in unequal group sizes.

2. APPLICATION

Cohen's *d* is a standardised mean difference between two groups, where the mean of one group is subtracted from the other $(\bar{x}_A - \bar{x}_B)$ and divided by a pooled standard deviation, s_P [2], of both sets of data:

$$Cohen's\ d = \frac{(\bar{x}_A - \bar{x}_B)}{s_P} \qquad [1]$$

$$s_P = \sqrt{\frac{s_A^2 + s_B^2}{2}} \quad [2]$$

For unequal group sizes, Cohen's d_s is used:

$$Cohen's\ d_s = \frac{(\bar{x}_A - \bar{x}_B)}{s_P} \quad [3]$$

$$s_P = \sqrt{\frac{(n_A - 1)s_A^2 + (n_B - 1)s_B^2}{n_A + n_B - 2}} \quad [4]$$

Where, n_A and n_B are the number of measurements of data set A and B, respectively. s_A and s_A are the standard deviation of data set A and B, respectively.

Cohen's *d* and d_s, represent the number of standard deviations by which the two groups differ; with a *d* or d_s of 1 indicating that the two groups differ by 1 standard deviation, a *d* of 2, by 2 standard deviations, and so on. More specifically, Cohen states that if *d*=0.2 it can be considered a 'small' effect size, *d*=0.5 represents a 'medium' effect size and *d*=0.8 a 'large' effect size with ≥1.2, as very large effect size.

2.1 Worked Example

Two laboratories are sent the same ore sample for copper determination of by ICP-OES. A statistical comparison of the means using a two-sample unequal variance t-test gives a *p*-value of 0.002, making the mean value A and B significantly different an alpha of 0.05. Figure 1 presents the results of both labs as Gaussian Distributions and Table 1 the figures of merit for the test work. How large is the difference between the two means? To answer this, Cohen's *d* is calculated and equation [1] and [2] are used as there are an equal number of results for each lab, i.e., *n*=6 for Lab A and *n*=6 for Lab B. The pooled standard deviation is calculated followed by Cohen's *d*:

$$s_P = \sqrt{\frac{0.0701^2 + 0.0363^2}{2}} = 0.0559$$

Substituting into equation [1]:

$$Cohen's\ d = \frac{(\ 12.07 - \ 11.93)}{0.0559}$$

$$Cohen's\ d = \frac{(\ 0.14)}{0.0559} = 2.5$$

In summary, a t-test shows that there is a significant statistical difference between the mean values (*p*=0.002 at α of 0.05). In terms of Cohen's *d* convention, a 2.5 standard deviation demonstrates that the difference between the means is large.

Table 1. Figures of merit for two mean values for the assay of a copper ore sample for % copper by ICP-OES.

Lab A % Cu	12.07
n	6
Lab B % Cu	11.93
n	6
Lab A *s*	0.0701
Lab B *s*	0.0363
s_{pooled}	0.0559
p-value (α=0.05)	0.002
Cohen's *d*	2.5

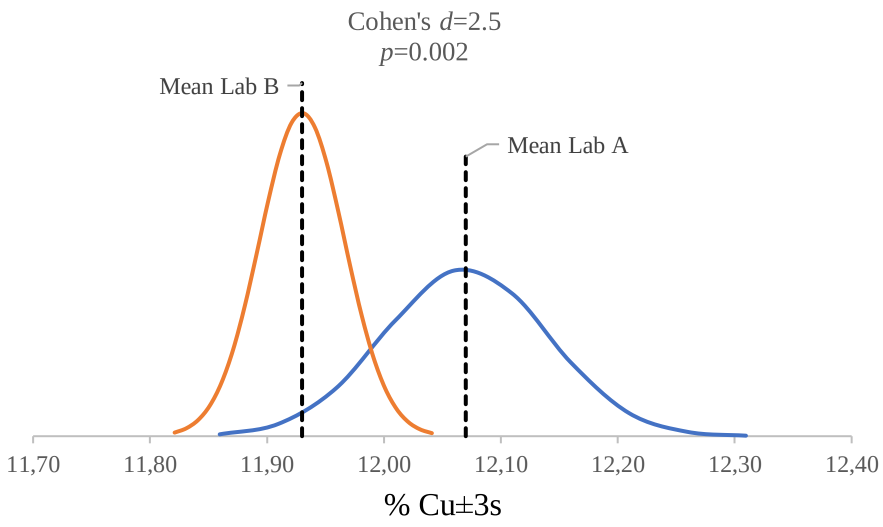

Figure 1. Means A and B are statistically significant (p=0.002 and <0.05), with a Cohen's *d* of 2.5. There are 2.5 standard deviations difference between the means making the effect size large according to Cohen's effect size convention.

3. Conclusion

In conclusion, Cohen's *d* values quantify the magnitude of the difference between mean values.

4. References

Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

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

How to cite this application note:

Fraser, A.W. (2020). Making sense of significant difference using Cohen's *d*. www.allanfraserandassociates.co.za. Applications note: 34. 4 pages.