Generalizing Effect Sizes for Differences with an Alternative to Cohen's d Coefficient

S Koby Taswell and Carl Taswell
ktaswell@bhavi.us, ctaswell@bhavi.us,
www.BrainHealthAlliance.org.

Introduction

Data obtained for population samples studied in clinical trials may not fit normal distributions. These non-normal distributions of samples create analytic and practical difficulties when estimating effect sizes for differences between sampled comparison groups. In these situations, it would be preferable to adopt a non-parametric model-free method to estimate an effect size for the difference between comparison groups in a manner that relies solely on computational analysis of the empirical data.

Methods

We generalize the classical effect size for differences as described by Cohen's d coefficient for the difference of the means divided by the standard deviation. Instead, we take an analogous approach for our estimate that we call the *Robust Generalized Effect Size* (RGES) by calculating the difference of the medians divided by the maximum of the peak half widths for the distributions of the two comparison groups. Thus, we use medians instead of means for the estimates of the measure of central tendency, and peak half widths instead of standard deviations for estimates of the measure of dispersion. Moreover, we define the peak half widths as a generalization of the traditional *Half Width at Half Maximum* (HWHM) of the main peak. Specifically, we propose use of the half width at half absolute height for measures of dispersion around the center of the data distribution. Our method has been defined with both mathematical formulas and MATLAB code with a function `rghwhah` for the *Robust Generalized Half Width Half Absolute Height*, and the function `rges` for the *Robust Generalized Effect Size*. The function `rghwhah` finds the width, height and center point of the main peak of the data distribution defined by the rectangular box with width determined by all data points with absolute values greater than half the height of the box. A peak half width obtained from `rghwhah` does not necessarily equal a peak half width obtained from a traditional HWHM. Then, the function `rges` computes the effect size comparing the difference between the two distributions each of which may have a different `rghwhah` width and height for their main peaks. Our method has been generalized sufficiently to analyse multi-modal and other non-normal distributions of data.

Results

Figure 1: RGHWHAH analyses with `rghwhah` boxes and box centers (in blue) of the main peaks (in green) for diverse functions: a normal distribution, a cosine function, a shifted cosine function, and a Daubechies Complex Orthogonal Most Asymmetric wavelet. Absolute value of the ordinates impacts the cosine waveform and DCOMA wavelet with both positive and negative values, but not the normal distribution or shifted cosine waveform with only positive values.

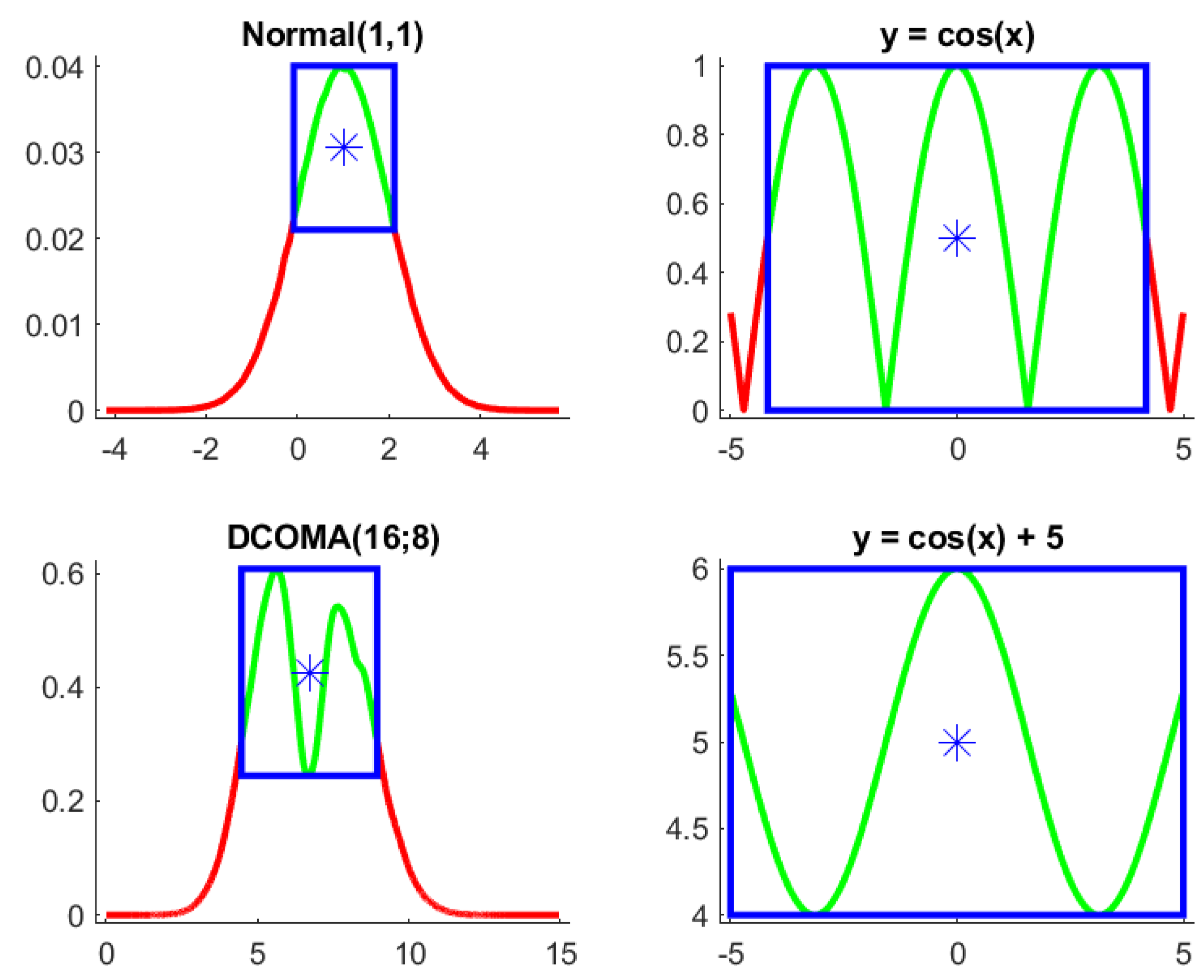


Figure 2: RGES analyses with estimated coefficient c effect sizes for pairs of `rghwhah` boxes and box centers comparing distribution Q(m,s) to P(m,s) with diverse examples: N(m,s) normal distribution, G(m,s) gamma distribution, L(m,s) logistic distribution, W(m,s) Weibull distribution.

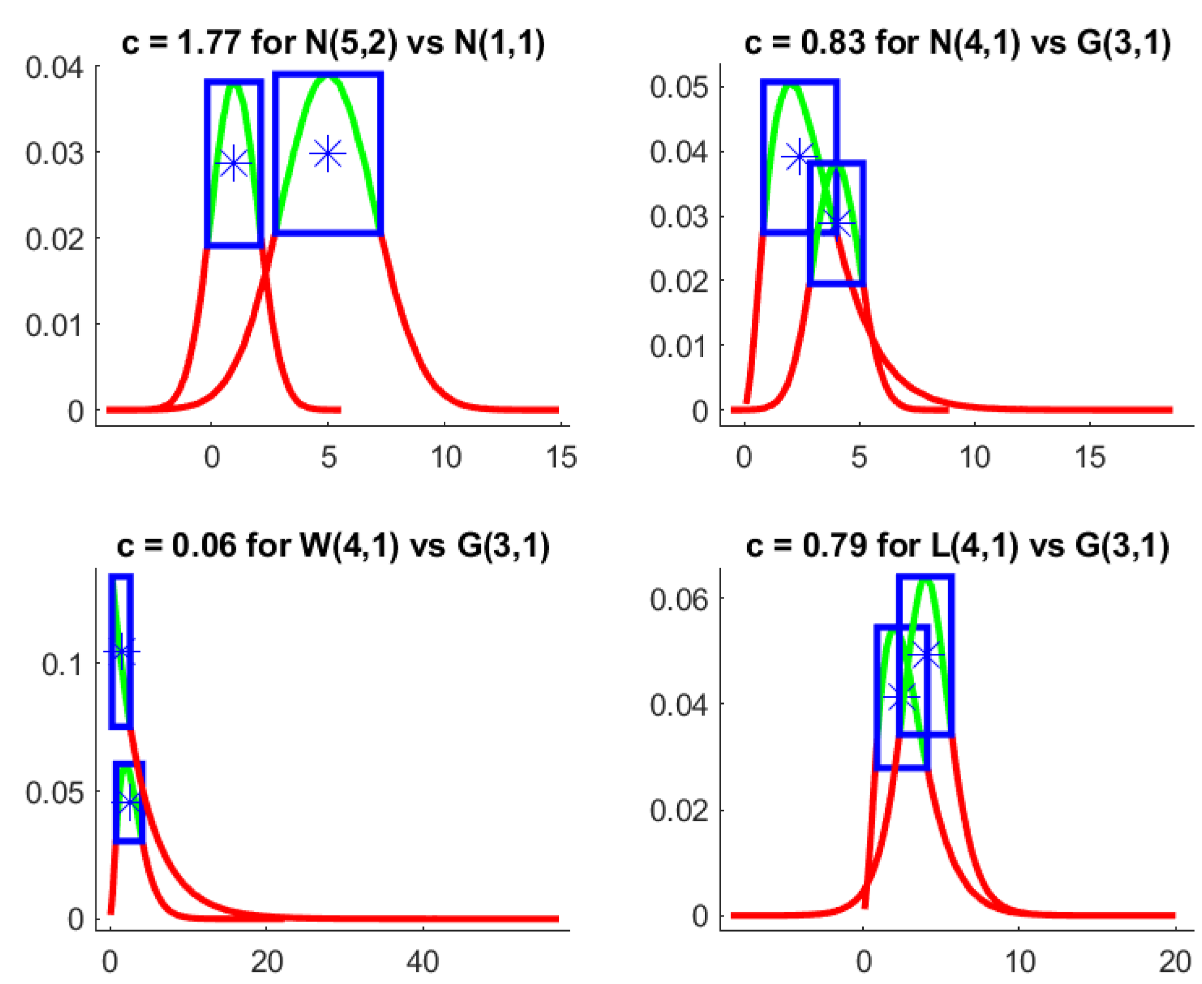


Figure 3: RGES analyses with estimated coefficient c effect sizes for pairs of `rghwhah` boxes and box centers comparing normal distribution Q(m,s) to normal distribution P(1,1).

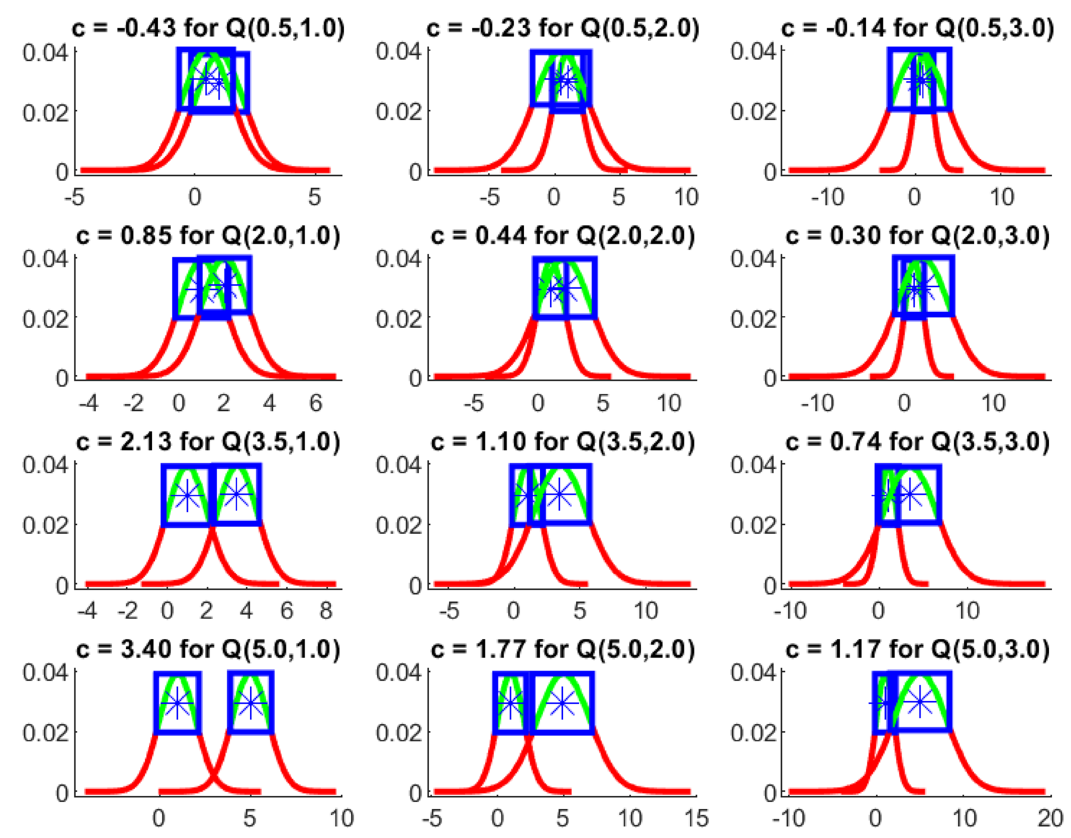


Figure 4: RGES analyses with estimated coefficient c effect sizes for pairs of normal distributions comparing Q(m,s) with varying parameters m and s to P(1,1) with fixed m = 1 and s = 1.

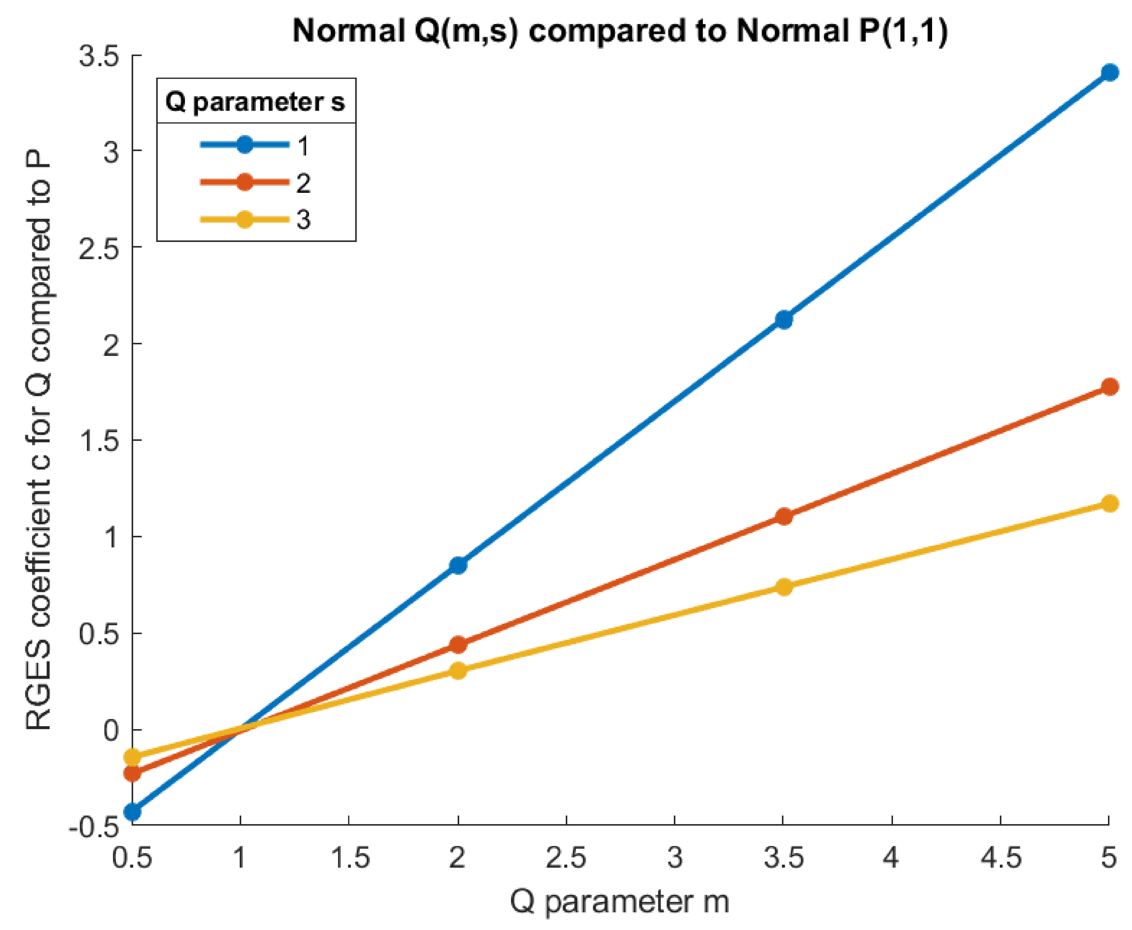
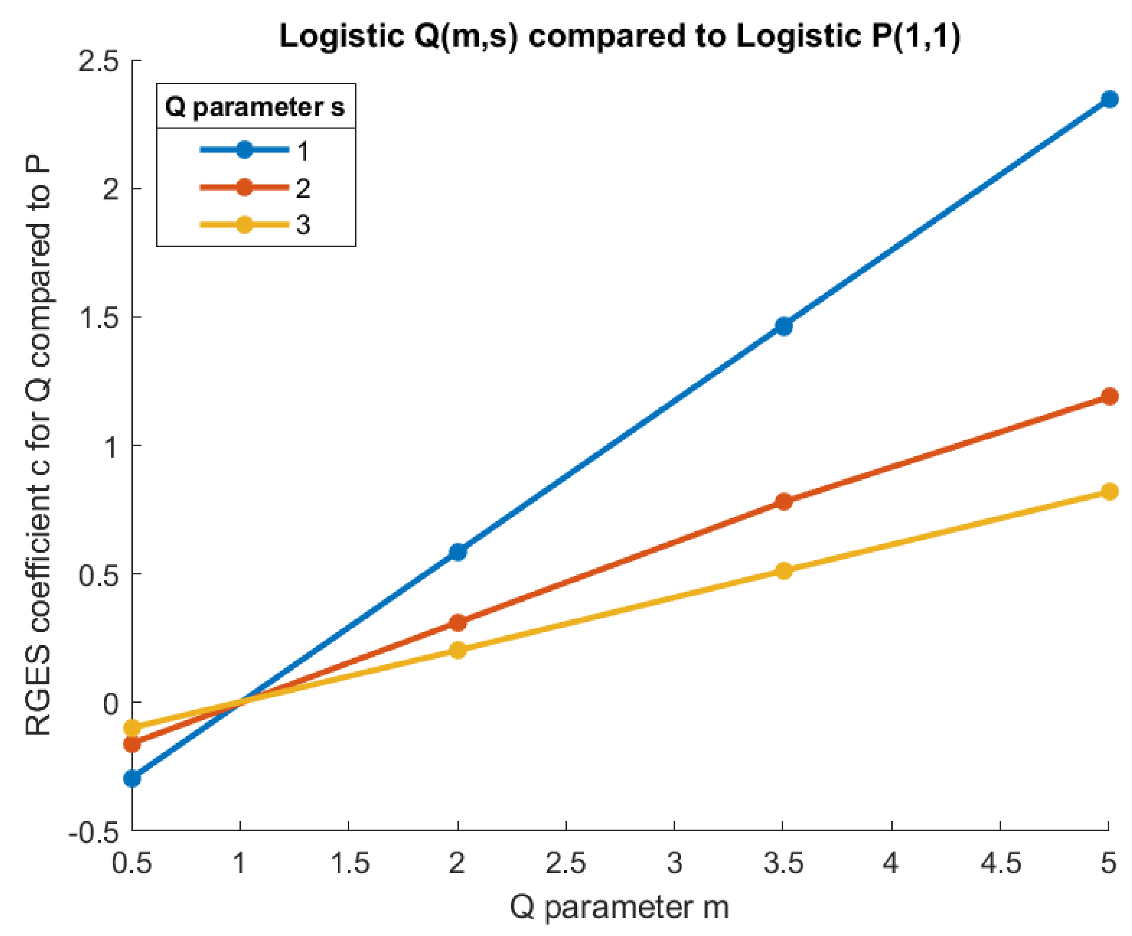


Figure 5: RGES analyses with estimated coefficient c effect sizes for pairs of logistic distributions comparing Q(m,s) with varying parameters m and s to P(1,1) with fixed m = 1 and s = 1.



Conclusion

We have tested our approach for the estimation of effect sizes with computer simulations comparing pairs of two different normal distributions with a wide range of varying means and standard deviations for the normal distributions. Results for our estimates of effect sizes have also been tabulated for a diverse variety of non-normal distributions compared with each other. These tabulated results of RGES coefficient c values will provide a reference for comparison of effect sizes for those statisticians and other clinical trial investigators who do not wish to rely solely on estimates for effect sizes based on standard deviations that might not be appropriate for non-normal distributions.

References

[1] Jacob Cohen, 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edition, Lawrence Erlbaum Associates.

[2] Wikipedia, 2019. *Effect size Cohen's d*. `https://en.wikipedia.org/wiki/Effect_size#Cohen's_d`.