

A Note on Cohen's Overlapping Proportions of Normal Distributions

James W. Grice, Ph.D.
Oklahoma State University

Paul T. Barrett, Ph.D.
University of Auckland

July 4th, 2011

In his popular book, *Statistical Power Analysis for the Behavioral Sciences* (1988, 2nd Ed.), Jacob Cohen does not correctly compute the percentage of overlap between two normal distributions when showing the relationship between this percentage and d (the standardized mean difference effect size). Specifically, on the bottom of page 21, Cohen writes "If we maintain the assumption that the populations being compared are normal and with equal variability, and conceive them further as equally numerous, it is possible to define measures of nonoverlap (U) associated with d which are intuitively compelling and meaningful." He then goes on to present examples to demonstrate the monotonic relationship between d and the percentage of nonoverlap between two normal distributions, reporting the results in Table 2.2.1 (p. 22). When $d = .1$, for instance, Cohen writes, " U_1 here equals 7.7%, that is, 7.7% of the area covered by both populations combined is not overlapped" (p. 21) and when $d = 2$, " U_1 then equals 81.1%, the amount of combined area not shared by the two population distributions" (p. 22). These descriptions, however, do not jibe with the meaning of the proportions as defined in Cohen's first statement. Consider when d equals 2, and refer to the top part of Figure 1 below. As can be seen, we assume both distributions are normal with standard deviations equal to unity. Their means are 0 and 2; thus, $d = [(2 - 0) / 1] = 2$. The "intersection point" is $(2 - 0) / 2 = 1$. The proportion of the alternative distribution that equals or exceeds 1 is .8413, and the proportion that equals or is less than 1 is .1587. These values can be found in any standard normal z -table, and they are shown in the bottom part of Figure 1. Incorporating the latter proportion (.1587) in the overlapped distributions in Figure 2, it can readily be seen that the proportion of nonoverlap (purple shaded area) for the alternative distribution is .6826. Because the two distributions are identical in shape and are "equally numerous", as Cohen states, the overall proportion of nonoverlap is also .6826; computed as $(.6826 + .6826) / 2 = .6826$, where 2 equals the total area of both distributions. This value does not match Cohen's .811 (81.1%), but it can be obtained from an online calculator: http://dmbru.net/Critical_Thinking/Normal_Intersection.html. The same value (.6826) can also be approximated from a chart provided by Linacre (1996; <http://www.rasch.org/rmt/rmt101r.htm>).

How did Cohen obtain 81.1%? The key is to realize that Cohen subtly shifted from thinking about frequency distributions to thinking of the distributions as two figures occupying a two-dimensional space. Note his first statement, "If we maintain the assumption that the populations being compared are normal and with equal variability, and conceive them further as equally numerous..." Clearly, he is speaking about frequencies of two populations of observations that are normally distributed. When speaking of actual U_1 values, however, he shifts to speaking about areas; for instance, " U_1 here equals 7.7%, that is, 7.7% of the *area covered by both populations combined* is not overlapped" (emphasis added). Focusing on area, as depicted in Figure 3, we can compute the total area as $.6826 + .3174 + .6826 = 1.6826$. The *combined nonoverlapping area* can then be computed as $(.6826 + .6826) / 1.6826 = .8114$, or 81.1%. All other values in Cohen's Table 2.2.1 can be computed in similar fashion. Consequently, they do not express the proportions of overall nonoverlap, although these are likely the values Cohen desired (as indicated by his initial frequency statement) because they are the values researchers typically want to know. We therefore recommend the reader of Cohen's text ignore the values in Table 2.2.1 and use the online calculator or the reasoning shown above instead.

The reader must also beware of statistics textbooks in which Cohen's proportions are reported, including when they are reported in their overlapping counterpart form. For instance, David Howell (*Statistical Methods for Psychology*, 6th Ed., 2007) reports Cohen's proportions (.85, .67, and .53) for three levels of d (.2, .5, .8, respectively) and erroneously describes them as "the degree to which the two distributions...overlap" (p. 218). The correct proportions of overlap are .92, .80, .69, respectively. As another example, consider this lengthier quote from Levine and Parkinson's (1994) *Experimental Methods in Psychology*,

“As was stated previously, when $d' = 0$ there is no detection. When $d' = 0$ there is no distance between the means of the signal-plus-noise and noise distributions. Another way of saying this is that the two distributions are 100% overlapped; there is 0% nonoverlap. As d' takes on values greater than zero, some portion of the area covered by both distributions combined will not be overlapped. The percentage nonoverlap provides us with a way of thinking about discriminability and values of d' . Consider a d' value of 0.50. The means of the signal-plus-noise and noise distributions are separated by 0.50 standard deviation. This means that a portion of the area covered by both distributions does not overlap. Cohen (1988, Table 2.2.1) listed the percentage nonoverlap between two normal distributions separated by distances from 0 to 4 standard deviation units. Referring to his table, we find that 33% of the combined area is nonoverlapped when $d' = 0.50$. That is, 33% of the area covered by the signal-plus-noise and noise distributions combined is either noise or signal plus noise, but not both.” (p. 232)

Here the authors use d' to refer to the standardized difference between means (d). They are also clearly thinking about frequencies of observations and interpreting Cohen's U_1 as the overall proportion of nonoverlap, but it does not represent this proportion. Instead, when d is equal to .50, the actual overall proportion of nonoverlap between two normal distributions with equal standard deviations is .20 (.1974), and the overall proportion of overlap is .80 (.8026). In Table 1 below we present a revised version of Cohen's table with the overall percent nonoverlap and overlap for different values of d . Examination of the values reveals that the general consequence of misinterpreting U_1 is an inflation of the magnitude of nonoverlap. Interpretations of effect sizes like those above which are based on U_1 will necessarily be inflated as well.

References

- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Ed). Hillsdale, NJ: Lawrence Erlbaum.
- Howell, D. C. (2007). *Statistical Methods for Psychology* (6th Ed). Belmont, CA: Thomson.
- Levine, G., & Parkinson, S. (1994). *Experimental Methods in Psychology*. Hillsdale, NJ: Lawrence Erlbaum.
- Linacre J.M. (1996) Overlapping Normal Distributions. *Rasch Measurement Transactions*, 10:1 p.487-8.

Figure 1. Overlapping null and alternative distributions, and alternative distribution with shaded proportions.

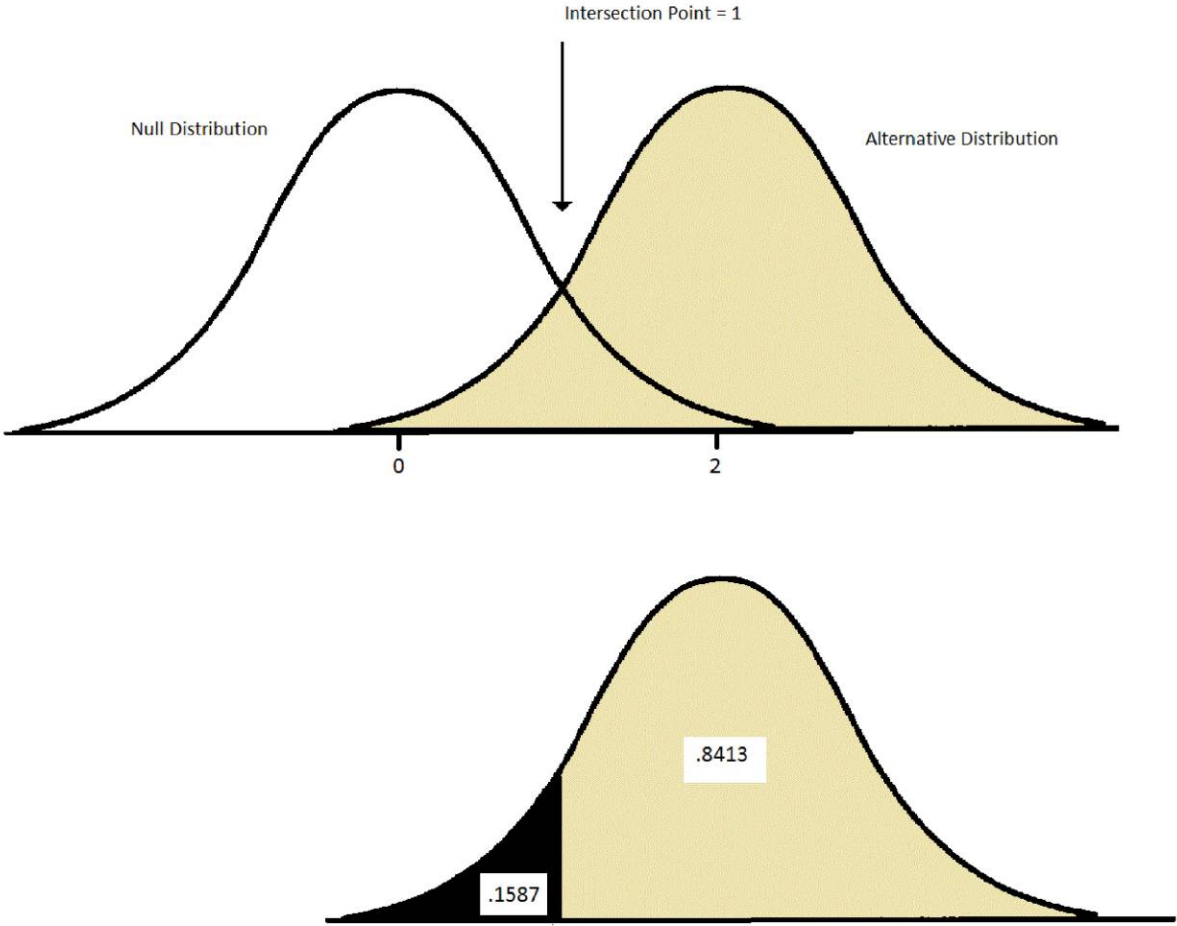


Figure 2. Overlapping distributions with shaded areas.

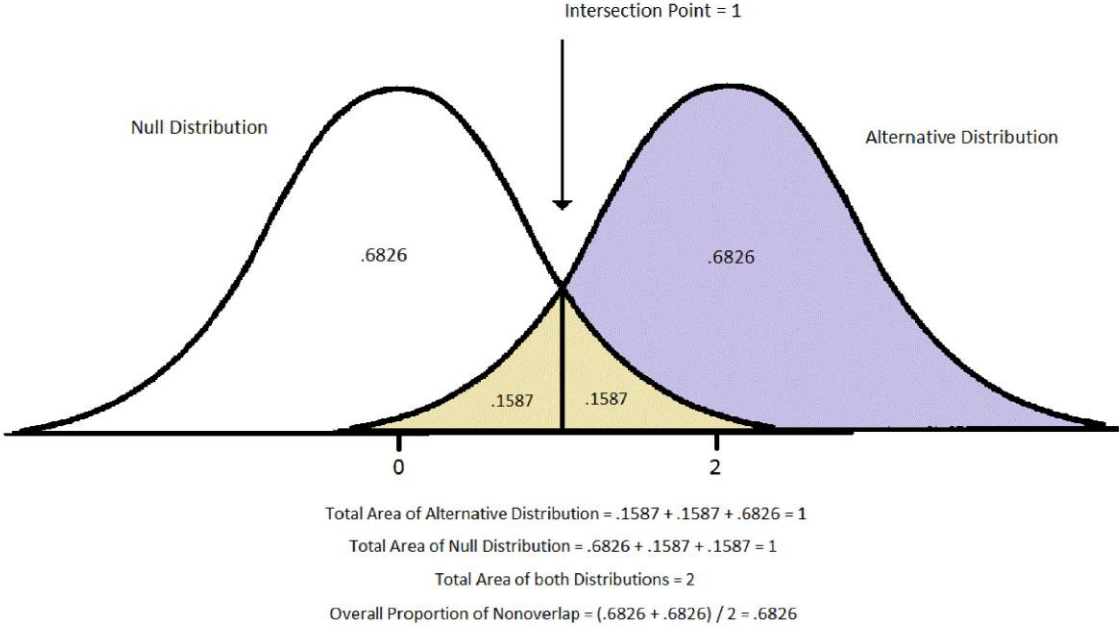


Figure 3. Overlapping distributions with the focus on combined area.

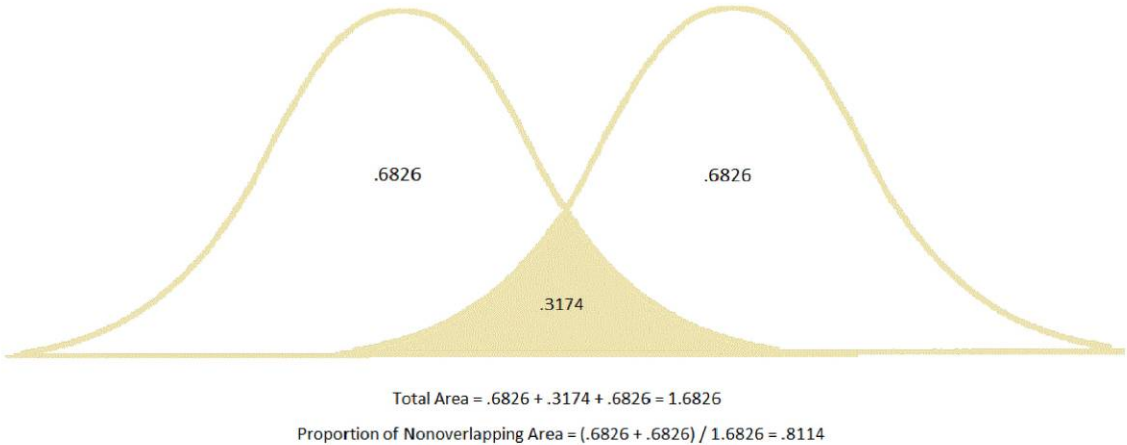


Table 1. Overall percent overlap for different values of d for two normal distributions with equal standard deviations.

| d | Cohen's U_1 | Overall Percent Nonoverlap | Overall Percent Overlap |
|-----|---------------|----------------------------|-------------------------|
| 0.0 | 0.0 | 0.00 | 100.00 |
| .1 | 7.7 | 3.99 | 96.01 |
| .2 | 14.7 | 7.97 | 92.03 |
| .3 | 21.3 | 11.92 | 88.08 |
| .4 | 27.4 | 15.85 | 84.15 |
| .5 | 33.0 | 19.74 | 80.26 |
| .6 | 38.2 | 23.58 | 76.42 |
| .7 | 43.0 | 27.37 | 72.63 |
| .8 | 47.4 | 31.08 | 68.92 |
| .9 | 51.6 | 34.73 | 65.27 |
| 1.0 | 55.4 | 38.29 | 61.71 |
| 1.1 | 58.9 | 41.77 | 58.23 |
| 1.2 | 62.2 | 45.15 | 54.85 |
| 1.3 | 65.3 | 48.43 | 51.57 |
| 1.4 | 68.1 | 51.61 | 48.39 |
| 1.5 | 70.7 | 54.67 | 45.33 |
| 1.6 | 73.1 | 57.63 | 42.37 |
| 1.7 | 75.4 | 60.47 | 39.53 |
| 1.8 | 77.4 | 63.19 | 36.81 |
| 1.9 | 79.4 | 65.79 | 34.21 |
| 2.0 | 81.1 | 68.27 | 31.73 |
| 2.2 | 84.3 | 72.87 | 27.13 |
| 2.4 | 87.0 | 76.99 | 23.01 |
| 2.6 | 89.3 | 80.64 | 19.36 |
| 2.8 | 91.2 | 83.85 | 16.15 |
| 3.0 | 92.8 | 86.64 | 13.36 |
| 3.2 | 94.2 | 89.04 | 10.96 |
| 3.4 | 95.3 | 91.09 | 8.91 |
| 3.6 | 96.3 | 92.81 | 7.19 |
| 3.8 | 97.0 | 94.26 | 5.74 |
| 4.0 | 97.7 | 95.45 | 4.55 |