Relation between Shieh's delta and Cohen's delta

Marie Delacre[1]

[1] Université Libre de Bruxelles, Service of Analysis of the Data (SAD), Bruxelles, Belgium

Author Note

Correspondence concerning this article should be addressed to Marie Delacre, CP191, avenue F.D. Roosevelt 50, 1050 Bruxelles. E-mail: marie.delacre@ulb.ac.be

Abstract

Keywords: keywords

Word count: X

Relation between Shieh's delta and Cohen's delta

Cohen's δ is the difference between both groups means, divided by a pooled error term:

$$\delta_{Cohen} = \frac{\mu_1 - \mu_2}{\sqrt{\frac{(n_1-1)\times\sigma_1^2+(n_2-1)\times\sigma_2^2}{n_1+n_2-2}}} \tag{1}$$

Shieh's δ is the difference between both groups means, divided by an unpooled error term (Shieh, 2013):

$$\delta_{Shieh} = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1/N} + \frac{\sigma_2^2}{n_2/N}}} \tag{2}$$

Unlike the classical Cohen's δ, Shieh's δ depends on the sample size ratio (i.e. $\frac{n_1}{n_2}$ that I will call later nratio). For the same amount of differences between two means, same standard deviations and σ-ratio, Shieh's δ will vary as a function of the nratio. Shieh's δ can therefore be expressed as a function of the nratio:

$$\delta_{Shieh} = \frac{(\mu_1 - \mu_2)\times\sqrt{nratio}}{(nratio+1)\times\hat{\sigma}}, \hat{\sigma} = \sqrt{(1-\frac{n_1}{N})\times\sigma_1^2 + (1-\frac{n_2}{N})\times\sigma_2^2}. \tag{3}$$

To illustrate the relation between Shieh's δ and the nratio, we can calculate the parameter across a range of nratio. We will first study the Shieh's δ (and its relation with Cohen's δ) when variances are equal between groups. We will then go through the relation when variances are unequal between groups.

When variances are equal between groups

As a first example, in Figure 1, Cohen's δ and Shieh's δ are calculated for different configurations where the observed mean difference $(\mu_1 - \mu_2)$ is 1, the total sample size is 200 and standard deviations σ_1 and σ_2 both equals 2[1].

First, when sample sizes are equal across groups, one can observe that Shieh's δ is half of the value of Cohen's δ. Shieh's δ equals 0.25 when nratio is 1, and Cohen's δ equals 0.50:

$$\delta_{Shieh,n_1=n_2} = \frac{\delta_{Cohen,n_1=n_2}}{2} \leftrightarrow \delta_{Shieh,n_1=n_2} = \frac{\mu_1 - \mu_2}{2 \times \hat{\sigma}} \tag{4}$$

Note that when variances are equal between groups, Cohen's δ is constant whatever groupe sample sizes are equal or not (i.e. Cohen's $\delta_{n1=n2}$ = Cohen's $\delta_{n1\neq n2}$). This equality will no longer be true when variances are unequal between groups. We use therefore $\delta_{Cohen,n_1=n_2}$ in Formula 4, in order that equality is still applicable when variances are unequal between groups (see later).

Moreover, when both sample sizes are equal between groups, Shieh's δ achieves its maximum value. When plotting both parameters against the log of the nratio, one can more easily observe that the Shieh's δ departs symmetrically from it's maximum value as long as the nratio moves away from 1 (i.e. when log(nratio) = 0; see Figure 2).

When variances are equal between groups, $\hat{\sigma}$ will be the same for all nratio ($\hat{\sigma} = \sigma_1 = \sigma_2$; i.e. $\hat{\sigma}$ in Figure 3 = $\hat{\sigma}$ in Figure 4), one can deduce that the relation between Shieh's δ value when nratio=1 and its value for all other nratio can be expressed as follows:

$$\delta_{Shieh,n_1=n_2} = \delta_{Shieh} \times \frac{\frac{\mu_1-\mu_2}{2\times\hat{\sigma}}}{\frac{(\mu_1-\mu_2)\times\sqrt{nratio}}{(nratio+1)\times\hat{\sigma}}} \leftrightarrow \delta_{Shieh,n_1=n_2} = \delta_{Shieh} \times \frac{nratio+1}{2 \times \sqrt{nratio}} \tag{5}$$

Note that because of formula 4 and because we know that Cohen's δ is constant for all

[1] Note that we chose a mean difference of 1 for convenience, however the mean difference does not impact the relation between both effect size measures

nratio, when variances are equal between groups, one can conclude that the relation between Cohen's δ and Shieh's δ can be expressed as following:

$$\delta_{Cohen} = \delta_{Shieh} \times \frac{nratio + 1}{\sqrt{(nratio)}} \tag{6}$$

This relations remains true as long as variances are equal between groups.

When variances are unequal between groups

In Figure 3, Cohen's δ and Shieh's δ are calculated for different configurations where the observed mean difference ($\mu_1 - \mu_2$) is 1, the total sample size is 200 and standard deviations σ_1 and σ_2 are respectively 4 and 3 (left) or 3 and 4 (right). As one can see, when variances are unequal between groups, Cohen's δ no longer remains constant for all nratio.

Once again, it is easier to study the influence of the nratio on both parameters when plotting them against the log of the nratio, as done in Figure 4. The way the standard error term is computed in Cohen's δ (see formula 1) implies that all samples are considered as issued from a common population variance (hence the assumption of homoscedasticity). When there is heteroscedasticity, if the larger variance is associated with the larger sample size (i.e. the colored parts on both plots in Figure 4), the error term is overestimated and therefore, the Cohen's δ is decreased. The smallest value is achieved when the sample size of the group associated with the largest variance equals n-1=199 (i.e. when one gives the largest weight to the largest standard deviation). On the other side, if the larger variance is associated with the smaller sample size (i.e. the non-colored parts of both plots), the error term is underestimated and therefore, the Cohen's δ is increased. The largest value is achieved when the sample size of the group associated with the largest variance equals 1 (i.e. when one gives the largest weight to the smallest standard deviation).

Unlike Cohen's δ, Shieh's δ is not influenced by the correlation between the sample size and the standard deviation.

While it remains true that when $n_1 = n_2$, the Cohen's δ is exactly as twice as large as Shieh's δ (Shieh's δ equals 0.25 and Cohen's δ equals 0.50), the maximum Shieh's δ value is no longer when the nratio equals 1 (i.e the log of the nratio equals 0). Moreover, Shieh's δ no longer departs symmetrically from it's maximum value as a function of the nratio. This is due to the fact that $\hat{\sigma}$ will vary a function of the nratio (and will therefore be different for all configurations presented in Figure 4): as shown in formula 3, one gives more weight to the standard deviation associated with the smallest group. For this reason, the maximum Shieh's δ is always achieved when there is a positive correlation between variances and sample sizes (i.e. we give more weight to the smallest standard deviation, associated with the smallest group) and the more unequal the variances, the further from 1 the nratio associated with the maximum , as illustrated in Figure 5.

As a consequence of different $\hat{\sigma}$ for all nratio's, the relation between Shieh's δ value when nratio=1 and its value for all other nratio cannot be as simplifed as it was when variances were equal:

$$\delta_{Shieh,n_1=n_2} = \delta_{Shieh} \times \frac{\frac{\mu_1-\mu_2}{2\times\sigma_{(n_1=n_2)}}}{\frac{(\mu_1-\mu_2)\times\sqrt{nratio}}{(nratio+1)\times\sigma_{(n_1\neq n_2)}}} \leftrightarrow \delta_{Shieh,n_1=n_2} = \delta_{Shieh} \times \frac{(nratio+1)\times\sigma_{n_1\neq n_2}}{2\times\sigma_{n_1=n_2}\times\sqrt{nratio}} \quad (7)$$

With

$$\sigma_{n_1=n_2} = \sqrt{\frac{\sigma_1^2+\sigma_2^2}{2}}$$

and

$$\sigma_{n_1\neq n_2} = \sqrt{(1-\frac{n_1}{N})\times\sigma_1^2 + (1-\frac{n_2}{N})\times\sigma_2^2}$$

Finally, because of formula 4, one can conclude that the relation between the Cohen's δ we would obtain if sample sizes were equal between groups and Shieh's δ can be expressed as following:

$$\delta_{Cohen} = \delta_{Shieh} \times \frac{(nratio + 1) \times \sigma_{n_1 \neq n_2}}{\sigma_{n_1 = n_2} \times \sqrt{nratio}} \tag{8}$$

Formula 8 gives us the general relation between Shieh's δ and Cohen's δ,applicable whatever variances are equal between groups or not.

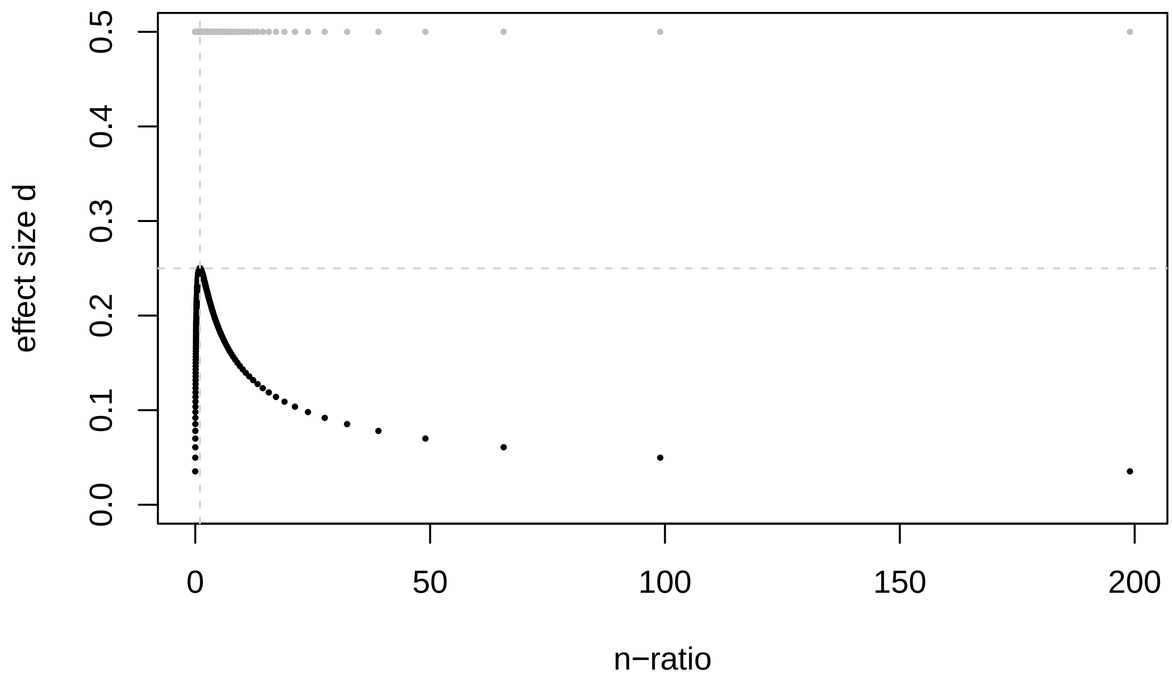

Figure 1. Comparison of Shieh's d (black dots) and Cohen's d (grey dots) when mu1 - mu2 = 1, N = 200 and sigma 1 and sigma 2 both equals 2

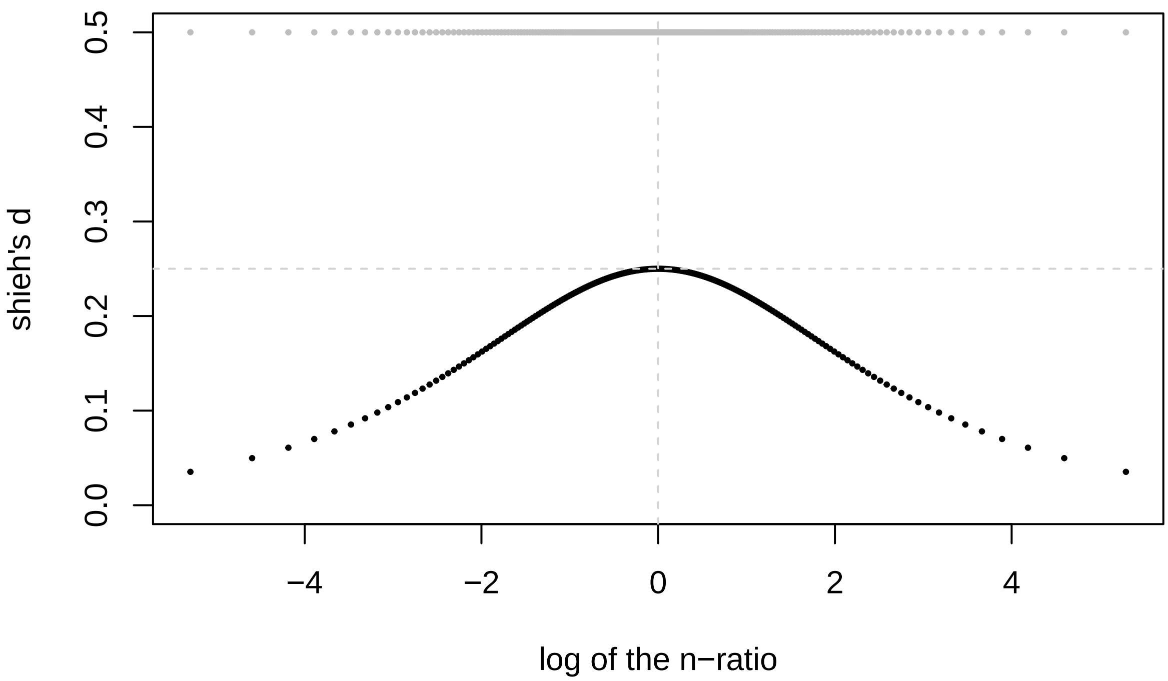

Figure 2. Comparison of Comparison of Shieh's d (black dots) and Cohen's d (grey dots) when mu1 - mu2 = 1, N = 200 and delta 1 = delta 2 = 2

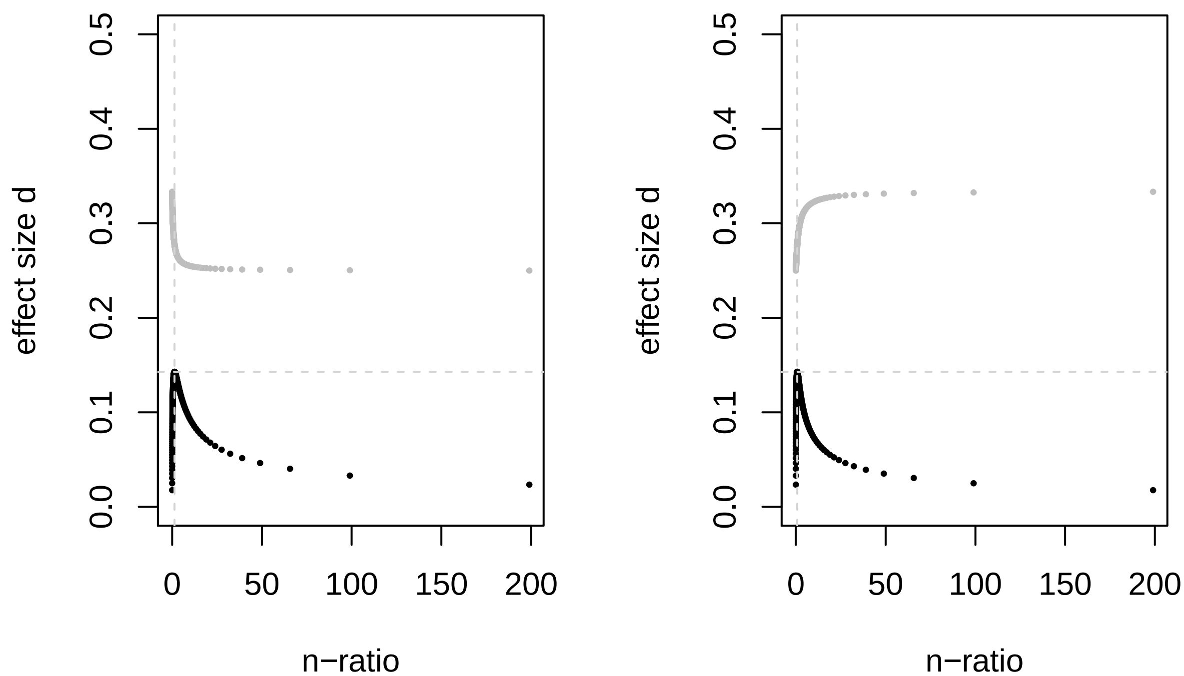

Figure 3. Comparison of Shieh's d (black dots) and Cohen's d (grey dots) when mu1 - mu2 = 1, N = 200 and sigma 1 and sigma 2 are respectively 4 and 2 (left) or 2 and 4 (right)

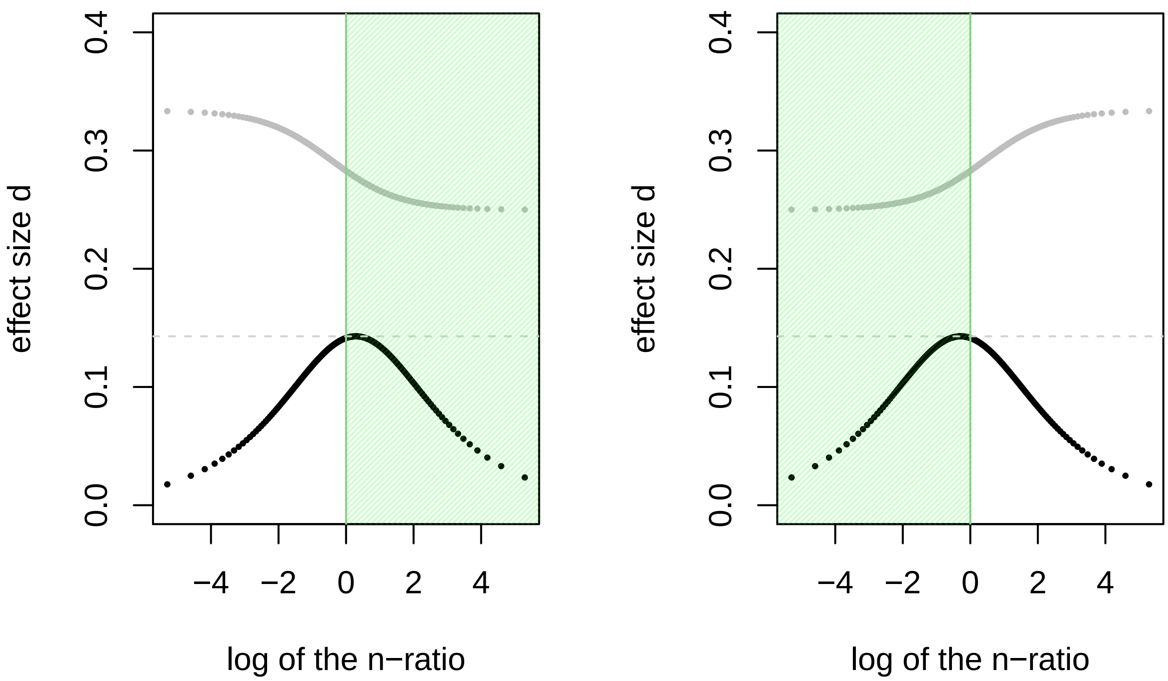

Figure 4. Comparison of Shieh's d (black dots) and Cohen's d (grey dots) when mu1 - mu2 = 1, N = 200 and sigma 1 and sigma 2 are respectively 4 and 2 (left) or 2 and 4 (right)

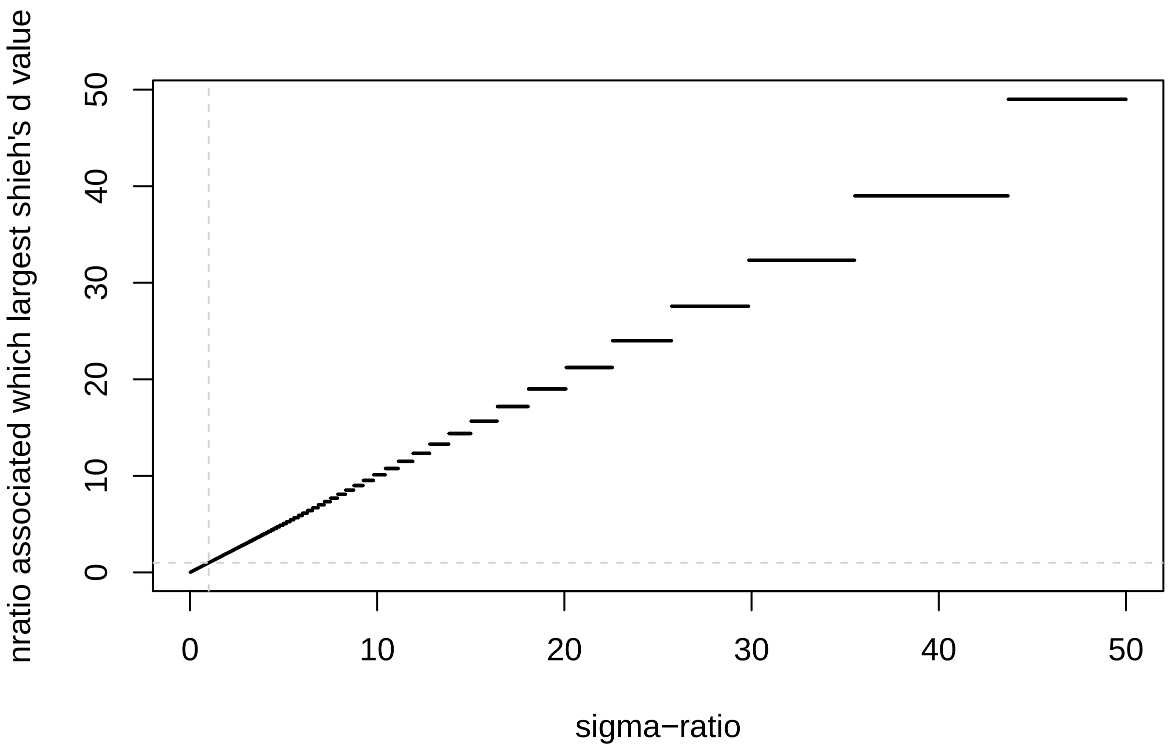

Figure 5. At which nratio the largest value of Shieh's d is achieved, as a function of the sigma-ratio (= sigma1/sigma2)

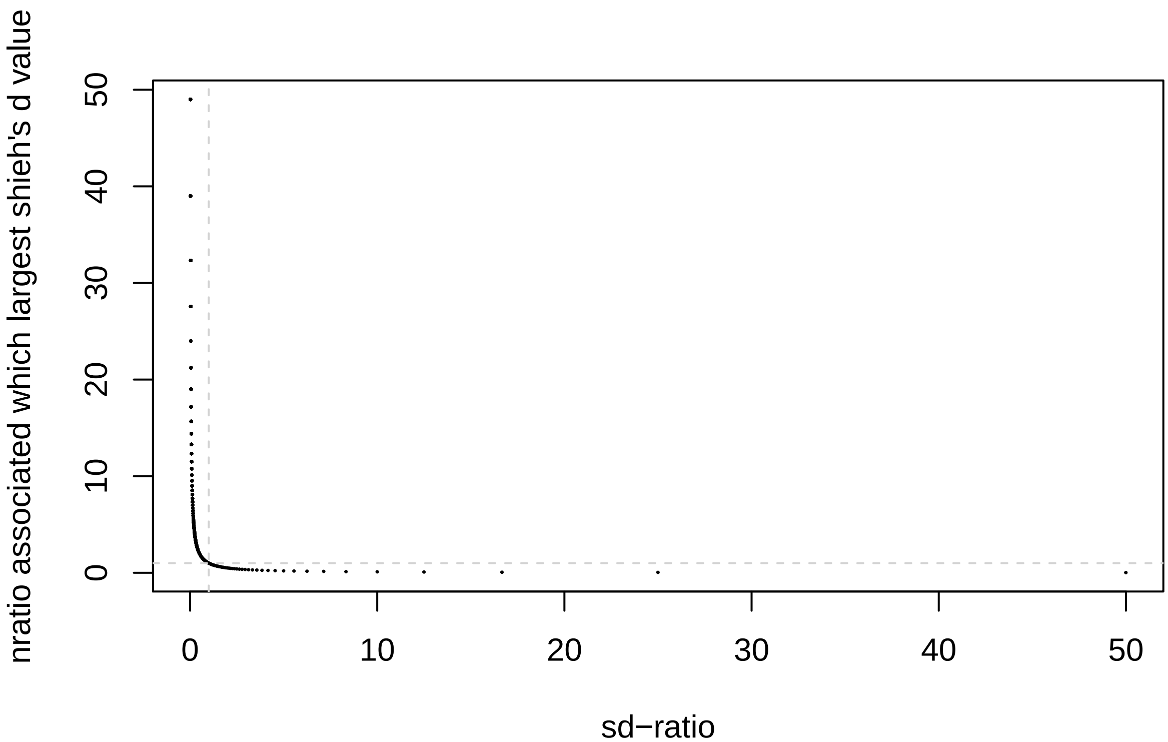

Figure 6. At which nratio the largest value of Shieh's d is achieved, as a function of the sigma-ratio (= sigma1/sigma2)