Reminder about Confidence Intervals

Marie Delacre[1]

[1] ULB

Author Note

Correspondence concerning this article should be addressed to Marie Delacre, Postal address. E-mail: marie.delacre@ulb.ac.be

Abstract

Keywords: keywords

Word count: X

Reminder about Confidence Intervals

Introduction: How to compute a confidence interval around $\mu_1 - \mu_2$. Considering the link between confidences intervals and NHST approach, we can think of confidence limits as the most extreme values of $\mu_1 - \mu_2$ that we could define as null hypothesis and that would not lead to rejecting the null hypothesis (Cumming & Finch, 2001), that is to say, the value associated with a p-value that exactly equals $\frac{alpha}{2}$.

Under the assumption of iid normal distribution of residuals with equal population variances across groups, in order to test the null hypothesis that $\mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$, we can compute the following quantity:

$$t_{Student} = \frac{(\bar{X_1} - \bar{X_2}) - (\mu_1 - \mu_2)_0}{SE} \tag{1}$$

With $SE = \sigma_{pooled} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ and $\sigma_{pooled} = \sqrt{\frac{(n_1-1)\times S_1^2 + (n_2-1)\times S_2^2}{n_1+n_2-2}}$.

Under the null hypothesis, this quantity will follow a central t- distribution with $n_1 + n_2 - 2$ degrees of freedom. [1]. The central t-distribution is always centered around $(\mu_1 - \mu_2)_0$ (e.g. when we expect that $\mu_1 = \mu_2$ under the null hypothesis, the central t-distribution is centered around 0, when we expect that $\mu_1 - \mu_2 = 3$ under the null hypothesis, the central t-distribution is centered around 3, etc.).

Considering all these information, we can therefore define $(\mu_1 - \mu_2)_L$, the lower limit of the confidence interval, such as $\frac{(\bar{X_1} - \bar{X_2}) - (\mu_1 - \mu_2)_L}{SE}$ exactly equals the quantile (1-$\frac{\alpha}{2}$) of the central t-distribution of the null hypothesis $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_L$ (i.e. the symmetric t-distribution that is centered around $(\mu_1 - \mu_2)_L$) and the upper limit $(\mu_1 - \mu_2)_U$ such as $\frac{(\bar{X_1} - \bar{X_2}) - (\mu_1 - \mu_2)_U}{SE}$ exactly equals the quantile $\frac{\alpha}{2}$ of the central t-distribution of the null

[1] Distribution is central because under the null hypothesis, the quantity is a (supposed normal) centered variable, divided by SE, an independant variable closely related with the χ^2.

hypothesis $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_U$ (i.e. the symmetric t-distribution that is centered around $(\mu_1 - \mu_2)_U$):

$$Pr[t_{n_1+n_2-2} \geq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_L}{SE}] = \frac{\alpha}{2} \tag{2}$$

$$Pr[t_{n_1+n_2-2} \leq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_U}{SE}] = \frac{\alpha}{2} \tag{3}$$

This is illustrated in Figure 1.

Under the assumption of iid normal distributions of residuals with unequal variances across groups, in order to test the null hypothesis that $\mu_1 - \mu_2 = (\mu_1 - \mu_2)_0$, we can compute the following quantity:

$$t_{Welch} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{SE} \tag{4}$$

With $SE = \sqrt{\frac{S_1^2}{n1} + \frac{S_2^2}{n2}}$. Again, under the null hypothesis, we know that this quantity will follow a central t- distribution with $DF = \frac{(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1 - 1} + \frac{(\frac{S_2^2}{n_2})^2}{n_2 - 1}}$ degrees of freedom (see Figure **??**). We can therefore easily define $(\mu_1 - \mu_2)_L$ such as $\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_L}{SE}$ exactly equals the quantile (1-$\frac{\alpha}{2}$) of the central t-distribution of the null hypothesis $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_L$, and the upper limit $(\mu_1 - \mu_2)_U$ such as $\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_U}{SE}$ exactly equals the quantile $\frac{\alpha}{2}$ of the central t-distribution of the null hypothesis $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_U$:

$$Pr[t_{DF} \geq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_L}{SE}] = \frac{\alpha}{2} \tag{5}$$

$$Pr[t_{DF} \leq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_U}{SE}] = \frac{\alpha}{2} \tag{6}$$

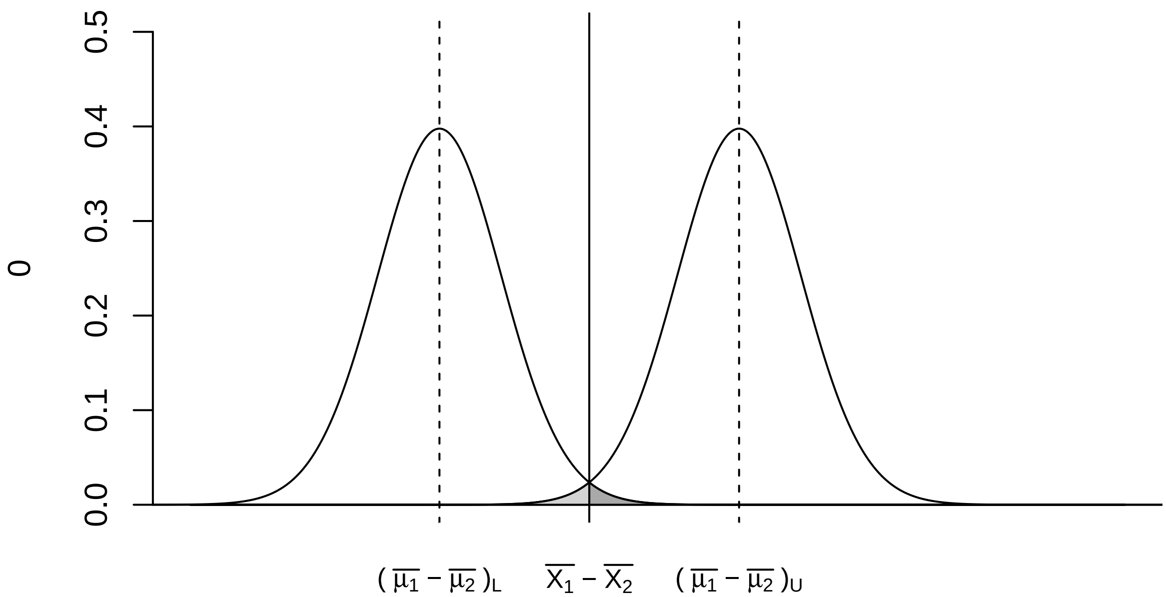

Figure 1. Sampling distribution. Both curves have the shape of a t distribution (df=n1+n2-2). Left curve is placed so that the proportion of estimates larger than the empirical mean difference (dark grey area) exactly equals alpha/2 and the right curve is placed so that the proportion of estimates lower than the empirical mean difference (light grey area) also exactly equals alpha/2.

It is not the most conventional way of computing confidences limits around any mean differences, however this approach is interesting as it helps to understand how to compute confidence limits around a measure of effect size.

How to compute a confidence interval around Cohen's δ. We previously mentioned that if the null hypothesis is true,$t_{Student}$ (see equation (1)) will follow a central t-distribution. However, if the null hypothesis is false, the distribution of this quantity will not be centered, and noncentral t-distribution will arise (Cumming & Finch, 2001), as

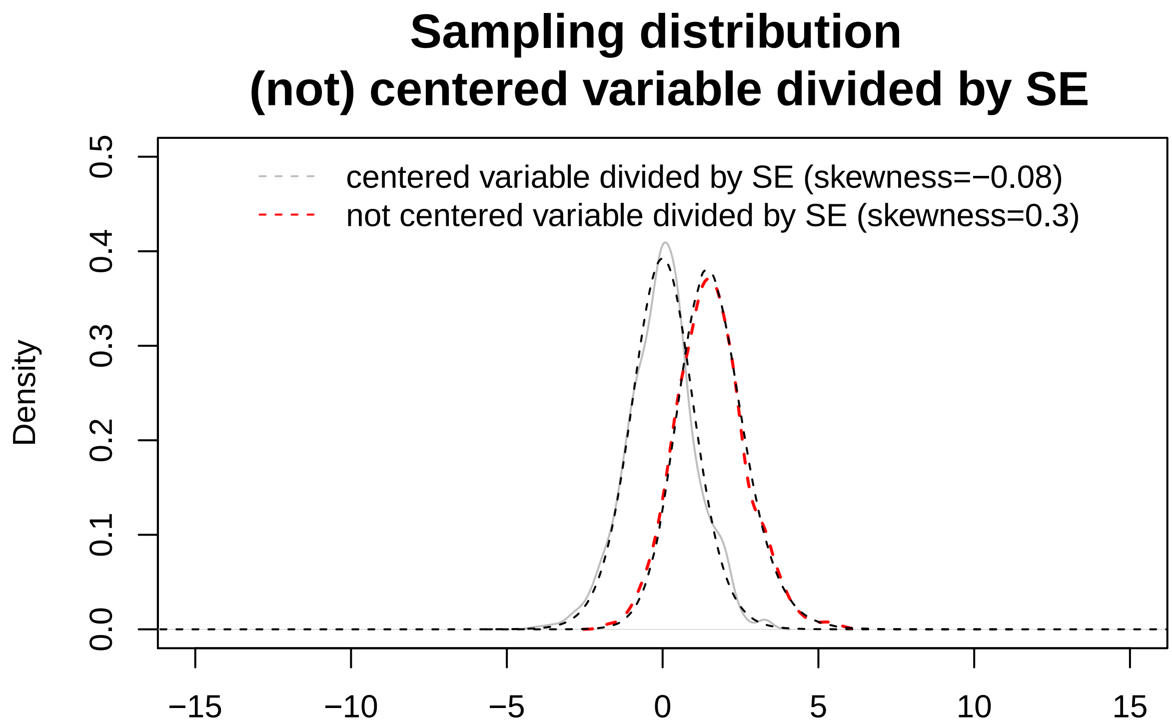

Figure 2. Sampling distribution of centered mean difference divided by SE (in grey, i.e. pivotal quantity) and not centered mean difference divided by SE (in red), assuming normality and homoscedasticity.

illustrated in Figure 2.

Noncentral t-distributions are described by two parameters: degrees of freedom (df) and noncentrality parameter (that we will call Δ; Cumming & Finch, 2001), the last being a function of δ and sample sizes n_1 and n_2:

$$\Delta = \frac{\mu_1 - \mu_2}{\sigma_{pooled}} \times \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}} \tag{7}$$

Considering the link between Δ and δ, it is possible to compute confidence limits for Δ, and divide them by $\sqrt{\frac{n_1 \times n_2}{n_1+n_2}}$ in order to have confidence limits for δ. In other word, we first need to determine the noncentrality parameters of the t-distributions for which $t_{Student}$ corresponds respectively to the $1 - \frac{\alpha}{2}$ and to the $\frac{\alpha}{2}$ th. quantile:

$$P[t_{df,\Delta_L} \geq t_{Student}] = \frac{\alpha}{2}$$

$$P[t_{df,\Delta_U} \leq t_{Student}] = \frac{\alpha}{2}$$

With $df = n_1 + n_2 - 2$. Second, we divide Δ_L and Δ_U by $\sqrt{\frac{n_1 \times n_2}{n_1 + n_2}}$ in order to define δ_L and δ_U:

$$\delta_L = \frac{\Delta_L}{\sqrt{\frac{n_1 \times n_2}{n_1 + n_2}}}$$

$$\delta_U = \frac{\Delta_U}{\sqrt{\frac{n_1 \times n_2}{n_1 + n_2}}}$$

How to determine the confidence interval around Shieh's $\delta*$

Like $t_{Student}$, t_{Welch} (see equation (4)) will follow a central t-distribution only if the null hypothesis is true. If the null hypothesis is false, it will follow a noncentral t-distribution, as illustrated in Figure 3.

The noncentrality parameter $\Delta*$ is a function of $\delta*$ and total sample size $N = n_1 + n_2$ (Shieh, 2013)

$$\Delta* = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1/N} + \frac{\sigma_2^2}{n_2/N}}} \times \sqrt{N} \qquad (8)$$

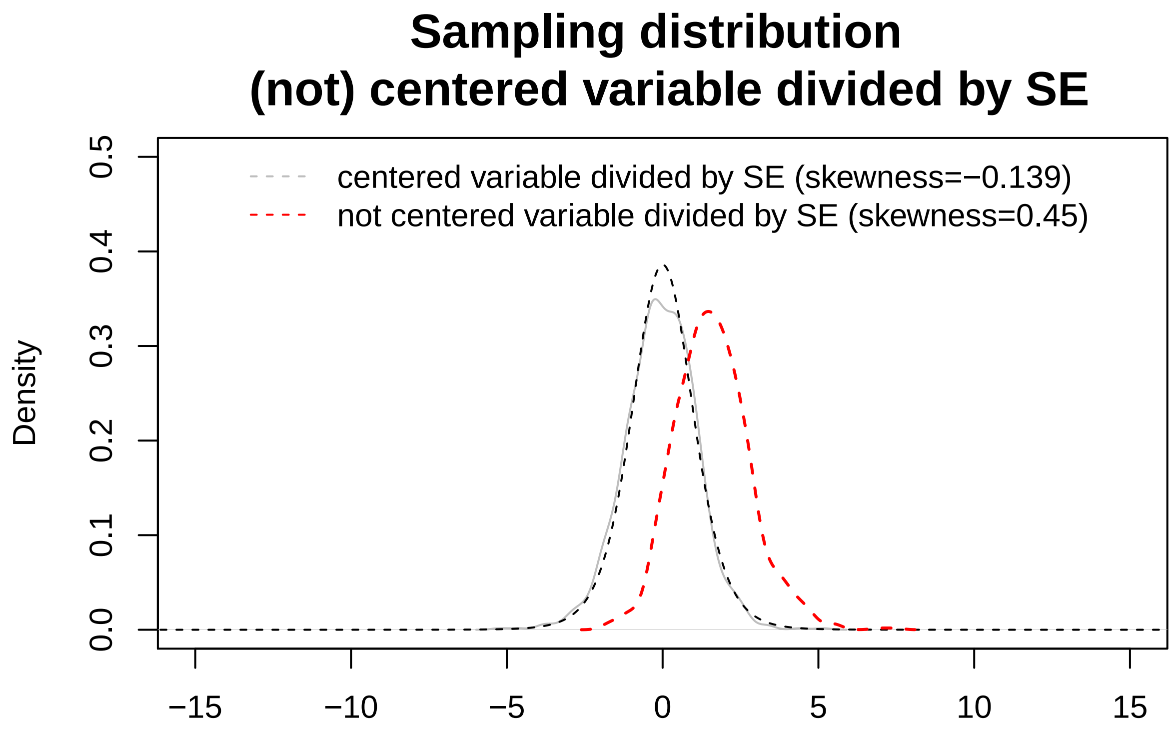

Figure 3. Sampling distribution of centered mean difference divided by SE (in grey, i.e. pivotal quantity) and not centered mean difference divided by SE (in red), assuming normality and homoscedasticity.

Considering the link between Δ and δ, we can compute confidence limits for $\Delta*$, and divide them by $\sqrt{N}$ in order to have confidence limits for $\delta*$. We first need to determine the noncentrality parameters of the distributions for which t_{Welch} corresponds respectively to the $1 - \frac{\alpha}{2}$ and to the $\frac{\alpha}{2}$ th. quantile.

$$P[t_{v,\Delta*_L} \geq t_{Welch}] = \frac{\alpha}{2}$$

and

$$P[t_{v,\Delta*_U} \leq t_{Welch}] = \frac{\alpha}{2}$$

.

With v approximated by $\hat{v} = \frac{(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1-1}+\frac{(\frac{S_2^2}{n_2})^2}{n_2-1}}$ (Shieh, 2013)

Second, we divide $\Delta *_L$ and $\Delta *_U$ by $\sqrt{N}$ in order to have $\delta *_L$ and $\delta *_U$ (i.e. confidences limits for Shieh's $\delta *$).

Cumming, G., & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervales that are based on central and noncentral distributions. *Educational and Psychological Measurement*, *61*(532), 532–574.

Shieh, G. (2013). Confidence intervals and sample size calculations for the weighted eta-squared effect sizes in one-way heteroscedastic ANOVA. *Behavior Research Methods*, *45*(1), 2–37. doi:10.3758/s13428-012-0228-7

Cumming, G., & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervales that are based on central and noncentral distributions. *Educational and Psychological Measurement*, *61*(532), 532–574.

Shieh, G. (2013). Confidence intervals and sample size calculations for the weighted eta-squared effect sizes in one-way heteroscedastic ANOVA. *Behavior Research Methods*, *45*(1), 2–37. doi:10.3758/s13428-012-0228-7