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.be

Reminder about Confidence Intervals

Introduction: How to compute a confidence interval around $\mu_1 - \mu_2$. Considering the link between confidence intervals and NHST approach, we can think of the 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 values 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:

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

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 $(\mu_1 - \mu_2)_0$. For example, 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.

Considering all these information we can 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 hypothesis $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_U$ (i.e. the symmetric t-distribution that is centered

[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.

around $(\mu_1 - \mu_2)_U$):

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

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

Under the assumption of iid normal distribution 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:

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

with $SE = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$. 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. 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 $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_U$:

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

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

It is not the most conventional way of computing confidence 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 a point estimator. As illustration, we will explain how to compute a confidence interval around Cohen's d_s (the explanation would be very similar for all other estimators). We previously mentioned that when the null hypothesis is true, $t_{Student}$ follows a central t-distribution. However, when the null hypothesis is false, the distribution of this quantity is not centered and noncentral t-distribution arises (Cumming & Finch, 2001), as illustrated in Figure 1.

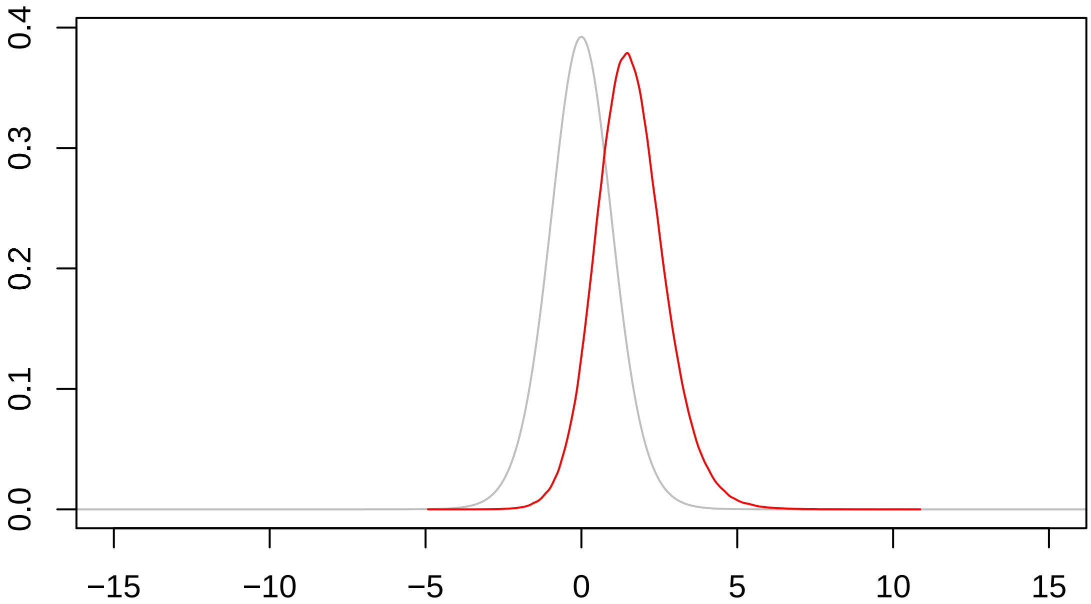

Figure 1. Sampling distribution of centered mean difference divided by SE (in grey) and not centered mean difference divided by SE (in red), assuming normality and homoscedasticity.

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:

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

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 words, we first need to determine the noncentrality parameters of the t-distributions for which $t_{Student}$ corresponds respectively to the quantiles $\left(1 - \frac{\alpha}{2}\right)$ and $\frac{\alpha}{2}$:

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

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

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:

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

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

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