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

Reference

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).

Shieh, G. (2013). Confidence intervals and sample size calculations for the standardized mean difference effect size between two normal populations under heteroscedasticity. Behavior Research Methods, 45,955–967

How to determine the confidence interval around a mean difference

Method 1: method based on the use of a pivotal quantity

When computing a (supposed normal) centered variable, divided by the standard error (i.e. an independant variable closely related with the χ^2 distribution), then computed quantity will follow a central *t*-distribution. This quantity is called a pivotal quantity (PQ), i.e. a quantity that is very interesting because its sampling distribution is not a function of the parameter we want to estimate (Cox & Hinkley, 1974 cited by Cumming and Finch, 2001). We can therefore use it, in order to define confidence limits for any parameter.

The method consists in four steps:

1) Compute a pivotal quantity (PQ) of the general form: (Estimator - parameter)/SE;

2) Determining the distribution of PQ;

3) Computing the confidence limits of PQ: determine a range of values, centered around 0, such as (1-alpha)% of the area under the distribution of PQ falls in this range;

4) Pivote in order to obtain the confidence interval around the parameter of interest.

As a first example, consider the case of 2 means difference, assuming normality and homoscedasticity. The pivotal quantity is defined as follows:

$$PQ = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{SE} \tag{1}$$

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

This quantity follows a t- distribution with $n_1 + n_2 - 2$ degrees of freedom (therefore, it depends only on n_1 and n_2, it does NOT depend on the parameter of interest, i.e. $\mu_1 - \mu_2$).

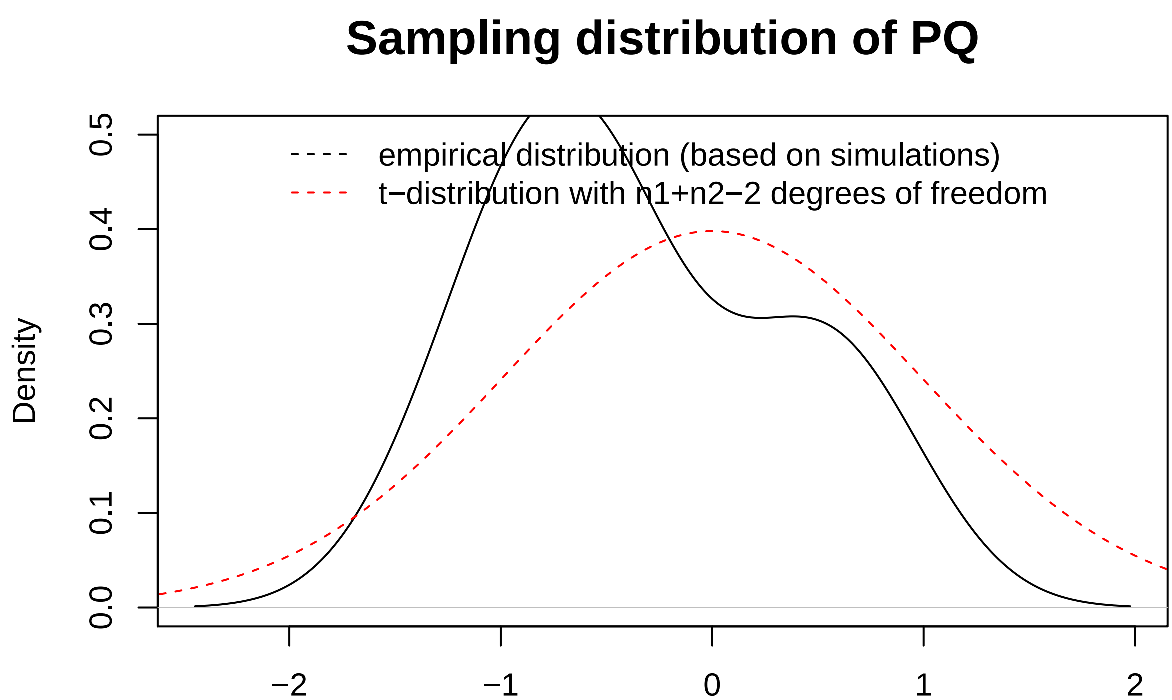

Figure 1. Sampling distribution of the pivotal quantity under the assumptions of normality and homoscedasticity

Because the theoretical distribution of PQ is known, one can compute the confidence limits, for any confidence level:

$$Pr[t_{n_1+n_2-2}(\frac{\alpha}{2}) < \frac{(\bar{X_1} - \bar{X_2}) - (\mu_1 - \mu_2)}{SE} < t_{n_1+n_2-2}(1 - \frac{\alpha}{2})] = 1 - \alpha \tag{2}$$

Because the t-distribution is symmetrically centered around 0, one can deduce that $t_{n_1+n_2-2}(\frac{\alpha}{2}) = -t_{n_1+n_2-2}(1 - \frac{\alpha}{2})$, and therefore:

$$Pr[-t_{n_1+n_2-2}(1 - \frac{\alpha}{2}) < \frac{(\bar{X_1} - \bar{X_2}) - (\mu_1 - \mu_2)}{SE} < t_{n_1+n_2-2}(1 - \frac{\alpha}{2})] = 1 - \alpha \tag{3}$$

In pivoting the inequation, one can deduce that:

$$Pr[-t_{n_1+n_2-2}(1 - \frac{\alpha}{2}) \times SE < (\bar{X_1} - \bar{X_2}) - (\mu_1 - \mu_2) < t_{n_1+n_2-2}(1 - \frac{\alpha}{2}) \times SE] = 1 - \alpha \tag{4}$$

$$\leftrightarrow Pr[-(\bar{X_1}-\bar{X_2})-t_{n_1+n_2-2}(1-\frac{\alpha}{2})\times SE < -(\mu_1-\mu_2) < -(\bar{X_1}-\bar{X_2})+t_{n_1+n_2-2}(1-\frac{\alpha}{2})\times SE] = 1-\alpha \tag{5}$$

$$\leftrightarrow Pr[(\bar{X_1}-\bar{X_2})+t_{n_1+n_2-2}(1-\frac{\alpha}{2})\times SE > \mu_1-\mu_2 > (\bar{X_1}-\bar{X_2})-t_{n_1+n_2-2}(1-\frac{\alpha}{2})\times SE] = 1-\alpha \tag{6}$$

$$\leftrightarrow Pr[(\bar{X_1}-\bar{X_2})-t_{n_1+n_2-2}(1-\frac{\alpha}{2})\times SE < \mu_1-\mu_2 < (\bar{X_1}-\bar{X_2})+t_{n_1+n_2-2}(1-\frac{\alpha}{2})\times SE] = 1-\alpha \tag{7}$$

As a second example, consider the case of 2 means difference, assuming normality and heteroscedasticity. The pivotal quantity is defined as follows:

$$PQ = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{SE} \tag{8}$$

With $SE = \sqrt{\frac{S_1^2}{n1} + \frac{S_2^2}{n2}}$

This quantity follows a t- distribution with $\frac{(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2})^2}{\frac{(\frac{S_1^2}{n_1})^2}{n_1-1}+\frac{(\frac{S_2^2}{n_2})^2}{n_2-1}}$ degrees of freedom (therefore, it depends on n_1 and n_2, S_1 and S_2, and does NOT depend on the parameter of interest, i.e. $\mu_1 - \mu_2$).

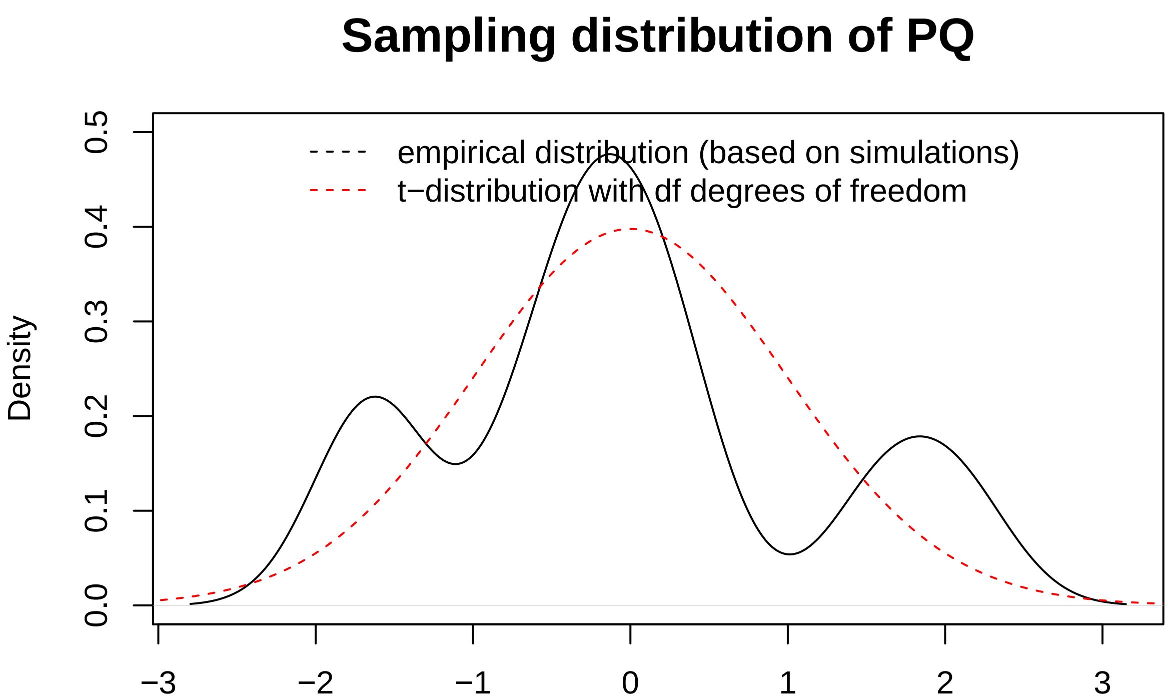

Figure 2. Sampling distribution of the pivotal quantity under the assumptions of normality and heteroscedasticity

Because the theoretical distribution of PQ is known, one can compute the confidence limits, for any confidence level (see the first example for more details):

$$Pr[(\bar{X}_1 - \bar{X}_2) - t_{n_1+n_2-2}(1-\frac{\alpha}{2}) \times SE < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{n_1+n_2-2}(1-\frac{\alpha}{2}) \times SE] = 1-\alpha \quad (9)$$

With SE $= \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

Method 2

We can also 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. In other words, we could define the lower limit $(\mu_1 - \mu_2)_L$ such as $\bar{X}_1 - \bar{X}_2$ 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 $\bar{X}_1 - \bar{X}_2$ exactly equals the quantile $\frac{\alpha}{2}$ of the central t-distribution of the null hypothesis $H_0 : \mu_1 - \mu_2 = (\mu_1 - \mu_2)_U$:

$$Pr[t_{n_1+n_2-2} \geq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_L}{SE}] = \frac{\alpha}{2} \quad (10)$$

$$Pr[t_{n_1+n_2-2} \leq \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_U}{SE}] = \frac{\alpha}{2} \quad (11)$$

This vision of the problem helps to understand how we calculate the confidence intervals around the effect size measures, as explained below.

How to determine the confidence interval around Cohen's δ

Consider the following quantity:

$$t_{Student} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{SE} \quad (12)$$

With $SE = \sigma_{pooled} \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$, $\sigma_{pooled} = \sqrt{\frac{(n_1-1)*S_1^2 + (n_2-1)*S_2^2}{n_1+n_2-2}}$, and $(\mu_1 - \mu_2)_0$ is the means difference under the null hypothesis. If the null hypothesis is true, this quantity is a

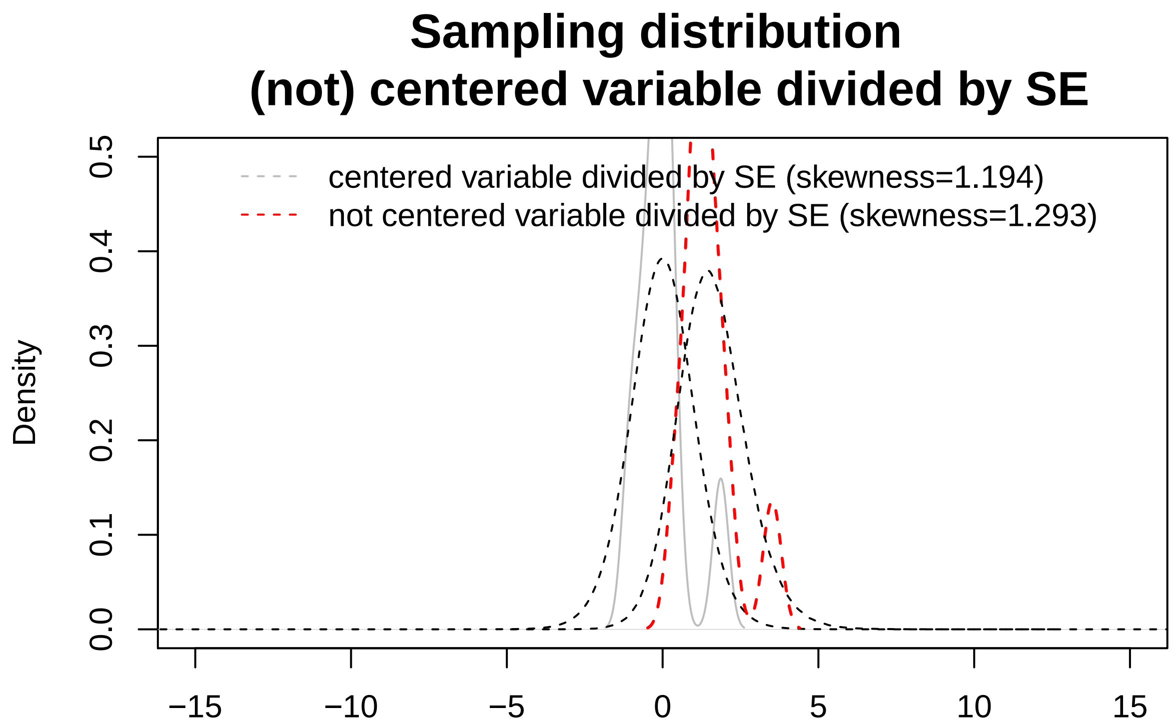

Figure 3. Sampling distribution of centered mean difference divided by SE (in grey, i.e. pivotal quantity) and not centered mean difference divided by SE (in red), assuming normality and homoscedasticity.

(supposed normal) centered variable, divided by an independant variable closely related with the χ^2. Therefore, as previously mentioned, it 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, as illustrated in Figure 3.

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

$$\Delta = \frac{\mu_1 - \mu_2}{\sigma_{pooled}} \times \sqrt{\frac{n_1 \times n_2}{n_1 + n_2}} \tag{13}$$

It is therefore 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 $\delta*$

Consider the following quantity:

$$t_{Welch} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)_0}{SE} \tag{14}$$

With $SE = \sqrt{\frac{S_1^2}{n1} + \frac{S_2^2}{n2}}$ and $(\mu_1 - \mu_2)_0$ is the means difference under the null hypothesis. As with $t_{Student}$, if the null hypothesis is true, this quantity is a (supposed normal) centered variable, divided by an independant variable closely related with the χ^2. It will therefore 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, as illustrated in Figure 4.

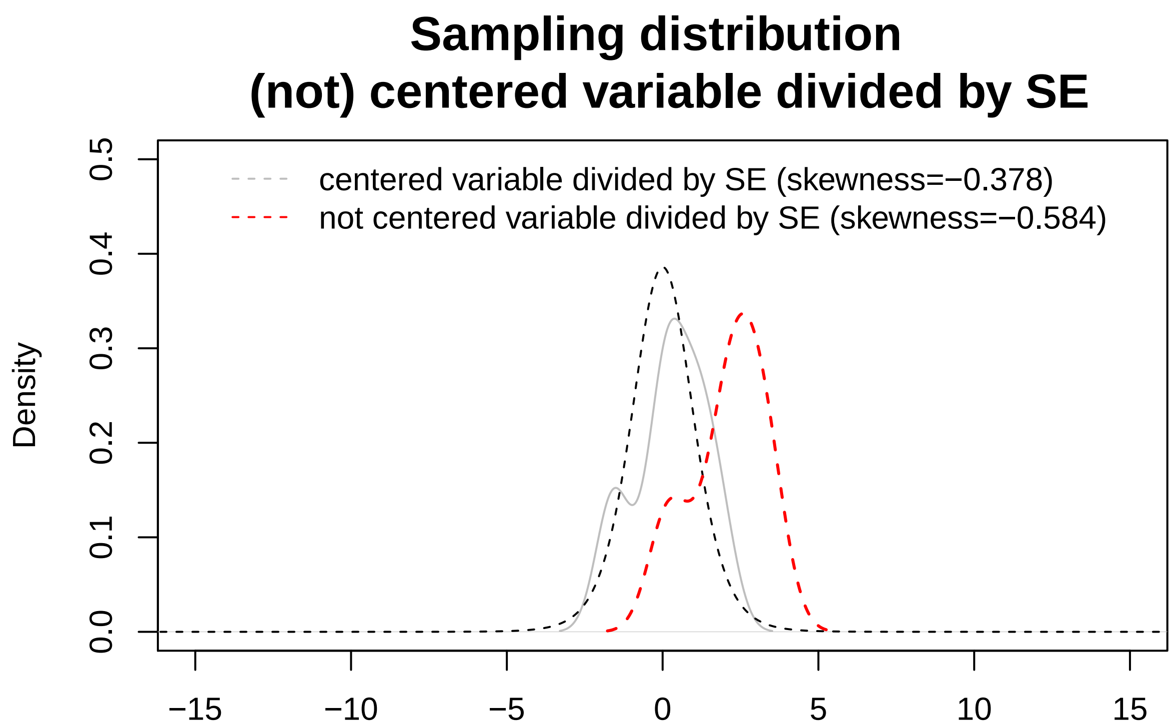

Figure 4. Sampling distribution of centered mean difference divided by SE (in grey, i.e. pivotal quantity) and not centered mean difference divided by SE (in red), assuming normality and homoscedasticity.

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

$$\Delta* = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{n_1/N} + \frac{\sigma_2^2}{n_2/N}}} \times \sqrt{N} \tag{15}$$

Again, it is therefore possible to 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 $v = \frac{(\frac{sd_1^2}{n_1}+\frac{sd_2^2}{n_2})^2}{\frac{(\frac{sd_1^2}{n_1})^2}{n_1-1}+\frac{(\frac{sd_2^2}{n_2})^2}{n_2-1}}$

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 $\delta*$).