5-1-2007

Another Look at the Confidence Intervals for the Noncentral T Distribution

Bruno Lecoutre
Centre National de la Recherche Scientifique and Université de Rouen, France

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Recommended Citation
DOI: 10.22237/jmasm/1177992600
Available at: http://digitalcommons.wayne.edu/jmasm/vol6/iss1/11
Another Look at Confidence Intervals for the Noncentral $T$ Distribution

Bruno Lecoutre
Centre National de la Recherche Scientifique and Université de Rouen, France

An alternative approach to the computation of confidence intervals for the noncentrality parameter of the Noncentral $t$ distribution is proposed. It involves the percent points of a statistical distribution. This conceptual improvement renders the technical process for deriving the limits more comprehensible. Accurate approximations can be derived and easily used.

Key words: Confidence intervals, noncentral $t$ distribution, lambda-prime distribution, Bayesian inference.

Introduction

In spite of several recent presentations (see especially, Fidler & Thompson, 2001; Bird, 2002), many potential users, as well as statistical instructors, consider computing or teaching confidence intervals for the noncentrality parameter of the Noncentral $t$ distribution to be very complex tasks. One of the conceptual difficulties is the lack of explicit formula. Although the considerable advances in computing techniques are supposed to render the task easy, they do not solve the conceptual difficulties.

The latter state is all the more deceptive in that when the number of degrees of freedom is large enough so that the Normal approximation holds the solution is very simple: the confidence limits are given by the percent points of a Normal distribution, as for the familiar case of an unstandardized difference between means. Thus, it can be expected that in the general case the limits would also be computed as the percent points of a statistical distribution. Unfortunately, this is not the case with the usual presentations.

Moreover, warnings about the accuracy of some computer programs of the Noncentral $t$ distribution (typically, the Noncentral $t$ algorithm fails for large sample size or effect size) cast doubt on some numerical results. Consequently, there remains the need for accurate approximations that are not currently easily available. Even when an exact computation is wanted, it needs an iterative algorithm, for which an accurate approximation constitutes a good starting point.

An alternative approach is proposed in this article that results in computing the confidence limits as the percent points of a statistical distribution. An interesting consequence of this conceptual improvement is that standard techniques to approximate statistical distributions can be used in order to find easy to use very accurate approximations. In conclusion, the question of the justification and interpretation of confidence intervals will be briefly examined.

Considerations and discussions regarding how and when to use confidence intervals for the Noncentral $t$ distribution, may be found elsewhere. Therefore, this article is not methodological. In this perspective, it will be sufficient, with no loss of generality, to consider the elementary case of the inference about a standardized difference between two means.
Computing confidence intervals from the Noncentral t distribution

When comparing two means, the t test statistic is the ratio \( \frac{(\bar{Y}_1 - \bar{Y}_2)}{E} \) of the two statistics, \( \bar{Y}_1 - \bar{Y}_2 \) that is an estimate of the population difference \( \mu_1 - \mu_2 \) and the standard error \( E \) of that estimate (see e.g., Fidler & Thomson, 2001, p. 587). In other words, \( E \) is an estimate of the standard deviation of the two statistics, \( \sigma \), one has \( E = \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \).

The sampling distribution of the ratio \( \frac{(\bar{Y}_1 - \bar{Y}_2)}{E} \) is a Noncentral t distribution with \( df \) degrees of freedom and a noncentrality parameter \( \lambda \), equal to \( (\mu_1 - \mu_2)/E \). This distribution is usually written \( t'_{df}(\lambda) \). The noncentrality parameter is termed \( \hat{\lambda} \), as in Algina and Keselman (2003), in order to avoid confusion with the population effect size. Formally, the Noncentral t distribution is the noncentrality parameter \( \lambda \) plus the standard Normal \( z \) distribution, all divided by the square root of the usual Chi-square distribution divided by the degrees of freedom (see e.g., Fidler & Thomson, 2001, p. 589):

\[
t'_{df}(\lambda) = (\lambda + z) / \sqrt{\chi^2_{df} / df}.
\]

The conceptual difficulties come from the fact that finding the limit \( \hat{\lambda}_L \) involves as many different distributions as considered \( \lambda \) values. A practical consequence is that it is a highly difficult task to derive accurate approximations.

An alternative approach: computing confidence intervals as percent points of the Lambda-prime distribution

An alternative solution consists in computing the confidence limits for the noncentrality parameter as percent points of a statistical distribution. When \( df \) is large enough so that the normal approximation holds, \( \hat{\lambda}_L \) is simply the 100\( \alpha/2 \) percent point of the standardized Normal distribution with mean \( t_{df} \). This can be generalized by introducing an appropriate statistical distribution. Even if it has not been made explicit in the usual presentations, this distribution is in fact not unfamiliar (without mentioning the fiducial and Bayesian presentations discussed in the conclusion).

Indeed, it is usual to plot \( p_{\lambda} \) (or its complement \( 1-p_{\lambda} \)) as a function of \( \lambda \). An illustration is given in Figure 1 for \( t_{df} = \pm 1.0076 \) with \( df = 22 \) (hence a \( p \)-value \( p = 0.3246 \), two-sided), which corresponds to the two-group A way data example given by Fidler & Thomson (2001, p. 586). The \( p_{\lambda} \) value increasingly varies from zero (when \( \lambda \) tends to -\( \infty \)) to one (when \( \lambda \) tends to +\( \infty \)), so that the corresponding curve is nothing else than the cumulative distribution function of a probability distribution. Such a graphical representation is commonly proposed to get a graphical solution for the confidence limits (see, for instance, Steiger & Fouladi, 1997, pp. 240), but the proponents fail to recognize that, in doing this, they implicitly define the confidence limits as the percent points of this probability distribution.

\[
p_{\lambda} = \Pr(t'_{df}(\lambda) > t_{df}) = \alpha/2.
\]
As for the Noncentral t, this distribution can be easily defined from the Normal and Chi-square distributions, but the result has not been so popularized. \((\overline{Y}_1 - \overline{Y}_2)/E > t_{CALC}\) can be equivalently written as \(\overline{Y}_1 - \overline{Y}_2 - t_{CALC} E > 0\). Consequently, \(p_{\lambda}\) is the probability that \(\overline{Y}_1 - \overline{Y}_2 - t_{CALC} E\) exceeds zero.

The sampling distribution of \(\overline{Y}_1 - \overline{Y}_2 - t_{CALC} E\) can be formally defined from independent standard Normal and Chi-square distributions as:

\[
e(\lambda + z - t_{CALC} \sqrt{\chi^2_{df}/df}),
\]

so that

\[
p_{\lambda} = \Pr(\overline{Y}_1 - \overline{Y}_2 - t_{CALC} E > 0) = \Pr(-z + t_{CALC} \sqrt{\chi^2_{df}/df} < \lambda)
\]

\[
= \Pr(z + t_{CALC} \sqrt{\chi^2_{df}/df} < \lambda),
\]

because the Normal distribution is symmetric around zero.

Thus, \(p_{\lambda}\) can be computed from the distribution characterized by \(z + t_{CALC} \sqrt{\chi^2_{df}/df}\). This distribution, which was considered (with no name) by Fisher (1990/1973, pp. 126-127) in the fiducial framework, was called Lambda-prime in Lecoutre (1999). It is also a noncentral distribution, again with \(df\) degrees of freedom, but with noncentrality \(t_{CALC}\). Formally:

\[
A'_{df}(t_{CALC}) = z + t_{CALC} \sqrt{\chi^2_{df}/df}.
\]

Consequently, it is possible to inverse in some sense the problem in (1) and compute \(p_{\lambda}\) as the probability that the Lambda-prime distribution with noncentrality \(t_{CALC}\) is smaller than \(\lambda\):

\[
p_{\lambda} = \Pr(A'_{df}(t_{CALC}) < \lambda).
\]

Thus, the curve in Figure 1 is the cumulative distribution function of the Lambda-prime distribution with 22 degrees of freedom and noncentrality +1.0076.

In order to find the limit, solve

\[
p_{\lambda} = \Pr(A'_{df}(t_{CALC}) < \lambda) = \alpha/2. \tag{2}
\]
(2) is technically equivalent to (1) and requires a similar iterative process, but it has a conceptual advantage. Indeed, it involves a unique distribution, so that $\lambda_L$ is the 100$\alpha/2$ percent point of the $\Lambda'_{df}(t_{CALC})$ distribution. In the same way, the upper limit $\lambda_U$ is its 100(1-$\alpha/2$) percent point. For instance, in Figure 1, the limits $\lambda_L = -0.986$ and $\lambda_U = +2.979$ of the 95% confidence interval are respectively the 2.5 and 97.5 percent points of the $\Lambda'_{22}(+1.0076)$ distribution.

Note again that the statistic $-\sum_{i=1}^{2} Y_i$ should not be regarded as less natural than the $t$ test statistic. Indeed, it is similar to the familiar limits $\bar{Y}_1 - \bar{Y}_2 \pm t_{1-\alpha/2} E$ of the 100(1-$\alpha$)% confidence interval for a raw difference. This analogy will be discussed in the conclusion.

Approximations of the 100$\pi$ percent point of the $\Lambda'_{df}(t)$ distribution

Beyond its conceptual simplification, the alternative approach allows to derive accurate approximations. In this Section, in order to simplify the notations, $t_{CALC}$ will be written $t$.

Numerical example

Consider Bird’s first example (Bird, 2002, p. 206), which is also considered in Algina and Keselman (2003). There were three independent groups of size 30 each, with means $\bar{Y}_1 = 22.467$, $\bar{Y}_2 = 24.933$, and $\bar{Y}_3 = 32.000$ and within group standard deviation 7.435. Bird reported the 97.5% confidence intervals of two standardized contrasts $(\bar{Y}_1 + \bar{Y}_2)/2 - \bar{Y}_3$ and $\bar{Y}_1 - \bar{Y}_2$. The computations for the first contrast will be detailed to illustrate the approximation methods. For this contrast, the $t$ test statistic is $t = -4.9924$ ($df = 87$). The exact confidence interval of $\lambda$ is: [-7.3766, -2.5844].

It can be computed by the usual method based on the Noncentral $t$ distribution using the available programs (for instance the Noncentral Distribution Calculator of Steiger, 2004), or alternatively as the 1.25 and 98.75 percent points of the Lambda-prime distribution with 87 degrees of freedom and eccentricity -4.9924.

Three approximation methods will be considered. The results for the two contrasts of interest are presented in Table 1. The limits for the standardized contrast $(\bar{Y}_1 + \bar{Y}_2)/2 - \bar{Y}_3$ in Table 1 are obtained by multiplying the limits for $\lambda$ by the appropriate constant (0.223607) referred as SE in Bird’s Table 7, page 208. Note that this constant can be simply computed as the ratio of the observed standardized contrast to the $t$ test value: $-1.1163/-4.9924 = 0.2236$.

### Table 1. Bird’s example: comparison of the three approximation methods

<table>
<thead>
<tr>
<th>Approximation</th>
<th>Bird</th>
<th>Normal</th>
<th>Chi-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>LowerLimit</td>
<td>-1.6264</td>
<td>-1.6489</td>
<td>-1.6495</td>
</tr>
<tr>
<td>UpperLimit</td>
<td>-0.5779</td>
<td>-0.5773</td>
<td>-0.5779</td>
</tr>
<tr>
<td>LowerLimit</td>
<td>-0.9123</td>
<td>-0.9123</td>
<td>-0.9123*</td>
</tr>
<tr>
<td>UpperLimit</td>
<td>+0.2506</td>
<td>+0.2506</td>
<td>+0.2506*</td>
</tr>
</tbody>
</table>

Note: Normal approximation

*Normal approximation
Three Approximation Methods - Bird’s Approximation

Reconsidered in the new approach, the Bird (2002, p. 203) approximation of the 100π percent point of the Lambda-prime consists in adding \( t \) and the 100π percent point of the standard (central) \( t \) distribution with the same degrees of freedom:

\[
\Lambda'_{df,\pi}(t) = t + t_{df,\pi}.
\]

The approximate 100(1-\( \alpha \))% confidence interval for \( \lambda \) is obtained:

\[
t \pm t_{df,1-\alpha/2},
\]

hence here for \( t_{87,0.9975} = +2.2809 \) the approximate confidence interval: \((-4.9924 \pm 2.2809 \rightarrow [-7.2733, -2.7116]\).

Algina and Keseleman (2003) found that the accuracy of this approximation does vary with the magnitude of the parameter, which can be verified in Table 1.

A simple normal approximation

The Lambda-prime distribution is generally asymmetric. However, when \( t = 0 \) it reduces to the standard Normal distribution, and when \( df \) is large it tends to the \( N(t,1) \) distribution. So we can expect that a Normal approximation with the same mean and variance is appropriate, at least for small \( t \) and for large \( df \).

The \( \Lambda'_{df} \) distribution has mean \( M = kt \) where

\[
k = \frac{2}{\Gamma(df)} \frac{1}{\sqrt{\Gamma(df+1)}} 
\]

and variance

\[
V = 1 + t^2 (1-k^2) = 1 + t^2 - M^2,
\]

so, that it can be approximated by the \( N(M,V) \) distribution:

\[
\Lambda'_{df,\pi}(t) = kt + z_{\pi} \sqrt{1 + t^2 (1-k^2)}.
\]

The approximate 100(1-\( \alpha \))% confidence interval for \( \lambda \) is obtained:

\[
kt \pm z_{1-\alpha/2} \sqrt{1 + t^2 (1-k^2)}.
\]

In order to find \( k \), one can compute its logarithm:

\[
\log(k) = \frac{2}{\log(df)} + \frac{1}{\log(\Gamma(df))} - \logGamma((df+1)/2) - \logGamma(df/2),
\]

and then take the exponential of \( \log(k) \). \( \logGamma(x) \) is the logarithm of the Gamma function \( \Gamma(x) \), that generalizes factorials to numbers beyond the integers. It is standard and for instance available in Excel. \( k \) can also be computed using the series expansion (Johnson & Welch, 1939):

\[
k = \frac{1}{4df} + \frac{1}{32df^3} + \frac{5}{128df^5} + \cdots
\]

Alternatively, Table 2 can be used for finding the wanted value.
In Bird’s example, it is found that $k = 0.997131$, $M = -4.9781$, and $V = 1.1428$, hence for $z_{0.9785} = +2.2414$ the approximate confidence interval for $\lambda$: -4.9781 ± 2.2414(0.951533) → [-7.3742,-2.5820] that is close to the exact interval.

A Chi-square approximation

For large $t$ values, a better approximation can be found that takes into account the asymmetry of the distribution. This needs to consider the third central moment that can be deduced from the mean:

$$W = 2k^2 - \frac{2df-1}{df} kt^3 = 2M^3 - \frac{2df-1}{df} t^2 M.$$ 

Next, compute the skewness of the distribution as the ratio of $W$ to the third power of the square root of the variance $V$ (i.e. $W/V^{3/2}$). The skewness is a measure of the degree of asymmetry of the distribution. When it is small, one can use the Normal approximation $N(M,V)$ above. For practical applications, it was empirically found that a more sophisticated approximation is not necessary when the skewness is smaller than 0.001. Otherwise, the following Chi-square approximation that fits the skewness can be used. It involves again reasonably simple computations. Let

$$c = \frac{W}{4V}, \quad q = \frac{V}{2c^2} \quad \text{and} \quad a = M - qc.$$ 

Then the approximation is given by percent points of the Chi-square distribution with $q$ degrees of freedom:

$$A'_{df,\nu} (t) = a + c \chi_{q,\nu}^2 \quad \text{if} \quad c > 0,$$
$$A'_{df,\nu} (t) = a + c \chi_{q,1-\nu}^2 \quad \text{if} \quad c < 0.$$
If \( t > 0 \) (which is equivalent to \( c > 0 \)), we get the approximate 100(1-\( \alpha \))% confidence interval for \( \lambda \): 
\[
[a + c \chi^2_{q,\alpha/2}, a + c \chi^2_{q,1-\alpha/2}]
\]
If \( t < 0 \), the limits are exchanged.

In practice \( q \) is generally very large and the Wilson and Hilferty (1931) approximation can be used (this is needed if your computer program does not work for high degrees of freedom values):

\[
\chi^2_{q,\pi} \approx q \left( \frac{2}{9q} z_\pi + 1 - \frac{2}{9q} \right)^3
\]

Some programs for the Chi-square distribution accepts only integer degrees of freedom. In this case, the 100\( \pi \) percent point of the Gamma distribution with parameter \( q/2 \) can be used alternatively:

\[
\chi^2_{q,\pi} = 2\text{Gamma}_{q}(q/2)
\]

In Bird’s example, it is found that \( W = -0.0041 \), \( c = -0.0009016 \), \( q = 702948.01 \) and \( a = 628.7998 \), hence the approximate confidence interval (computations have been performed with the maximum number of decimals for intermediate values): 
\[
[628.7998 \times 705608.34, 628.7998 - 0.0009016 \times 702993.06] \rightarrow [-7.3766, -2.5844]
\]

where \( \chi^2_{702948.01,0.0125} = 700293.06 \) and \( \chi^2_{702948.01,0.09875} = 705608.34 \) are computed using the Wilson-Hilferty approximation. This interval coincides with the exact interval with four decimal place accuracy.

A Comparison of the Three Methods

Table 3 gives a more systematic comparison of the three approximation methods. The exact probability levels associated with the different approximations of the 100\( \pi \) percent point of the \( \Lambda_{df}(t) \) distribution are reported for 100\( \pi \) = 2.5 and 100\( \pi \) = 97.5 (which gives the limits of the 95% confidence interval), and for 100\( \pi \) = 0.5 and 100\( \pi \) = 99.5 (which gives the limits of the 99% confidence interval). In the two cases, results are given for 10 and 50 degrees of freedom. They are reported only for positive values of \( t \); the results for negative values can be deduced by symmetry.

Bird’s approximation is very inaccurate for small \( df \) or large \( t \) and can hardly be recommended. By contrast, the simple Normal approximation works very well. The Chi-square approximation is quasi exact for most practical applications.
Table 3. Exact probability levels associated with the three approximations of the 100π percent point of the $\Lambda_{df}(t)$ distribution

<table>
<thead>
<tr>
<th>$t$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.29</td>
<td>98.71</td>
<td>2.23</td>
<td>97.77</td>
<td>2.50</td>
<td>97.50</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>0.25</td>
<td>1.33</td>
<td>98.72</td>
<td>2.24</td>
<td>97.77</td>
<td>2.50</td>
<td>97.50</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>0.50</td>
<td>1.38</td>
<td>98.70</td>
<td>2.26</td>
<td>97.77</td>
<td>2.50</td>
<td>97.50</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>0.75</td>
<td>1.46</td>
<td>98.66</td>
<td>2.28</td>
<td>97.76</td>
<td>2.50</td>
<td>97.50</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>1.00</td>
<td>1.57</td>
<td>98.60</td>
<td>2.31</td>
<td>97.74</td>
<td>2.49</td>
<td>97.49</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>1.50</td>
<td>1.86</td>
<td>98.41</td>
<td>2.39</td>
<td>97.69</td>
<td>2.48</td>
<td>97.48</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>2.00</td>
<td>2.27</td>
<td>98.10</td>
<td>2.50</td>
<td>97.61</td>
<td>2.46</td>
<td>97.46</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>2.50</td>
<td>2.81</td>
<td>97.68</td>
<td>2.63</td>
<td>97.50</td>
<td>2.42</td>
<td>97.43</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>3.00</td>
<td>3.50</td>
<td>97.14</td>
<td>2.80</td>
<td>97.37</td>
<td>2.39</td>
<td>97.39</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>4.00</td>
<td>5.31</td>
<td>95.75</td>
<td>3.22</td>
<td>97.01</td>
<td>2.30</td>
<td>97.31</td>
<td>2.50</td>
<td>97.50</td>
</tr>
<tr>
<td>5.00</td>
<td>7.63</td>
<td>94.01</td>
<td>3.77</td>
<td>96.53</td>
<td>2.22</td>
<td>97.24</td>
<td>2.49</td>
<td>97.50</td>
</tr>
<tr>
<td>10.00</td>
<td>21.04</td>
<td>84.57</td>
<td>8.19</td>
<td>92.66</td>
<td>1.95</td>
<td>97.04</td>
<td>2.47</td>
<td>97.51</td>
</tr>
<tr>
<td>15.00</td>
<td>30.54</td>
<td>77.72</td>
<td>14.10</td>
<td>87.58</td>
<td>1.84</td>
<td>96.97</td>
<td>2.46</td>
<td>97.52</td>
</tr>
<tr>
<td>20.00</td>
<td>36.36</td>
<td>73.27</td>
<td>19.72</td>
<td>82.78</td>
<td>1.80</td>
<td>96.94</td>
<td>2.45</td>
<td>97.52</td>
</tr>
<tr>
<td>25.00</td>
<td>40.12</td>
<td>70.25</td>
<td>24.39</td>
<td>78.77</td>
<td>1.78</td>
<td>96.93</td>
<td>2.45</td>
<td>97.52</td>
</tr>
</tbody>
</table>

Note. Normal approximation (exact for $t = 0$)

<table>
<thead>
<tr>
<th>$t$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
<th>$df = 10$</th>
<th>$df = 50$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.08</td>
<td>99.92</td>
<td>0.37</td>
<td>99.63</td>
<td>0.50</td>
<td>99.50</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>0.25</td>
<td>0.08</td>
<td>99.92</td>
<td>0.37</td>
<td>99.63</td>
<td>0.50</td>
<td>99.50</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>0.50</td>
<td>0.08</td>
<td>99.92</td>
<td>0.38</td>
<td>99.63</td>
<td>0.50</td>
<td>99.50</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>0.75</td>
<td>0.09</td>
<td>99.92</td>
<td>0.38</td>
<td>99.63</td>
<td>0.50</td>
<td>99.50</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>1.00</td>
<td>0.11</td>
<td>99.91</td>
<td>0.39</td>
<td>99.62</td>
<td>0.50</td>
<td>99.50</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>1.50</td>
<td>0.14</td>
<td>99.88</td>
<td>0.41</td>
<td>99.60</td>
<td>0.49</td>
<td>99.49</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>2.00</td>
<td>0.20</td>
<td>99.83</td>
<td>0.44</td>
<td>99.58</td>
<td>0.48</td>
<td>99.48</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>2.50</td>
<td>0.30</td>
<td>99.74</td>
<td>0.48</td>
<td>99.55</td>
<td>0.46</td>
<td>99.46</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>3.00</td>
<td>0.44</td>
<td>99.61</td>
<td>0.53</td>
<td>99.50</td>
<td>0.45</td>
<td>99.45</td>
<td>0.50</td>
<td>99.50</td>
</tr>
<tr>
<td>4.00</td>
<td>0.92</td>
<td>99.17</td>
<td>0.67</td>
<td>99.38</td>
<td>0.41</td>
<td>99.41</td>
<td>0.49</td>
<td>99.49</td>
</tr>
<tr>
<td>5.00</td>
<td>1.77</td>
<td>98.44</td>
<td>0.86</td>
<td>99.21</td>
<td>0.37</td>
<td>99.37</td>
<td>0.49</td>
<td>99.49</td>
</tr>
<tr>
<td>10.00</td>
<td>11.13</td>
<td>91.64</td>
<td>3.03</td>
<td>97.23</td>
<td>0.25</td>
<td>99.27</td>
<td>0.47</td>
<td>99.52</td>
</tr>
<tr>
<td>15.00</td>
<td>21.38</td>
<td>84.59</td>
<td>7.25</td>
<td>93.53</td>
<td>0.21</td>
<td>99.23</td>
<td>0.45</td>
<td>99.53</td>
</tr>
<tr>
<td>20.00</td>
<td>28.67</td>
<td>79.33</td>
<td>12.32</td>
<td>89.21</td>
<td>0.19</td>
<td>99.22</td>
<td>0.44</td>
<td>99.53</td>
</tr>
<tr>
<td>25.00</td>
<td>33.67</td>
<td>75.54</td>
<td>17.15</td>
<td>85.12</td>
<td>0.18</td>
<td>99.21</td>
<td>0.44</td>
<td>99.53</td>
</tr>
</tbody>
</table>

Note. *Normal approximation (exact for $t = 0$)
Conclusion

Returning to the analogy between the statistic 
\[ \overline{Y}_1 - \overline{Y}_2 - t_{\text{CALC}} E \]  and the familiar limits 
\[ \overline{Y}_1 - \overline{Y}_2 \pm t_{1-\alpha/2} E \]  of the 100(1-\alpha)\% confidence interval for a raw difference, it can be tempting to consider the interval 
\[ \overline{Y}_1 - \overline{Y}_2 \pm t_{\text{CALC}} E. \]
Assume for instance that \( t_{\text{CALC}} \) is positive, then one can remark that \( t_{\text{CALC}} \) is the 100(1-p/2) percent point \( t_{1-p/2} \) of the t distribution, where \( p \) is the two-sided p-value of the usual t test. Thus, the analogy seems again more compelling.

By the definition of \( t_{\text{CALC}} \), for the data in hand the bounds of this interval are zero, the traditional null hypothesis value, and two times the observed difference, what Rosnow and Rosenthal (1996) called the counter-null value. In their methodological article, Rosnow and Rosenthal (page 336) considered such an interval. Taking the example of an observed difference between two means +0.266 and a p-value \( p = 0.23 \), they interpreted the specific null counter-null interval \([0, +0.532]\) as a 77% confidence interval, that is as a 100(1-p\%) confidence interval. This interpretation reveals a typical confusion between Frequentist and Bayesian probabilities.

In the Frequentist conception of confidence intervals, the confidence level is the proportion of repeated intervals that contain the (fixed) parameter; it is usually termed the coverage probability. The procedure, and in particular the confidence level, must be determined before knowing the data. In the case of the Rosnow and Rosenthal interval \([0, +0.532]\), two possibilities can be envisaged to define the procedure and thus to compute the coverage probability.

Nevertheless, the procedure can proceed by computing the interval 
\[ \overline{Y}_1 - \overline{Y}_2 \pm t_{\text{CALC}} E, \]  with the data dependent value \( t_{\text{CALC}} \). For each repeated sample the bounds of this interval are zero and the particular counter-null value for this sample. Of course, the coverage probability of this interval varies with the parameters and it is not equal to 0.77 (except for two particular values of the ratio \((\mu_1-\mu_2)/\varepsilon\) symmetrical around zero).

The procedure can also proceed by computing the interval 
\[ \overline{Y}_1 - \overline{Y}_2 \pm t_{1-p/2} E, \]  with the fixed value \( p = 0.23 \) for each repeated sample. The coverage probability of this interval is 0.77. However, this is not a Frequentist approach, because 0.77 has been determined by the data in hand. Clearly, 0.77 is a data dependent probability, which needs a Bayesian approach to be correctly interpreted. The Bayesian inference associates to the interval \([0, +0.532]\) the posterior probability that this interval contains the parameter, given the data.

Although confidence intervals refer to a Frequentist justification, they are often (mis)interpreted in Bayesian terms. The distinction between the Frequentist coverage probability and the Bayesian posterior probability is all the more subtle in the present situation that it turns out that it is correct from a Bayesian viewpoint to say that there is a 77% chance that the interval \([0, +0.532]\) contains \((\mu_1-\mu_2)/\varepsilon\) or again in the example in Figure 1 to say that there is a 95% chance that the interval \([-0.986, +2.979]\) contains the noncentrality parameter \( \lambda \). This simply assumes a prior distribution that does not favor any particular value of the parameters, what Bayesian called a non-informative prior.

This distribution is revised by the data and the corresponding posterior distribution for the noncentrality parameter \( \lambda \) is just the \( \Lambda_{df}(t_{\text{CALC}}) \) distribution. Consequently, the Lambda-prime distribution, in addition to its status of sampling distribution gains the status of a probability distribution that expresses the uncertainty about the unknown parameter \( \lambda \). One can use the confidence interval for \( \lambda \) with the benefits of both the Frequentist and Bayesian interpretations and without worrying about the correct justification (not to speak of Fisher’s fiducial argument).

The fact that even experts in statistics are not immune to conceptual confusions and interpret Frequentist intervals in Bayesian terms should not be regarded as an error. Rather this means that, as most statistical inference users, they are also, perhaps without knowing it, interested in the Bayesian probability. This should invite us not to radicalize the opposition between the Bayesian and Frequentist inferences.
but rather to consider their interplay. This is a difficult challenge, but it is already well advanced in the statistical literature (see Bayarri & Berger, 2004).

References
