Exploratory Software for Confidence Intervals

www.psy.latrobe.edu.au/esci

User Notes

Geoff Cumming

School of Psychological Science
La Trobe University
Bundoora Australia 3086

ESCI (Pronounced "esky")
esci@latrobe.edu.au
Comments and suggestions are welcomed.

These notes are provided via free download from **www.psy.latrobe.edu.au/esci**
Watch the ESCI site for **news and updates**. See also the FAQ section.
User Notes, version: 23 July 2001

ESCI is published by La Trobe University,
Bundoora Australia 3086
2001

Acknowledgements
Michael Smithson raised noncentral distributions as a topic of interest.
Bruce Thompson prompted the Cumming & Finch paper, and advised at every stage.
Rodney Carr showed what Excel can do.
Fiona Fidler, Sue Finch and Neil Thomason collaborated and advised.
Joanna Leeman advised, and suggested the name ESCI.
Geoff Robinson provided the routines for calculating noncentral ***t***.
David Walsh built the website, developed graphics and advised.

Contents

Overview

ESCI is a set of interactive simulations that runs under Microsoft Excel. ESCI-delta is the first set of ESCI simulations, and comprises 6 Excel workbooks.

With ESCI-delta you can:

- Explore many **Confidence Interval** (CI) concepts.
- Calculate CIs for Cohen's **standardised effect size δ**.
- Explore **noncentral *t*** distributions and their role in **statistical power**.
- Use CIs for simple **meta-analysis**, using original units or ***d*** (standardised) effect sizes.
- Explore all these concepts via vivid interactive **graphical simulations**.

> ***Note*** Cohen's standardised effect size is a mean, or mean difference, divided by a standard deviation. It is thus a number of standard deviations, like a *z* score. In the Cumming and Finch (2001) article we use the symbol **δ** for the population parameter, and ***d*** for the sample statistic. (On the ESCI website ***d*** is used for both, because the symbol **δ** cannot be reliably displayed on all browsers.)

The rationale for ESCI

Statistics reform requires much wider use of **confidence intervals** (CIs) and **effect size** measures, in original units and in standardised units. We therefore need **CIs for standardised effect sizes**. Unfortunately in many cases these are not easily calculated, and require use of noncentral distributions. ESCI enables you to find the CI for Cohen's **δ** for your own data, which requires use of **noncentral *t*** distributions and an iterative algorithm. ESCI also allows you to explore noncentral ***t*** itself, and its most familiar application, which is the calculation of statistical power.

Statistics reform also requires wider use of meta-analysis. The best way to explore (and teach) simple meta-analysis may be via CIs, graphically presented. ESCI supports the development of **'meta-analytic thinking'** in original measurement units and in standardised units using Cohen's **δ**.

Future developments

Further ESCI simulations are in development. They will be announced on the ESCI website.

The CI primer article, Cumming and Finch (2001)

The following article explains confidence intervals for Cohen's **δ**. It gives an introduction to noncentral ***t*** distributions and discusses related concepts including power and simple meta-analysis. It is illustrated with part-images from ESCI. ESCI-delta is the set of six Excel workbooks that are referred to in this article.

Cumming, G., & Finch, S. (2001). A primer on the understanding, use and calculation of confidence intervals based on central and noncentral distributions. **Educational and Psychological Measurement, 61**(4, August), 530-572.

Abstract Reform of statistical practice in the social and behavioural sciences requires wider use of confidence intervals (CIs), effect size measures and meta-analysis. The authors discuss four reasons for promoting use of CIs: they (i) are readily interpretable; (ii) are linked to familiar statistical significance tests; (iii) can encourage meta-analytic thinking; and (iv) give information about precision. The authors discuss calculation of CIs for a basic standardised effect size measure, Cohen's **δ** (also known as Cohen's ***d***), and contrast these with the familiar CIs for original score means. CIs for **δ** require use of noncentral ***t*** distributions, which the authors apply also to statistical power and simple meta-analysis of standardised effect sizes. They provide the **ESCI** graphical software, which runs under Microsoft Excel, to illustrate the discussion. Wider use of CIs for **δ** and other effect size measures should help promote highly desirable reform of statistical practice in the social sciences.

In this article we outline the case for reform of statistical practice by researchers in psychology and other disciplines. We give formulas and explanations for calculating CIs in original measurement units, and for Cohen's **δ**. Some examples are included. Noncentral ***t*** distributions are described, and used in discussions of statistical power and of simple meta-analysis.

It may be most useful to read the paper while working with the ESCI-delta software.

Since writing the paper the new edition of the APA manual has appeared:

American Psychological Association. (2001). **Publication manual of the American Psychological Association** (5th ed.). Washington, DC: Author. http://www.apa.org/books/4200060.html

The following quotes from the **Manual** indicate that effect size measures and confidence intervals are likely to be required routinely by journal editors, and will be seen often by anyone reading published research. These are highly desirable developments.

> For the reader to fully understand the importance of your findings, it is *almost always necessary* to include some index of effect size or strength of relationship in your Results section. You can estimate the magnitude of effect or the strength of the relationship with a number of common effect size estimates... The general principle to be followed... is to provide the reader not only with information about

> statistical significance but also with enough information to assess the magnitude of the observed effect or relationship. (pp. 25-26, emphasis added)

> The reporting of confidence intervals... can be an extremely effective way of reporting results. Because confidence intervals combine information on location and precision and can often be directly used to infer significance levels, they are, in general, the best reporting strategy. The use of confidence intervals is therefore *strongly recommended*. (p. 22, emphasis added)

The aim of ESCI is to provide tools that may assist researchers, teachers and students to explore the relevant concepts and to analyse and present their data in accordance with these new requirements.

Downloading, installing and running ESCI

Downloading
At the ESCI site the downloads page gives download of these User Notes, and access to the software downloads webstore, which offers:

- Free download of the DEMO version of ESCI-delta.
- Download at low cost of the full version of ESCI-delta, with an INDIVIDUAL licence.
- Download of the full version of ESCI-delta, with a LOCAL NETWORK licence, which permits use on a network of not more than 50 computers.

For bulk deals for student licences, or any other licence enquiries please email **esci@latrobe.edu.au**

The download is of a 1.2MB zipped .exe file. Save this to disk, then double-click it for automatic unzipping into a folder, which contains the six Excel workbooks of ESCI-delta, total size 3.5MB. The largest of these is CIdelta, at 1.8MB. If the folder includes a README file please read that first.

DEMO version
The DEMO version is not time-limited. It is the same as the full version except that on each worksheet a few cells (parameters, or data entry cells) are greyed out. These cells may not be changed by the user. In addition the sheets are protected with a password and may not be changed.

Using ESCI
No installation is required. To run ESCI you need a licensed copy of Microsoft Excel97, or Excel2000, running under Windows 95 or later. (Unfortunately, as usual for Excel, the transfer to Macintosh is usually not fully successful.)

You must, if asked while opening a workbook, enable macros.

On each Intro sheet there are two hints:

1. Check that **comments** are visible. Some cells are marked with small red triangles. Rest the mouse there and see a popout comment. If these are not visible, go to Tool/Options/View and click 'Comment indicator only'. (Click 'None' to hide comments.) Comments give the main guidance for using the worksheets.

2. If your screen resolution is 1024x768, the usual Excel **zoom factor** of 100% should give a suitable size of image. With a screen resolution of 800x600 it may be necessary to change the zoom factor (on the toolbar, or go to View/Zoom) of each sheet to about 85%. Other zoom factors may be used, but avoid using 80% or less, or there is a risk that some numbers may not be displayed properly: If this problem occurs the number is replaced by hashes (####).

Nothing is hidden, so you can examine formulas as you wish. In the full version the worksheets are protected but without password (except the first Intro sheet, which may not be changed). You can therefore unprotect a sheet (Tools/Protection/Unprotect sheet) and

explore as you choose. **Caution:** It is very easy to make unwanted changes on an unprotected sheet. Be extremely cautious about saving.

Of course it is essential to keep an archive copy of the original workbooks. It is good practice to start work each time with a fresh copy of the original workbook, and to save only with a new name.

NonCentral *t* distributions and interactivity
Probabilities under the noncentral ***t*** distributions are calculated by summing a 100-term series. Curves are formed of 200 points, each requiring such a calculation. (In NonCentral t and Power you can see the 200x100 array of values by scrolling right.) In CIdelta the points of the two red noncentral ***t*** curves are calculated by summing a 50-term series, which gives acceptable accuracy in this situation and not such a large file. (The tail probabilities are calculated with the 100-term series, for accuracy.)

The extensive computation needed for the noncentral ***t*** curves means that updating of the graphs is not instantaneous. Only on a very fast computer does dragging of a slider give smooth responsive movement of the corresponding curve.

There are comments about accuracy in various places. In some extreme situations the calculation is not sufficiently accurate, so a value or a curve is not displayed, or information is given about the degree of precision achieved.

Entering data
In CIoriginal and CIdelta you type in your own data. You can paste data into the data cells, from the clipboard, after you have copied to the clipboard from, for example, a different Excel workbook, or Microsoft Word. There is a problem however: Material pasted into Excel bypasses any data validation checks, so for example text would not be rejected when pasted into an ESCI data cell. Even worse, the validation settings are wiped by the paste operation.

Therefore, paste data into ESCI with care, and check you have just one number in each cell. Don't save the file after pasting data because ESCI's data validation tests have been removed.

ESCI's graphs
As you move the mouse pointer over the graphs (Excel 'charts') in ESCI, you may see many yellow popouts with numbers and labels. At Tools/Options/Chart you can tick or clear two boxes to specify whether you will 'Show names' and/or 'Show values'. It may be best to clear the 'names' box since most names are not informative. However you may prefer to see 'values' since in many cases these give a precise value for a plotted point.

Printing
ESCI offers no additional printing facilities. You can use the Print Screen key to copy a full screen image to the clipboard, for pasting into, for example, Microsoft Paint or Word. Alternatively you can print from Excel: Print just the first page, or set an area for printing (File/Print Area) and print this.

The simulations that make up ESCI-delta

See the CIprimer article for the statistical context and formulas relevant for each simulation, and for suggested conclusions that may be drawn from exploration with the software. For several of the workbooks, follow the red numbers and note the popout comments.

Throughout ESCI-delta underlying populations are assumed Normal, and of homogeneous variance where that is relevant. All NHST (Null Hypothesis Significance Testing) is two-tailed. You can usually set the Type 1 error rate **α**, or the confidence level, to any value you wish. In Power and MAthinking the following symbols are used to mark conventional statistical significance levels (two-tail probabilities throughout):

ns	$.10 \leq p$
?	$.05 \leq p < .10$
*	$.01 \leq p < .05$
**	$.001 \leq p < .01$
***	$p < .001$

NonCentral t

*Explore noncentral **t** distributions and calculate accurate probabilities.*

The familiar ***t*** distribution has the single parameter ***df***. Noncentral ***t*** distributions have in addition a noncentrality parameter **Δ**. Here you can use the controls to see how the shape of the noncentral distribution (red) changes with ***df*** and the noncentrality parameter. Click boxes to display or hide a central ***t*** distribution (blue), for comparison, and various other features. You can also display accurate noncentral ***t*** tail probability values for your chosen ***df***, **Δ**, and ***t*** values. Noncentral ***t*** distributions are essential for statistical power and for CIs for Cohen's **δ**.

Power

*Explore statistical power, which in most cases requires noncentral **t** calculations.*

Click to display the central ***t*** distribution (blue) that applies when $\mathbf{H_0}$ is true, and/or the noncentral ***t*** distribution (red) that is appropriate when $\mathbf{H_a}$ is true. Display the two-tail rejection region for the statistical significance test, then the tail area(s) under the red curve that give statistical power. Vary ***df***, **α** (Type 1 error rate), and **Δ** (which is determined by the effect size **δ**, i.e. the distance between the null and alternative hypothesised mean values). You can also display the percentages of experiments that, if $\mathbf{H_a}$ is true, would be expected to give various significance test outcomes: These percentages may be surprising.

CIjumping

Repeated sampling, to illustrate basic concepts of confidence intervals (CIs).

Take independent samples from a normal population. Set the population parameters and select the sample size ***n***. Click to display or hide the population distribution and various

other features. Click to take a single sample, shown as a dotplot. Means of successive samples are shown cascading down the screen. Run the simulation and see the means 'dance', illustrating the (surprisingly large?) extent of sampling variability for your chosen population and ***n***. Set a confidence level and show a confidence interval for _ for each sample. Set a $_0$ comparison value (usually = _) for the mean and see which CIs capture this value. One key idea is that the CIs vary, not _.

CIoriginal

Calculate and display CIs for your data, for three simple experimental designs.

These three simple tools allow you to enter your own data for three designs: Case 1, Single Group; Case 2, Two Independent Groups; and Case 3, Paired Data. The data are shown as a dotplot. For Case 1 you select a $_0$ value as a reference value for the mean. The CI, based on your dataset, for the population mean is shown. Noting whether the CI captures $_0$ is equivalent to conducting the corresponding ***t*** test. These tools use original measurement units throughout and are provided as an introductory step towards the more complex CIdelta, in which standardised units (Cohen's ***d*** and **δ**) are shown alongside original units.

CIdelta

Calculate and display CIs for standardised effect sizes (Cohen's **δ***), for One and Two Group designs. Requires use of NonCentral* ***t*** *distributions.*

Enter your data for Case 1, Single Group, or Case 2, Two Independent Groups. The dotplot and CI are shown, in original units, as in CIoriginal. Set a $_0$ comparison value, which serves as the zero for measurement of Cohen's ***d*** and **δ**, the standardised effect size. Click to display the CI for **δ**, but note that this is not yet accurate. To find accurate values, the red noncentral ***t*** distributions must be positioned so that their tails that overlap ***d*** are of size (**α**/2). To achieve this you can move the sliders, or you can click the buttons to have Excel do the work.

MAthinking

Explore simple Meta-analysis, based on graphical display of CIs, for effect sizes in original measurement units, and for standardised effect sizes (Cohen's **δ***).*

The most basic meta-analysis consists of simple pooling over studies, weighted by sample size. This may be carried out in original units, or in standardised units. A simulation is provided for each of these, for the Case 1, Single Group design. For the standardised units, enter the ***d*** and ***n*** values for up to 10 previous studies. Click the buttons to have Excel find and display the CI for **δ** for each study, and for all the studies combined. You can also enter ***d*** and ***n*** for an additional study (perhaps our current study) and click to show the CI for this study, and for all studies combined. Click to show the statistical significance test results: Exploration of examples will probably suggest that effect sizes (and sample sizes) are much more influential and informative than significance levels.