HOW THE NONCENTRAL *t* DISTRIBUTION GOT ITS HUMP

Geoff Cumming
La Trobe University, Australia
g.cumming@latrobe.edu.au

Once upon a time there was only one t *distribution, the familiar central* t, *and it was monopolised by the Null Hypotheses (the Nulls), the high priests in Significance Land. The Alternative Hypotheses (the Alts) felt unjustly neglected, so they developed the noncentral* t *distribution to break the monopoly, and provide useful functions for researchers—calculation of statistical power, and confidence intervals on the standardised effect size Cohen's* d. *I present pictures from interactive* software *to explain how the noncentral* t *distribution arises in simple sampling, and how and why it differs from familiar, central* t. *Noncentral* t *deserves to be more widely appreciated, and such pictures and software should help make it more accessible to teachers and students.*

ONCE UPON A TIME

Once upon a time, many long years ago, there was only one *t* distribution. It was the familiar symmetric *t* distribution with a single parameter, the *degrees of freedom* (*df*). The Null Hypotheses (the Nulls), who were the high priests in Significance Land, had a monopoly on this *t* distribution. The Nulls tended to be grumpy because most of them had the very uninteresting value $\mu_0 = 0$, and also because—although they tried to keep this quiet—almost all of them were false! Furthermore, although researchers constantly invoked Nulls, they seemed to be interested merely in rejecting them. Little wonder that morale among Nulls was low, and that they guarded their monopoly on the only *t* distribution jealously!

One day several smart young Alternative Hypotheses (Alts) were discussing *their* predicament. There were many more Alts than Nulls, and the Alts tended to be much more interesting and lively than the cantankerous old Nulls. Alts tended to have interesting values (μ_A values) and they were proud of the fact that many Alts were actually true! Surely, researchers should be much more interested in finding true μ_A values, knowledge that could have practical use in the world, or could perhaps give strong support to a scientist's theory? There were a few enlightened researchers who used estimation and confidence intervals (CIs) to make their best estimate of true effects (Cumming and Finch, 2005), but most researchers still seemed interested only in examining a Null and hoping desperately that they could reject it—as if this told them anything specific about what was actually true!

One especially smart Alt had been watching the Nulls carefully; she wondered whether the Nulls received so much attention because of their *t* distribution. Researchers seemed to use *t* values all the time, and students, as part of their induction into the weird significance rituals declared compulsory by the high priests, were given lots of instruction about using the *t* distribution. Perhaps, if Alts could invent some sort of *t* distribution to suit their wonderful μ_A values, the monopoly of the Nulls would be broken, and researchers might take Alts more seriously? The Alt think-tank thought this an excellent idea. Soon they had developed what they called the *noncentral t* distribution, in contrast to the Nulls' ordinary old *central t* distribution.

The think-tank did a great job. Noncentral *t* is asymmetric, and thus much more interesting to behold. As well as *df*, it has a *noncentrality parameter* Δ, whose value depends on μ_A or, more precisely, on ($\mu_A - \mu_0$), the difference between μ_A and the null hypothesised value. Cool young Alts took to carrying their own noncentral *t* distributions with them, clearly visible in the shape of their backpacks. The think-tank designed the new distribution to have at least two really useful functions. First, unless you know the population SD σ—and you rarely do—then to calculate statistical power it is necessary to find an area under the noncentral *t* distribution. Second, if you use Cohen's *d*, a simple and very useful standardised effect size measure, you need to use noncentral *t*, and an iterative computer procedure, to calculate a CI on your *d* value.

Alas, even those valuable uses were not sufficient for the new distribution to be a best-seller. Perhaps the Nulls were too entrenched, or perhaps the Alts did not have sufficiently clever media relations? The beautiful, intriguing, and useful noncentral *t* distribution was rarely taught to students, and remained little-known among researchers. Cumming and Finch (2001) did help, by describing noncentral *t*, with formulas and pictures, and explaining how it permits calculation of power, and CIs on *d* values. They also provided software to help people explore noncentral *t* and its uses: *ESCI*, *Exploratory Software for Confidence Intervals*, www.latrobe.edu.au/psy/esci. *ESCI* runs under Microsoft *Excel*. Further descriptions of applications of noncentral *t* were given by Kline (2004) and Grissom and Kim (2005).

PICTURES TO EXPLAIN HOW SAMPLING GIVES NONCENTRAL *t*

Cumming and Finch (2001) did not, however, explain vividly and graphically how noncentral *t* arises from simple sampling. I believe I now have pictures and software that can do that. Consider the simplest sampling situation: Take random samples of size n from a population that is normally distributed, with mean μ and SD σ. Our sample has mean M and SD s. We can test hypotheses about μ (for example, the hypothesis $\mu = \mu_H$) by using as a test statistic, if we assume σ is known:

$$z = (M - \mu_H)/(\sigma/\sqrt{n}) \quad (1)$$

or, if we assume σ is not known:

$$t = (M - \mu_H)/(s/\sqrt{n}). \quad (2)$$

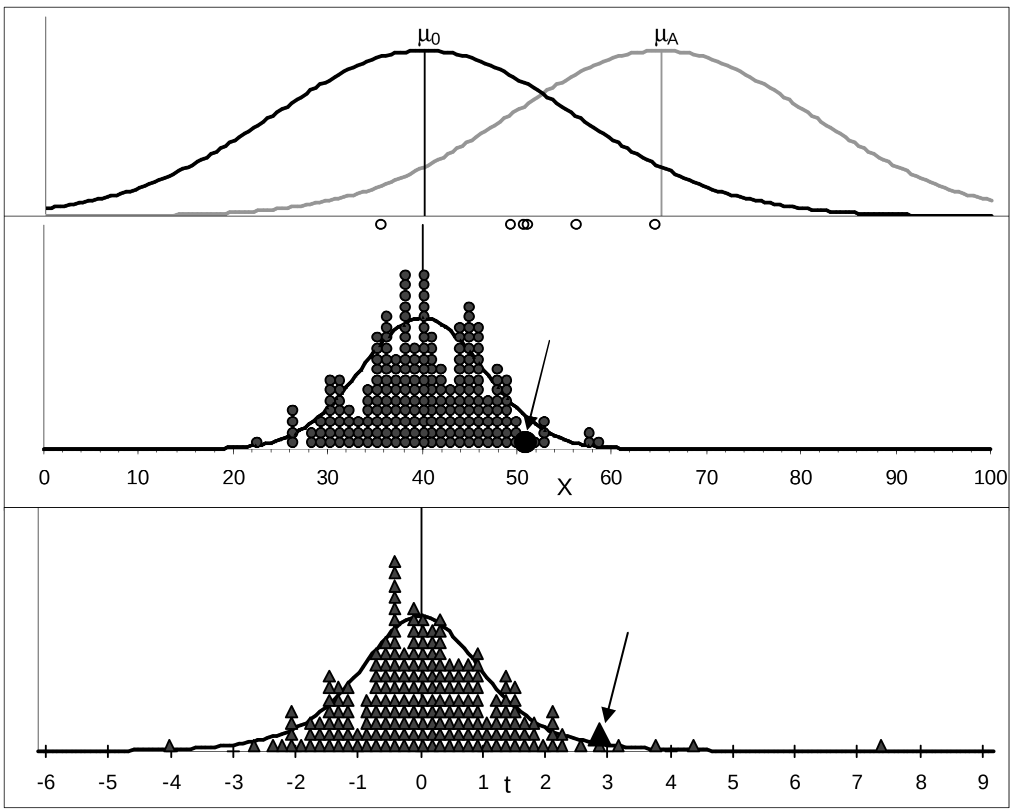

Figure 1: The simulation after taking 208 samples, each of size n = 6, assuming the Null Hypothesis (H_0: $\mu = \mu_0$) is true, so that sampling is from the left population in the upper panel, which has $\mu = 40$, $\sigma = 16$. The middle panel shows the dot plot of sample means M, and the sampling distribution of M, which is normal. The lower panel shows the dot plot of *t* values, and the (central) *t* distribution, which has df = n – 1 = 5. The six open circles are the data points of the latest sample; their M value (51.4) is marked by the large dot (arrowed) in the middle panel, and the corresponding *t* value (2.93) by the large triangle (arrowed) in the lower panel. For that sample, *s* happens to be 9.5, much less than σ.

Figure 1 shows the simulation after a run of samples, assuming H_0 is true. The middle panel shows the familiar sampling distribution of M. The horizontal axis in the lower panel is marked in units of population SE, which is $\sigma/\sqrt{n}$, either side of μ_0. This panel displays *t* values for each sample, calculated using Equation 2. The simulation uses colour, and has numerous controls that allow setting of parameter values, control of sampling, and display (or hiding) of many aspects of the display. If the control for 'assume σ known' is clicked on, the lower display shows *z* values, and the dot plot and curves are identical to those in the middle panel. If clicked off, *t* values are displayed, as in Figure 1. Clicking on and off repeatedly shows dynamically how the central *t* distribution compares with the corresponding normal distribution. A sample size of 6 is used here, to highlight the fat tails of the central *t* distribution when *df* is very small.

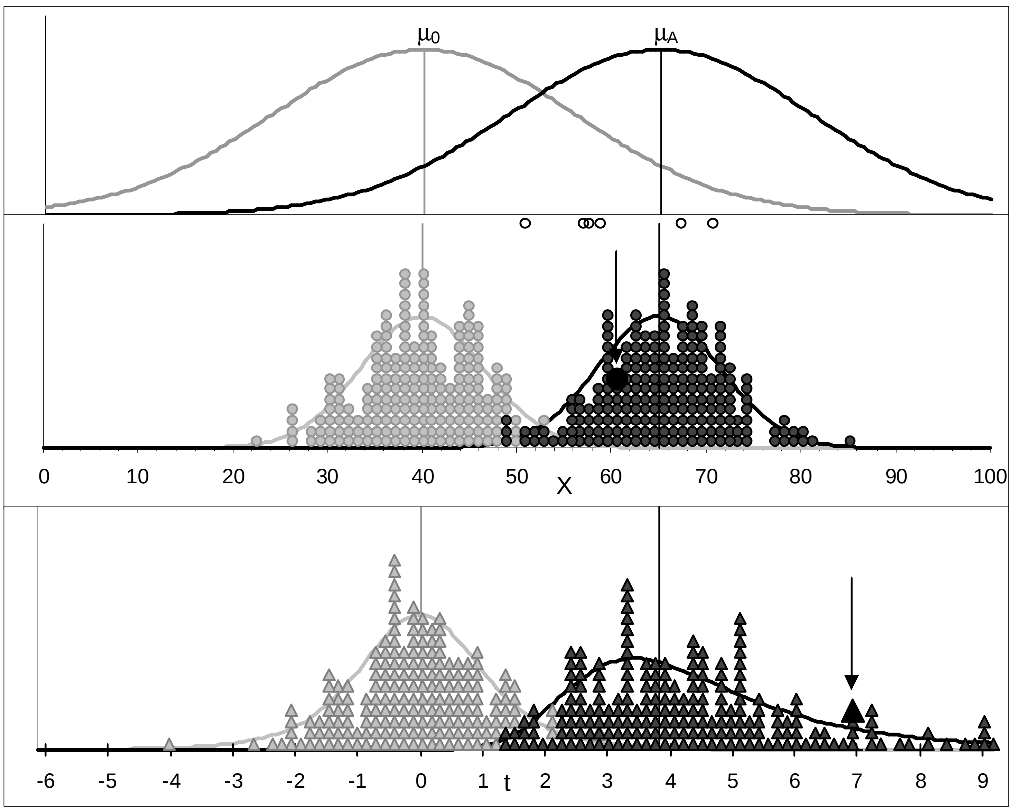

Figure 2: The simulation after taking 209 samples, each of size n = 6, assuming the Alternative Hypothesis (H_A: $\mu = \mu_A$) is true, so sampling is from the right population in the upper panel, which has $\mu = 65$, $\sigma = 16$. The middle panel shows the dot plot of sample means M, and the sampling distribution of M, which is normal. The lower panel shows the dot plot of $t = (M - \mu_0)/(s/\sqrt{n})$ values, and the noncentral *t* distribution, with df = 5 and noncentrality parameter $\Delta = (\mu_A - \mu_0)/(\sigma/\sqrt{n}) = 3.83$. The latest sample has M = 60.5, marked by the large dot in the middle panel (arrowed), s = 7.28, and t = 6.9 (the large triangle in the lower panel, arrowed).

Still referring to Figure 1, it is useful to think of *z*, in Equation 1, as measuring how far *M* falls from μ_0, in units of $\sigma/\sqrt{n}$. These units are constant—they do not depend on any sample value—and so can be used to mark the horizontal axis in the lower panel. If 'assume σ known' is on, the lower panel thus plots *z* values. Contrast this with *t*, in equation 2, which measures how far *M* falls from μ_0, but in units of $s/\sqrt{n}$. The tricky thing is that these units are different for every sample, because *s* varies from sample to sample. Consider the latest sample, whose *M* and *t* are marked in Figure 1 by large symbols, and arrows. Because *s* for that sample happens to be small (9.5, much less than $\sigma = 16$), *t* is much larger than *z* would be for that sample. Now, consider that the sampling distribution of s^2 has the shape of the χ^2 distribution, which has strong positive skew, especially for low *df*. The sampling distribution of *s* thus also has positive skew, meaning that a majority of values are less than the mean, which is approximately σ. Therefore, many samples have small *s* and, correspondingly, large *t* values—and that is why central *t* has fat tails!

In print this argument seems tortuous, but working with the interactive simulation makes it much more vivid and convincing.

In Figure 1, sampling is from the distribution centred on μ_0 and t is measured from μ_0, so the distributions in the lower panels are symmetric. In Figure 2, by contrast, although t is still measured from μ_0, sampling is from the distribution centred on μ_A. If 'assume σ known' is on, once again the distributions of M in the centre panel and z in the lower panel are normal, and thus symmetric. However, if 'assume σ known' is off, t values calculated using Equation 2 are displayed, and each is strongly influenced by the s of its sample. This influence is especially strong because t is measuring the distance from μ_0, a considerable distance from μ_A. The distribution of t values shown lower right in Figure 2 thus has large variance. Moreover, because many s values are small, there are many large t values and the distribution is asymmetric and has a long fat tail to the right —that being the side opposite to μ_0. Once again, controlling the simulation, examining the points for individual samples, and seeing the curves grow as more samples are taken, all give a much clearer appreciation than this description in words can give.

THE BEAUTY OF NONCENTRAL *t*

About 20,000 cells in the Microsoft Excel spreadsheet are needed to calculate the noncentral t distribution displayed lower right in Figure 2. It is pleasing to watch the sampling process give a pile of t values that always, whatever parameter values are selected, approximately fits the displayed curve. Noncentral t arises naturally from simple sampling when the Alt is true—which in practice is very often.

One important feature of the simulation seems to be that sampling from the Alt distribution (Figure 2) is compared directly with sampling from the Null distribution (Figure 1). Another is that marking the lower axis in units of $\sigma/\sqrt{n}$ permits direct comparison of the normally distributed sampling distributions of M in the middle panel with the t distributions in the lower panel.

Noncentral t takes on a wide range of shapes, with the degree of asymmetry varying with both df and Δ, which in turn depends on the distance from μ_0 to μ_A. It approaches the normal distribution in shape, as n (and therefore df) increases, but it does so very slowly, as Cumming and Finch (2001) explained and illustrated. It is necessary for the calculation of power, and CIs on Cohen's d values. Noncentral t is an important distribution: The Alt think-tank that designed it did a fine job! Noncentral t deserves to be more widely appreciated, and my experience is that these pictures and the interactive simulation can help give that appreciation.

ACKNOWLEDGEMENT

I thank Fiona Fidler for comments on a draft.

REFERENCES

Cumming, G. and 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, 530–572.

Cumming, G. and Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist*, 60, 170-180.

Grissom, R. J. and Kim, J. J. (2005). *Effect Sizes for Research. A Broad Practical Approach*. Mahwah, NJ: Erlbaum.

Kline, R. B. (2004). *Beyond Significance Testing. Reforming Data Analysis Methods In Behavioral Research*. Washington, DC: American Psychological Association.