APPENDIX 10.4: PLANNING THE MARGIN OF ERROR

Specifying the Desired Margin of Error

We saw in Appendix 6.1 that researchers have a great deal of control over the *moe* in a confidence interval. When σ is known, the sample size required to obtain an *moe* of f * σ is $n = (z_{\alpha/2}/f)^2$. We use this approach to control precision (*moe*) when σ is not known but there are some slight complications.

We spent a long time explaining the difference between $z_{\alpha/2}$ and $t_{\alpha/2}$ in Chapter 10. Therefore, it should not be surprising that computing the sample size (n) required to obtain an *moe* of f * σ involves replacing $z_{\alpha/2}$ with $t_{\alpha/2}$ as follows:

$$n = \left(\frac{t_{\alpha/2}}{f}\right)^2. \quad (10.A4.1a)$$

That's the good news. The bad(ish) news is that $t_{\alpha/2}$ depends on *df*, which in turn depends on n. Equation 10.A4.1b makes this explicit by adding $n-1$ as a subscript to $t_{\alpha/2}$ as follows:

$$n = \left(\frac{t_{\alpha/2,(n-1)}}{f}\right)^2. \quad (10.A4.1b)$$

Therefore, thinking back to algebra, we recognize that we have n defined in terms of a $t_{\alpha/2}$ value that itself depends on n, making it impossible to solve for n with the usual tools of algebra.

What should we do? There are interesting methods that will allow us to calculate the exact sample size required to achieve our margin of error. Such methods take us beyond what is necessary for the purposes of this chapter. Fortunately, there is a simple approximation that works reasonably well if the required f is less than 1. (We noted in Appendix 6.1 that no useful f is greater than 1.) If we were to perform the exact calculation of n for values of f between .1 and 1, we would find that the required n is always just 2 or 3 greater than the n we'd get by using $n = (z_{\alpha/2}/f)^2$. Therefore, we can simply compute $n = (z_{\alpha/2}/f)^2$ and then add 3 to get a good approximation to the correct answer.

Planning a Study

Let's reconsider the case of the Prescott Pharmaceuticals Company (PPC) from Appendix 6.3 (available online at study.sagepub.com/gurnsey). PPC wished to estimate, with 95% confidence, the mean attentional focus (AF) in adolescents with attention deficit disorder (ADD). They also require the *moe* to be σ/10; i.e., $f = .1$. They offered $50,000.00 to someone to carry out this study, and your research showed that it would cost $95.00 per subject to conduct the analysis. The $95.00 covers all costs involved, including lab work, research assistant pay, transportation, and so forth. Would it be worth your while to accept this contract?

To figure out how much profit you would make if you were to accept this contract, you need to determine the number of participants required to achieve the required *moe*. You start by estimating the number that would be required if σ were known. As before, you find that you would need 385 participants. The approximate method that we just described leads you to add 3 to this total. Therefore, you estimate you will need 388 participants. (An exact method would tell you that you need 387 participants to achieve an *moe* of σ/10.) Therefore, you would make only $285.00 less this time around than you did when σ was known, because this time you would need three more participants at a cost of $95.00 each.

A Caveat

You may have noticed that we've ignored a complication that distinguishes planning when σ is known from planning when σ is not known. Specifically, we stated that our desired *moe* was σ/10 in the preceding example. However, Figure 10.3 shows that the *moe* is different from sample to sample. Therefore, how is it possible that the sample size we computed using the method described above guarantees an *moe* of σ/10? The short answer is that it doesn't. Rather, our calculations produce a sample size that will have an *moe* of σ/10 *on average*. Because *moe* differs from sample to sample, the *moe* associated with a given sample might be slightly larger or smaller than the *moe* we wished to obtain.

The solution to this problem is to find a sample size that is large enough that only a small proportion of *moe* values would be greater than the one we desire. In the example we've been considering, we might want to find a sample size (n) such that only 1% of *moe* values would be greater than $\sigma/10$. We won't discuss how to achieve this level of *assurance* at this point. There is a very good discussion of this in Cumming (2012) if you would like to know more.

Reference

Cumming, G. (2012). *Understanding the new statistics effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge, Taylor & Francis Group.

APPENDIX 10.5: PROOF THAT $m \pm t_{\alpha/2}(s_m)$ WILL CAPTURE μ $(1-\alpha)$100% OF THE TIME

The purpose of this appendix is to show that $m \pm t_{\alpha/2}(s_m)$ will capture μ $(1-\alpha)$100% of the time, no matter what the population mean and standard deviation are or what sample size is. The only assumptions here are that the samples are randomly drawn from the same normal distribution.

We saw in Chapter 10 that any distribution of means can be transformed to a t-distribution by applying the following transformation to each sample mean:

$$t = \frac{m-\mu}{s_m}.$$

We then stated that $(1-\alpha)$100% of all possible intervals computed as

$$m \pm t_{\alpha/2}(s_m)$$

will capture μ. So, let's see why this is true.

By definition we know that $(1-\alpha)$100% of the t-distribution lies in the interval $-t_{\alpha/2}$ to $t_{\alpha/2}$. We can express this as

$$\Pr(-t_{\alpha/2} < t < t_{\alpha/2}) = 1-\alpha. \qquad (10.A5.1)$$

As before, equation 10.A5.1 should be read as follows: *The probability that a randomly chosen t-score will fall in the interval* $\pm t_{\alpha/2}$ *is* $1-\alpha$.

We will now show that the inequality $(-t_{\alpha/2} < t < t_{\alpha/2})$ can be rearranged into a more familiar form without changing its meaning. The colored text will be retained to remind us that the following transformations do not change the meaning of what is inside the parentheses.

When we insert the definition of the t-score into equation 10.A5.1, we obtain the following:

$$\Pr\left(-t_{\alpha/2} < \frac{m-\mu}{s_m} < t_{\alpha/2}\right) = 1-\alpha. \qquad (10.A5.2)$$

Multiplying all three terms by s_m leaves us with this:

$$\Pr(-t_{\alpha/2}(s_m) < m-\mu < t_{\alpha/2}(s_m)) = 1-\alpha. \qquad (10.A5.3)$$

If we then subtract m from all three terms, we are left with the following:

$$\Pr(-t_{\alpha/2}(s_m) - m < -\mu < t_{\alpha/2}(s_m) - m) = 1-\alpha. \qquad (10.A5.4)$$

This transformation leaves $-\mu$ in the center of the inequality. If we now multiply all three terms by -1 (to make μ positive) and reorder the terms, we obtain

$$\Pr(m + t_{\alpha/2}(s_m) > \mu > m - t_{\alpha/2}(s_m)) = 1-\alpha. \qquad (10.A5.5)$$

Equation 10.A5.5 shows that μ is between $m + t_{\alpha/2}(s_m)$ and $m - t_{\alpha/2}(s_m)$ $(1-\alpha)$100% of the time. In other words, the probability is $1-\alpha$ that the interval $m \pm t_{\alpha/2}(s_m)$ will contain μ.

APPENDIX 10.6: G*POWER

There are many fabulous statistical tools that are freely available on the web. One of the best is G*Power, which can be downloaded from gpower.hhu.de/en.html. G*Power does many things, but our focus here will be on determining the sample size required to achieve a desired level of power for the one-sample t-test described in Chapter 10.

Using G*Power

When you launch G*Power, you will be presented with the dialog shown in Figure 10.A6.1. (The default settings will be different from those in Figure 10.A6.1.) I will step through these settings so that you can see how to do a prospective power analysis. There are five panels in Figure 10.A6.1 that we'll discuss in detail: Test family, Statistical test, Type of power analysis, Input parameters, and Output parameters. Before we begin, make sure that Central and noncentral distributions are selected in the top panel.

Test Family

Power is an issue for many experimental and non-experimental designs. Therefore, in the box titled Test Family, there is a drop-down list that will allow you to choose the type of test of interest. I've selected t tests from this list.

FIGURE 10.A6.1 ■ The G*Power Dialog

Settings of G*Power to determine the sample size to achieve the desired 80% power for a one-tailed, one-sample *t*-test when δ = .1 and α = .05.

Statistical Test

To the right of the Test Family drop-down list, there is a second drop-down list under the heading Statistical Test. As we'll see, there are many different *t*-tests we can perform. I've chosen Means: Difference from constant (one sample case), which describes the test we've been discussing.

Type of Power Analysis

To compute the required sample size for a given effect size, we must select the appropriate type of power analysis. For our purposes, the correct selection from the drop-down list is A priori: Compute required sample size - given α, power, and effect size.

Input Parameters

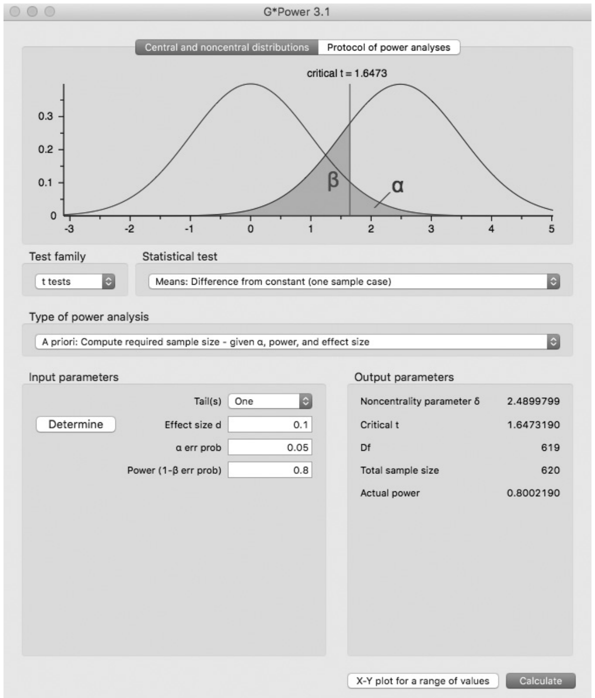
The region in the bottom left of Figure 10.A6.1 titled Input parameters allows you to specify the number of tails in the test, the effect size of interest, α, and power. You will see that I've chosen to compute the sample size required for a one-tailed test, with effect size (δ) = 0.1, α = .05, and power = .8. Once you've entered these, press the Calculate button and G*Power will begin searching for the required sample size.

Output Parameters

The results are shown in the bottom right of Figure 10.A6.2 in the region labeled Output parameters. The first output parameter is called the *non-centrality parameter*. In G*Power, the non-centrality parameter is denoted with the Greek symbol δ; *this is not Cohen's δ*. We will say a few things about the non-centrality parameter in the next section, but for now we'll ignore it. Critical t corresponds to what we called $t_{critical}$ in Chapter 10. The sample size required for the specified input parameters is 620, and the degrees of freedom is therefore 619. We were aiming to have power = .8; G*Power has found that when sample size is 620, the actual power is .8002.

So, that's pretty simple. You only need to know δ, α, and the power you would like to have (along with your alternative hypothesis) and G*Power delivers the required sample size instantly.

FIGURE 10.A6.2 ■ G*Power Results



G*Power has computed the sample size required to achieve 80% power for a one-tailed, one-sample *t*-test, when δ = .1 and α = .05.

**The Non-central *t*-Distribution

Let's now return to the non-centrality parameter, which G*Power denotes as δ, but I'll call *ncp*. To understand the *ncp*, we need to think about the distribution of our test statistic (t_{obs}) when the null hypothesis is true and when it is false. In Chapter 10, we introduced the *t*-statistic defined as

$$t = \frac{m - \mu_0}{s/\sqrt{n}}. \qquad (10.A6.1)$$

When *m* is drawn from a distribution with mean μ_0, then the result is a *central* t-*distribution* based $n-1$ degrees of freedom. The mean of this distribution is 0, and this is the reason we call it a central *t*-distribution: the distribution is symmetrical and centered on 0. In the top panels of Figure 10.A6.2, the distribution to the left is a central *t*-distribution with $n-1$ degrees of freedom. This is the sampling distribution of t_{obs} when the null hypothesis is true.

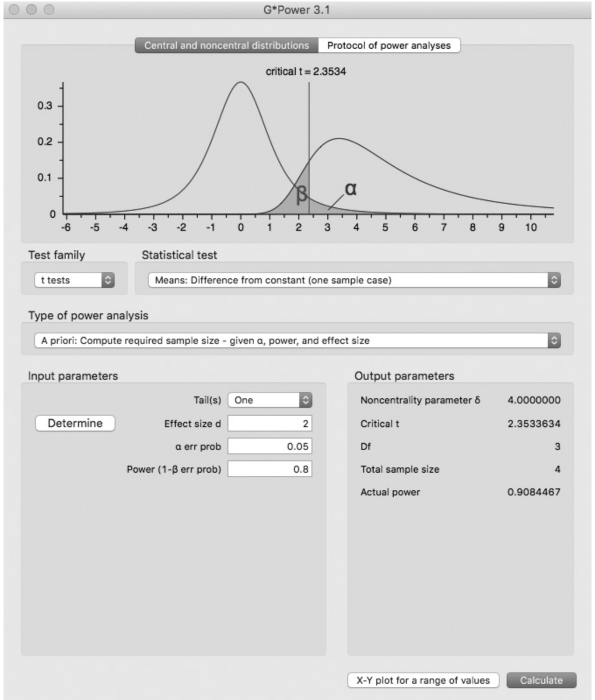
In prospective power analysis, we have to think about the distribution of the test statistic (t_{obs}) described in equation 10.A6.1 *when the alternative hypothesis is true*. This means computing the statistic described in equation 10.A6.1 when samples are actually drawn from a distribution having a mean μ_1. The resulting distribution is a *non-central* t-*distribution*, which is shown as the curve to the right in the top of Figure 10.A6.2. This is the density function we obtain when we form all samples of size *n* from the distribution of scores under the alternative hypothesis and compute the statistic shown in equation 10.A6.1. The proportion above the non-central *t*-distribution above $t_{critical}$ is power.

To appreciate what is happening here, remember that the samples were drawn from a distribution with mean μ_1 but the statistic was computed with $m - \mu_0$ in the numerator. The consequence is not a central *t*-distribution shifted along the *x*-axis but a skewed distribution shifted along the *x*-axis. The exact shape of a non-central *t*-distribution depends on the effect size (Cohen's δ) and *df*.

Note that μ_1 and μ_0 don't appear anywhere in Figure 10.A6.2 because they are unnecessary. The figure conveys the non-central *t*-distribution for all combinations of μ_1, μ_0, and σ for which $\delta = (\mu_1 - \mu_0)/\sigma$. These distributions have uses far beyond computing the power of a significance test. For example, when we compute exact confidence intervals for *d* in **R**, we are making use of non-central *t*-distributions.

Figure 10.A6.3 shows a second example of a non-central *t*-distribution for a case in which δ = 2. Here, the distribution shows much greater right skew and is

FIGURE 10.A6.3 ■ Non-central *t*-Distribution



G*Power has computed the sample size required to achieve 80% power for a one-tailed, one-sample *t*-test, when δ = 2 and α = .05.

shifted farther along the *x*-axis; compare the scales of the *x*-axes in Figures 10.A6.2 and 10.A6.3. Non-central *t*-distributions are defined by degrees of freedom (*df*) and the non-centrality parameter (*ncp*).

As a final note, in the situation we're considering, there is a very simple relationship between the non-centrality parameter (*ncp*), Cohen's δ, and sample size:

$$ncp = \delta\sqrt{n}. \qquad (10.A6.2)$$

(Remember, G*Power calls the non-centrality parameter δ.) In Figure 10.A6.2, we had effect size (δ) = .1 and $n = 620$. Therefore,

$$ncp = \delta\sqrt{n} = .1\sqrt{620} = 2.4899799,$$

as shown in Figure 10.A6.2. In Figure 10.A6.3, we had effect size (δ) = 2 and $n = 4$. Therefore,

$$ncp = \delta\sqrt{n} = 2\sqrt{4} = 4,$$

as shown in Figure 10.A6.3.