Power and Sample Size Calculations for the 2-Sample Z-Statistic

James H. Steiger

November 4, 2004

1 Topics for this Module

1. Reviewing Results for the 1-Sample Z
   (a) Power and Sample Size in Terms of a Noncentrality Parameter
2. The 2-Sample Independent Sample Z
   (a) Deriving the Distribution of Z
      i. The Mean
      ii. The Variance
   (b) Power Calculation
      i. The Noncentrality Parameter
   (c) Sample Size Calculation
3. The Generalized Independent Sample Z-Statistic
   (a) Deriving the Distribution of Z
      i. The Mean
      ii. The Variance
   (b) Power and Sample Size Calculation

2 Reviewing Results for the 1-Sample Z-Test

In our previous lecture, we found that power, P, can be calculated from the following equation.

$$P = \Phi\left(\sqrt{N}E_s - R\right)$$

where $E_s = (\mu - \mu_0)/\sigma$ is the "standardized effect size," and R is the rejection point for the Z-statistic. We also found that sample size, N, can be computed as

$$N = \text{ceiling}\left(\frac{\Phi^{-1}(P) + R}{E_s}\right)^2$$

2.1 Power in Terms of a Noncentrality Parameter

On the way to deriving the above equations for power and N, we also derived the mean and standard deviation of the Z-statistic, and found that they are $\mu_Z = \sqrt{N}E_s$ and $\sigma_Z = 1$, respectively. Suppose we define a *noncentrality parameter* δ as

$$\delta = \mu_Z = \sqrt{N}E_s$$

Then

$$P = \Phi(\delta - R)$$

Note, however, that by "collapsing" $\sqrt{N}E_s$ into δ in this way, we have "lost" N, and it is no longer clear that there is a simple equation for N that can be derived as a function of E_s and R.

There is, as we will see later, an advantage to defining a noncentrality parameter. It turns out that, for any Z test in a very general family, power is simply $\Phi(\delta - R) = \Phi(\mu_Z - R)$. In what follows, we will assume equal variances in all populations, because our goal is to use the results for Z-tests as an approximation of the corresponding results for t-tests

3 The 2-Sample, Independent Sample Z-Statistic

We will study the behavior of independent sample Z-tests. As you recall, the test statistic for the most basic 2-sample, independent sample Z-statistic when variances are equal is

$$Z = \frac{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}}{\sqrt{\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\sigma^2}}$$

Notice that there are many ways to write the above statistic. A key to being able to re-express the formula is to realize the following simple identity:

$$\frac{1}{N_1} + \frac{1}{N_2} = \frac{N_1 + N_2}{N_1 N_2}$$

So another way to write Z is

$$Z = \frac{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}}{\sqrt{\sigma^2}\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} = \frac{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}}{\sigma\sqrt{\frac{N_1+N_2}{N_1 N_2}}} = \sqrt{\frac{N_1 N_2}{N_1 + N_2}}\left(\frac{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}}{\sigma}\right)$$

3.1 Deriving the Distribution of Z

Once we have written Z in this form, it is easy to derive its mean and variance, using our standard results on linear combinations and linear transformations.

3.1.1 The Mean

Notice that the only random variables in the equation for Z are $\overline{X}_{\bullet 1}$ and $\overline{X}_{\bullet 2}$. Recall that the mean of $\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}$ must be $\mu_1 - \mu_2$, and multiplication comes straight through in the mean, so

$$\mu_Z = \sqrt{\frac{N_1 N_2}{N_1 + N_2}}\left(\frac{\mu_1 - \mu_2}{\sigma}\right) = \sqrt{\frac{N_1 N_2}{N_1 + N_2}}E_s$$

where in this case, E_s is the difference between the two means in standard deviation units, i.e.,

$$E_s = \frac{\mu_1 - \mu_2}{\sigma}$$

3.1.2 The Variance

We can quickly show that the Z-statistic has a variance of 1. Recall that multiplying a variable by a constant multiples its variance by the square of that constant. We will re-express the equation for the 2-sample Z as

$$Z = \sqrt{\frac{N_1 N_2}{N_1 + N_2}}\left(\frac{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}}{\sigma}\right) = \sqrt{\frac{N_1 N_2}{N_1 + N_2}}\left(\frac{1}{\sigma}\right)(\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2})$$

Recall from our linear combination theory that the variance of the linear combination $\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}$ is

$$\sigma^2_{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}} = \frac{\sigma^2}{N_1} + \frac{\sigma^2}{N_2} = \sigma^2 \left(\frac{1}{N_1} + \frac{1}{N_2} \right) = \sigma^2 \left(\frac{N_1 + N_2}{N_1 N_2} \right)$$

But the Z statistic multiplies $\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}$ by some constants, so the variance of Z must be multiplied by the square of these constants. Specifically

$$\begin{aligned} \sigma^2_Z &= \left[\sqrt{\frac{N_1 N_2}{N_1 + N_2}} \left(\frac{1}{\sigma} \right) \right]^2 \sigma^2_{\overline{X}_{\bullet_1} - \overline{X}_{\bullet 2}} \\ &= \left[\sqrt{\frac{N_1 N_2}{N_1 + N_2}} \left(\frac{1}{\sigma} \right) \right]^2 \sigma^2 \left(\frac{N_1 + N_2}{N_1 N_2} \right) \\ &= \frac{N_1 N_2}{N_1 + N_2} \frac{1}{\sigma^2} \sigma^2 \left(\frac{N_1 + N_2}{N_1 N_2} \right) = 1 \end{aligned}$$

In the final analysis, we have proven that the 2-sample Z-statistic has a mean of $\sqrt{\frac{N_1 N_2}{N_1 + N_2}} E_s$ and a standard deviation of 1. Notice that, if the sample sizes are equal, i.e., $N_1 = N_2 = N$, the mean of the Z-statistic reduces to

$$\mu_Z = \sqrt{\frac{N}{2}} E_s = \delta$$

Once again, we find that the noncentrality parameter is simply the mean of the distribution of the Z-statistic.

3.2 Calculating Power

Since the standard deviation of the Z-statistic is still 1, power is simply a function of how far μ_Z (or, if you prefer, δ) is from the rejection point R. So, once again, for a rejection point R, power (P) is

$$P = \Phi \left(\mu_Z - R \right)$$

3.2.1 Calculating Sample Size

Note, however, that by leaving N in the equation, we can also solve for the sample size required to assure a given power. Specifically, since

$$P = \Phi \left(\mu_Z - R \right) = \Phi \left(\sqrt{\frac{N}{2}} E_s - R \right)$$

we have, after taking Φ^{-1} of both sides of the equation and doing a bit of rearranging

$$\begin{aligned}
\Phi^{-1}(P) &= \sqrt{\frac{N}{2}}E_s - R \\
\frac{\Phi^{-1}(P) + R}{E_s} &= \sqrt{\frac{N}{2}} \\
2\left(\frac{\Phi^{-1}(P) + R}{E_s}\right)^2 &= N
\end{aligned}$$

To guaratee power at or above P, we use the ceiling function, and we have

$$N = \text{ceiling}\left[2\left(\frac{\Phi^{-1}(P) + R}{E_s}\right)^2\right]$$

3.2.2 An Example

Let's try an example. Suppose $E_s = .5$ and $N = 25$ *per group.* With $\alpha = .05$, and a 2-sided test, the rejection point R is 1.96. In this case, power is computed as

$$\begin{aligned}
P &= \Phi\left(\sqrt{\frac{N}{2}}E_s - R\right) \\
&= \Phi\left(\sqrt{\frac{25}{2}}.5 - 1.96\right) \\
&= \Phi\left(\sqrt{\frac{25}{2}}.5 - 1.96\right) \\
&= \Phi\left[(3.5355)(.5) - 1.96\right] \\
&= \Phi\left(-.1922\right) \\
&= .43
\end{aligned}$$

The area to the left of $-.1922$ in the standard normal distribution is about .43. Clearly, sample size is inadequate in this case. What sample size would we need to achieve a power of .80? First, we need $\Phi^{-1}(P) = \Phi^{-1}(.80)$. From the normal curve table, this is about .84. From the above equation, we get

$$\begin{aligned} N &= \text{ceiling}\left[2\left(\frac{\Phi^{-1}(P)+R}{E_s}\right)^2\right] \\ &= \text{ceiling}\left[2\left(\frac{.84+1.96}{.5}\right)^2\right] \\ &= \text{ceiling}(62.72) = 63 \end{aligned}$$

4 The Generalized Independent Sample Z-Test

Entire books have been written about t-tests that go beyond simple tests of equality for two means. So how does one generalize to this new situation? The answer is that you proceed in exactly the same way you did for the one sample and 2-sample t tests. You examine the distribution of the Z-statistic, obtain the mean and variance, and then write a formula for P and for N. Let's take a simple special case first. Suppose you wished to test the hypothesis that $\mu_1 - 2\mu_2 = 0$. Assuming equal variances, the Z-statistic for testing this hypothesis is

$$Z = \frac{\overline{X}_{\bullet_1} - 2\overline{X}_{\bullet 2}}{\sqrt{\left(\frac{1}{N_1}+\frac{4}{N_2}\right)\sigma^2}}$$

We can rewrite the above statistic as

$$Z = \sqrt{\frac{N_1 N_2}{N_1 + 4N_2}}\left(\frac{\overline{X}_{\bullet_1} - 2\overline{X}_{\bullet 2}}{\sigma}\right)$$

You can quickly derive, in the same manner as previously, that the standard deviation of the statistic is 1 and its mean is

$$\mu_Z = \sqrt{\frac{N_1 N_2}{N_1 + 4N_2}}\left(\frac{\mu_1 - 2\mu_2}{\sigma}\right)$$

If the sample sizes are both equal to N, this reduces to

$$\mu_Z = \delta = \sqrt{\frac{N}{5}}\left(\frac{\mu_1 - 2\mu_2}{\sigma}\right) = \sqrt{\frac{N}{5}}E_s$$

Note that the 2 in the denominator has been replaced by a 5. In the sections that follow, we present a general formula for the mean and variance of the Z-statistic

4.1 Deriving the Distribution of Z

Suppose the null hypothesis is

$$H_0 : \kappa = \sum_{j=1}^{J} c_j \mu_j = \kappa_0$$

The Z-statistic for testing this hypothesis, if variances are assumed equal, is

$$Z = \frac{\sum_{j=1}^{J} c_j \overline{X}_{\bullet j} - \kappa_0}{\sqrt{\left(\sum_{j=1}^{J} \frac{c_j^2}{N_j}\right) \sigma^2}}$$

With equal sample sizes, this can be written

$$\begin{aligned} Z &= \frac{\sum_{j=1}^{J} c_j \overline{X}_{\bullet j} - \kappa_0}{\sqrt{\left(\sum_{j=1}^{J} \frac{c_j^2}{N}\right) \sigma^2}} \\ &= \sqrt{\frac{N}{\sum c_j^2}} \frac{\sum_{j=1}^{J} c_j \overline{X}_{\bullet j} - \kappa_0}{\sigma} \end{aligned}$$

4.1.1 The Mean

The mean of Z can be derived immediately from the preceding formula, since each $\overline{X}_{\bullet j}$ has a mean of μ_j. The result is

$$\begin{aligned} \mu_Z &= \delta = \sqrt{\frac{N}{\sum c_j^2}} \frac{\sum_{j=1}^{J} c_j \mu_j - \kappa_0}{\sigma} \\ &= \sqrt{\frac{N}{\sum c_j^2}} E_s \end{aligned}$$

where the c_j are the linear weights used in the null hypothesis, and E_s is defined generally as the amount by which the null hypothesis is wrong in standard deviation units, i.e.,

$$E_s = \frac{\sum_{j=1}^{J} c_j \mu_j - \kappa_0}{\sigma}$$

.

4.1.2 The Variance

The variance standard deviation of the Z-statistic are always 1. See if you can prove this result for yourself.

4.2 Power and Sample Size Calculations

This allows us to write very general formulae for power and sample size when the sample sizes are equal to N. For power, we have

$$P = \Phi\left(\sqrt{\frac{N}{\sum c_j^2}} E_s - R\right)$$

For sample size, we manipulate the above equation

$$\begin{aligned} \Phi^{-1}(P) &= \sqrt{\frac{N}{\sum c_j^2}} E_s - R \\ \frac{\Phi^{-1}(P) + R}{E_s} &= \sqrt{\frac{N}{\sum c_j^2}} \end{aligned}$$

Squaring both sides and manipulating a bit more, we end up with

$$N = \text{ceiling}\left[\sum c_j^2 \left(\frac{\Phi^{-1}(P) + R}{E_s}\right)^2\right]$$

Once you become familiar with a couple of "key values" for the rejection point R and the power value $\Phi^{-1}(P)$, you can deduce power and sample size across a wide variety of situations. For example, suppose you are doing a 2-tailed test with $\alpha = .01$, and you need power of $P = .90$ to detect a standardized effect size of .50.. You need to estimate the sample size per group required to test the hypothesis

$$H_0 : \mu_1 - \mu_2 = \mu_3 - \mu_4$$

or, equivalently

$$\kappa = \mu_1 - \mu_2 - \mu_3 + \mu_4 = 0$$

In this case, the sum of squared linear weights is

$$\sum c_j^2 = (1)^2 + (-1)^2 + (-1)^2 + (1)^2 = 4$$

The power value $\Phi^{-1}(P)$ is $\Phi^{-1}(90) = 1.282$. The rejection point for the 2-sided test is 2.576. So the required sample size is

$$\begin{aligned} N &= \text{ceiling}\left[\sum c_j^2 \left(\frac{\Phi^{-1}(P) + R}{E_s}\right)^2\right] \\ &= \text{ceiling}\left[4\left(\frac{1.282 + 2.576}{.50}\right)^2\right] \\ &= \text{ceiling}[238.15] \\ &= 239 \end{aligned}$$