

Computation of Statistical Power dale.berger@cgu.edu

This document shows the computation of statistical power for a one-tailed z-test as demonstrated in the [WISE](#) Statistical Power Applet and tutorial.

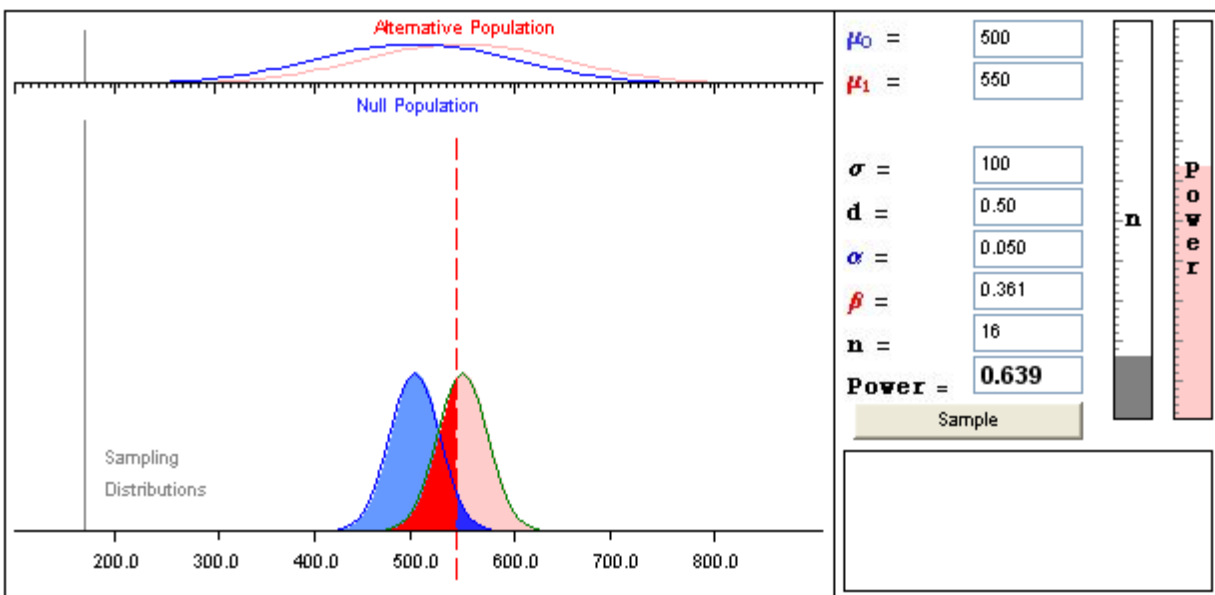
Consider an example where the null population mean is 500 and the alternate mean is 550, and the standard deviation is 100 for both populations. We wish to determine the statistical power for a one-tailed z-test using a sample of 16 cases, with $\alpha = .05$, one-tailed.

We can summarize this information as follows:

$$\mu_0 = 500, \mu_1 = 550, \sigma = 100, n = 16, \text{ and } \alpha = .05 \text{ one-tailed.}$$

With these values, the WISE Statistical Power Applet shows us that power = .639.

WISE Statistical Power Applet

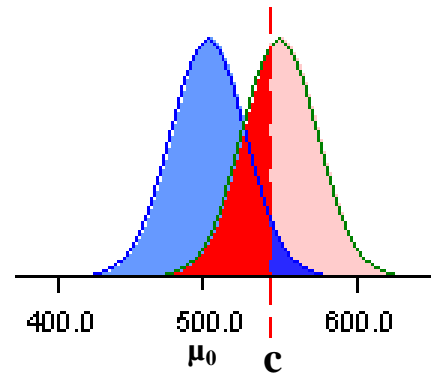


The sampling distributions shown in the applet represent the possible values of sample means based on samples of size n if the null hypothesis is true (blue distribution on the left) or if the alternate hypothesis is true (red/pink distribution on the right). The dashed red line represents the decision criterion: we reject the null hypothesis if we observe a sample mean greater than the criterion. Power is represented by the pink area, the probability that a sample mean will be greater than the criterion if we are sampling from the Alternate Population.

The following sections show how the power value is calculated. We first derive a general formula for power analysis, and then apply the formula to the scenario shown here to compute power. Next we adapt the formula to compute the sample size needed to attain a desired level of power, and to compute the minimum effect size needed to attain a desired level of power.

Derivation of a formula for power computations

The vertical dashed red line represents the critical value (**c**) for deciding whether or not to reject the null hypothesis. It is determined solely by our choice of acceptable alpha error. In the example, we set alpha at .05, one-tailed. This means we will reject the null hypothesis if we observe a sample mean that is in the largest 5% of possible means on the null sampling distribution, the dark blue area to the right of the red line. Because the sampling distributions are assumed to be normal in shape, we can find the Z score that corresponds to the top 5% of the blue distribution = 1.645. We will call this value Z_α because it is defined on the null distribution by alpha.



Here is a formula for Z_α :

$$Z_\alpha = \frac{(c - \mu_0)}{\sigma/\sqrt{N}} \quad \text{We can rewrite this equation to solve for } \mathbf{c}. \quad \mathbf{c} = Z_\alpha \sigma/\sqrt{N} + \mu_0$$

In our example, $c = (1.645)(100)/4 + 500 = 41.125 + 500 = 541.1$.

We will use Z_β as the label for the standardized score corresponding to **c** on the red distribution because the beta error is defined as the portion of the red curve that falls below **c**. The formula

for Z_β is $Z_\beta = \frac{(c - \mu_1)}{\sigma/\sqrt{N}}$ and the equation for **c** is $\mathbf{c} = Z_\beta \sigma/\sqrt{N} + \mu_1$

We can set the two equations for **c** equal to each other, giving

$$c = \mu_0 + Z_\alpha (\sigma/\sqrt{N}) = \mu_1 + Z_\beta (\sigma/\sqrt{N})$$

Rearranging terms gives $(Z_\alpha - Z_\beta) = \sqrt{N}(\mu_1 - \mu_0)/\sigma$ or $(Z_\alpha - Z_\beta) = \sqrt{N} d$

where d is a common measure of effect size: $d = (\mu_1 - \mu_0) / \sigma$.

The result is this elegant formula: $(Z_\alpha - Z_\beta) = \sqrt{N} d$ (Formula 1)

Using our BEAN

Formula 1 expresses the relationship between four concepts. If we know any three, we can compute the fourth. An easy way to remember these four concepts is with the mnemonic BEAN:

B = beta error, represented by Z_β , also called Type II error, is the probability of failing to reject a false null hypothesis. Power, the probability that we do reject a false null hypothesis, is $1 - \text{beta error}$.

E = effect size, represented by d , where $d = (\mu_1 - \mu_0) / \sigma$

A = alpha error, represented by Z_α , is set by the researcher, often at .05 or .01.

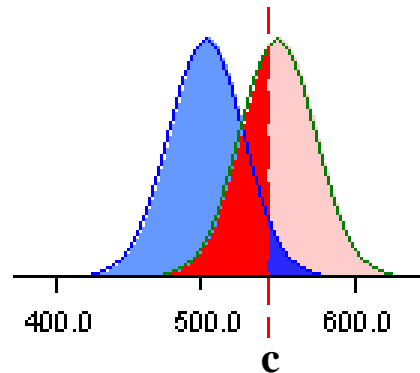
N = sample size, is usually determined by the researcher.

Power calculation

For power calculation, we adapt Formula 1

and solve for Z_β using the equation

$$Z_\beta = Z_\alpha - \sqrt{N} d \quad (\text{Formula 2})$$



In our example, we have an effect size $d = (\mu_1 - \mu_0) / \sigma = (550 - 500) / 100 = .500$.

We plan to use $N = 16$ (giving $\sqrt{16} = 4$), and one-tailed alpha = .05, which gives $Z_\alpha = 1.645$.

Applying Formula 2, $Z_\beta = 1.645 - 4.00 (0.500) = 1.645 - 2.000 = -0.355$.

Z_β corresponds to the Z score for the critical value on the red/pink (alternate hypothesis) sampling distribution. We will reject the null hypothesis if we observe a score greater than this critical value. We can use the [WISE p-z converter](#) or a Z table to find that the probability of observing a Z score greater than -0.355 is .639. This is the value shown for power in the WISE Power Applet.

WISE p-z Converter

right-tail p	z-score
<input type="text" value=".639"/>	<input type="text" value="-.355"/>
<input type="button" value="p --> z"/>	<input type="button" value="z --> p"/>
<input type="button" value="Graphic"/>	<input type="button" value="Help"/>

Sample size calculation

We can rearrange the terms in Formula 1 to solve for N.
$$N = \left[\frac{(Z_\alpha - Z_\beta)}{d} \right]^2 \quad (\text{Formula 3})$$

Suppose we wish to have a sample large enough to have power of 90% to detect a difference between means of 4 points when the standard deviation is 10, using one-tailed alpha error rate of .01. Recalling BEAN, we need to specify B, E, and A to solve for N.

B: Z_β corresponds to the Z score on the pink distribution where 90% of the distribution falls above that score. We can use the [WISE p-z converter](#) or a Z table to find $Z_\beta = -1.282$.

E: The effect size $d = (\mu_1 - \mu_0) / \sigma = (4/10) = 0.400$.

A: The alpha error rate of .01 corresponds to $Z_\alpha = 2.326$.

Applying Formula 3 to solve for N:

$$N = \left[\frac{Z_\alpha - Z_\beta}{d} \right]^2 = \left[\frac{2.326 - (-1.282)}{.400} \right]^2 = \left[\frac{3.608}{.400} \right]^2 = [9.02]^2 = 81.36$$

Thus, we need a sample size of about 82 to attain the desired level of power in this scenario.

It is important to be aware that the sign on Z_β is likely to be negative. Subtracting a negative value is equivalent to adding a positive value.

Computing the minimum effect size that can be detected with a specified level of statistical power

We can rearrange the terms in Formula 1 to solve for the effect size. Suppose we have only 25 cases available for study. What is the minimum effect size that would attain power = 80% using one-tailed alpha = .05?
$$d = \left[\frac{(Z_\alpha - Z_\beta)}{\sqrt{N}} \right] \quad (\text{Formula 4})$$

In this scenario, we wish to determine the effect size, so we need to specify B, A, and N.

$Z_\beta = -.842$, $Z_\alpha = 1.645$, and $N = 25$.

Applying Formula 4 we find $d = (1.645 - (-.842)) / 5.000 = .497$

This tells us that if the actual effect is smaller than about .50, then power for the contemplated study is less than 80%. We may decide that the study is not worth conducting if we believe that an effect size much smaller than .50 is important to detect and we wish to have at least an 80% chance to detect an effect size that small (e.g. $d=.40$).