Sampling Calculator Formulas

The following formulas are used for the sampling calculators on the ActivityInfo website.

Snapshot calculator

This calculator uses the following formula to find the required sample size:

$$n = \frac{m}{1 + \frac{m-1}{N}}$$

Where m is the sample size required for a large population, and N is the actual population size.

The required sample size for a large population is:

$$m = \frac{z_{\alpha/2}^2 \hat{p}(1 - \hat{p})}{\epsilon^2}$$

Where $\hat{p}$ is the expected proportion in the population, ϵ is the allowable margin of error, and $z_{\alpha/2}^2$ is the z-Score that corresponds to the 95% confidence level.

For a proof of this formula, see Penn State's excellent Introduction to mathematical statistics course.

Baseline and Endline Sample Calculator

The formula used by this calculator is based on the following equality from Wang (2007):

$$n = (Z_{\alpha/2} + Z_\beta)^2 \frac{f p_1(1 - p_1) + f p_2(1 - p_2)}{(p_1 - p_2)^2}$$

Where f is the finite population correction factor, which is:

$$f = \sqrt{\frac{N - n}{N - 1}}$$

Substituting and solving for n yields:

$$n = \frac{XA}{1 + XB}$$

Where:

$$X = \frac{(Z_{\alpha/2} + Z_\beta)^2}{(p_1 - p_2)^2}$$

$$A = \frac{N p_1(1 - p_1)}{N - 1} + \frac{N p_2(1 - p_2)}{N - 1}$$

$$B = \frac{p_1(1 - p_1)}{N - 1} + \frac{p_2(1 - p_2)}{N - 1}$$