

Is this an appropriate use of the margin of error equation in order to calculate the necessary sample size?

$$(1) \quad ME = z_{score} * \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

where,

ME is the desired margin of error

z_{score} is the z-score for the desired confidence level

\hat{p} is the best guess at the correct value of p

n is the sample size that we are calculating

Then in our particular example, if we assume that the desired margin of error is 0.002 at the 95% confidence level. And we also assume that 5% of the counts are in the tail, we would have the following equation for establishing the required sample size:

$$(2) \quad 0.002 = 1.96 * \sqrt{\frac{(0.05)(0.95)}{n}}$$

The sample size required to achieve the desired margin of error around \hat{p} is then 45,619. Taking it one step further, we would be confident at the 95% confidence level that the proportion of the total count found in the tail is between [0.0498, 0.0502].