Notation

N	total population size
n	sample size
$z_{1-\alpha/2}$	level of confidence according to the standard normal distribution
P	estimated proportion of the population
d	margin of error

Assuming a simple random sample, the estimated population proportion with a specified absolute precision is

$$n = z^2_{1-\alpha/2} P(1-P)/d^2_{srs} \quad \text{(Lwanga and Lemeshow 1991, p. 25).}$$

This is equivalent with

$$d_{srs} = z_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}}.$$

For a 95 % confidence level $z_{1-\alpha/2} \approx 1.96$ and for an unknown estimated proportion we use $P = 0.5$ where the term $P(1-P)$ reaches its maximum value. Thus, we get

$$d = 1.96\sqrt{\frac{0.5(1-0.5)}{n}} = \frac{0.98}{\sqrt{n}}.$$

Example

If $n = 100$ then $d_{srs} = 0.98/\sqrt{100} = 0.098$.

A sample size of 100 would have a 95 % confidence interval of ±9.8 % or less.

Margin of error in survey data

A sample that includes complex survey design such as different sampling probablilities or cluster sampling, should account for the design effect. The design effect

$$DEFF = var(survey)/var(srs),$$

where $var(survey)$ is the variation based on the survey design and $var(srs)$ is the variance assuming a simple random sampling.

If the sample includes a significant amount of the total population, a finite population correction can be applied. The formula for the correction is

$$FPC = \sqrt{\frac{N-n}{N-1}}.$$

A margin of error that accounts for design effect and finite population correction is

$$d_{survey} = z_{1-\alpha/2}\sqrt{\frac{P(1-P)}{n}} \times DEFF \times FPC$$

(Lehtonen and Pahkinen 2004).