Suppose the desired confidence level is 95%. Then Z values defining 95% confidence level are ±1.96, thus

$$(\overline{X}-\mu)=\pm 1.96\times\frac{\sigma}{\sqrt{n}} \quad \text{or} \quad E=\pm 1.96\times\frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \sqrt{n}=\pm\frac{1.96\times\sigma}{E}$$

$$n=\left(\frac{1.96\times\sigma}{E}\right)^2 \quad \text{or} \quad n=\left(\frac{Z\sigma}{E}\right)^2$$

Let's consider an example to explain this procedure. If a researcher wants to calculate the average income (in lac) of the given population having an accuracy of 0.5 of in income. This means that the researcher can tolerate a loss of 50% in income on either side of the true average income at 95% confidence level. Thus, it means that the researcher wants to be 95% confident about his decision.

The formula for confidence limits is: $(\overline{X}-\mu)=Z\frac{\sigma}{\sqrt{n}}$

Where,

μ	=	Population mean
$\overline{X}$	=	Average income calculated from the sample
Z	=	Value of z at 95% confidence level
$\frac{\sigma}{\sqrt{n}}$	=	Standard error of $\overline{x}$
σ	=	Standard deviation
n	=	Sample size

If the researcher has decided to tolerate an error of 1/2 lac that is $(\overline{X}-\mu)=0.5$

So $0.5=Z\frac{\sigma}{\sqrt{n}}$ value of Z at 95% confidence level are ±1.96

So $0.5=\pm 1.96\frac{\sigma}{\sqrt{n}}$ or $\sqrt{n}=\frac{1.96\times\sigma}{0.5}$

The value of standard deviation (σ) can be calculated by either;

1) Assuming or guessing,
2) Consulting by the expert,
3) By experimenting to get the value, and
4) Obtaining σ from other comparable studies.

Let σ = 3 lac

Then, $\sqrt{n}=\frac{1.96\times 3}{0.5}=11.76$

$n=(11.76)^2=138.3\approx 138$

Thus, 138 respondents are required to constitute a sample at 95% level of accuracy. Similarly, it can be calculated for 99% or others.

Example 1: the placement department of a leading management institution is planning a survey of annual earning of its MBA alumni numbering 2,000. How large a sample size it should take in order to estimate the mean annual earning within plus and minus ₹ 1,500 and at 95% confidence level? The standard deviation of annual earning of the entire population is known to be ₹ 2,500.

Solution: As the desired upper and lower limit is ₹ 1,500, i.e., we want to estimate the annual earning within ± 1,500.

$$\text{So } Z\times\frac{\sigma}{\sqrt{n}}=1500$$

As the level of confidence is 95%, the Z value is 1.96

$1.96\times\frac{\sigma}{\sqrt{n}}=1500;\quad \frac{\sigma}{\sqrt{n}}=\frac{1500}{1.96};\quad \frac{\sigma}{\sqrt{n}}=765.31;$

$\frac{2500}{\sqrt{n}}=765.31;\quad \sqrt{n}=\frac{2500}{765.31}=3.3$

So n = 10.89

Or n = 11.

Example 2: Lumber companies need to be able to estimate the amount of lumber that they can harvest in a tract of land to determine whether the effort will be profitable. To do so, they must estimate the mean diameter of the trees. It has been decided to estimate that parameter to be within one inch with 99% confidence. A forester familiar with the territory guesses that the diameters of the trees are normally distributed with a standard deviation of 6 inches. How large a sample should be taken?

(It is given that in a normally distributed population 99% area is covered within mean ± 2.57 standard deviations).

Solution: The forester determines the sample size as follows:

Error of Estimation (E) = 1

Confidence level is 99%, So Z = 1.96

The population standard deviation is assumed to be σ = 6.

Thus,

$$\text{Sample Size (n)}=\frac{Z^2\sigma^2}{E^2}=\frac{(1.96)^2(6)^2}{(1)^2}=\frac{36\times 3.8416}{1}=138.3\approx 138$$

Example 3: A simple random sample is to be taken from a population of 50,000 sales invoices to estimate the mean amount per invoice. The standard deviation of the population is 4,000. The allowable error is 200 and the confidence coefficient is 90% (Z = 1.64). What size of sample is appropriate?

Solution: Given, Z = 1.64, σ = 4000 and E = 200

$$\therefore\quad n=\frac{Z^2\sigma^2}{E^2}=\frac{(1.64)^2\times(4000)^2}{(200)^2}=2.6896\times 400=1075.84\approx 1076$$

Example 4: Ascertain the size of the sample from the following particulars:
Standard deviation of population $\sigma_p = 4$

Mean of population $\mu = 24$
Mean of sample or $\overline{X}_s = 22$ and
Level of confidence = 99%
(Z value at 99% = 2.5758)

Solution: Given, Standard deviation of population (σ_p) = 4, Mean of population (μ) = 24, Mean of sample ($\overline{X}_s$) = 22 and Z = 2.5758.

Since, $\sigma_{\bar{x}} = \dfrac{\sigma_p}{\sqrt{n}}$ $\Rightarrow \sigma_{\bar{x}} = \dfrac{4}{\sqrt{n}}$

$$\because \; Z = \frac{\overline{X}_s - \mu}{\sigma_{\bar{x}}} = \frac{22-24}{4/\sqrt{n}}$$

$$\Rightarrow 2.5758 = \frac{-2}{4/\sqrt{n}}; \qquad \Rightarrow 2.5758 \times 4 = -2\sqrt{n}$$

$$\Rightarrow \sqrt{n} = -5.1516; \qquad \Rightarrow n = 26.53 \approx 27$$

Determining Sample Size (n) When Estimating the Population Proportion

In the estimation of population proportion (p) sample proportion $\hat{p}$ is used for a sample of size n.

The standard error of $(\hat{p}) = \sqrt{\dfrac{pq}{n}}$ (p is known).

For large samples,

$$Z = \frac{\hat{p} - p}{S.E.(\hat{p})} = \frac{E}{\sqrt{\frac{pq}{n}}}$$

Where, $E = \hat{p} - p$ (sampling error)

$$\sqrt{n} = \frac{Z\sqrt{pq}}{E} \qquad \text{Or } n = \frac{Z^2 pq}{E^2}$$

Example 5: Suppose p is estimated as 60 % and the confidence level is set at 95%. If the allowable error in estimating the population proportion is not to be greater than 3%, calculate the required sample size.

Solution: At 95% confidence level Z = 1.96.

The sample size is given by $\sqrt{n} = \dfrac{Z\sqrt{pq}}{E}$

Here p = 0.6, q = 0.4, E = 0.03

$$\text{So } \sqrt{n} = \frac{1.96\sqrt{0.6 \times 0.4}}{0.03} = \frac{1.96\sqrt{0.24}}{0.03}$$

$$n = \frac{1.96 \times 1.96 \times 0.24}{0.0009} = \frac{0.921984}{0.0009} = 1024.43 \approx 1024$$