Derivation of the **Standard Error of the Mean (SEM)**

$$SEM = \frac{SD}{\sqrt{n}}$$

We have n individual values of some variable (property):

$$X_1, X_2, \ldots, X_n$$

The population mean (expected value) is μ.

This is also the expected value for each individual X_i.

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu$$

The population standard deviation is σ , hence the variance is σ^2.

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

In practice, we do not know υ, so we replace it with the sample mean.

This results in a bias of the estimate of the variance (because the sample mean provides the smallest possible estimate of the variance). This bias can be easily corrected by dividing by (n-1) instead of n. We know must use s^2 instead of σ^2 but we could also use $\widehat{\sigma^2}$. The hat indicates an estimate.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The variance of the mean is:

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)$$

According to the rule Var (c . X) = c^2 . Var (X), the 1 / n becomes 1 / n^2.

All X_i are identically distributed with Var (X_i) = σ^2.

Since all X_i are independent from one another (no covariances), the variance of the sum is the sum of these variances (n times the variance of one X_i)

$$Var(\bar{X}) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Thus, the variance of the mean is σ^2/ n.

Since we want the standard error (standard deviation of the mean), we take the square root of the variance, which is

$$\sigma_{\bar{X}} = \sqrt{Var(\bar{X})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

The population variance σ^2 is typically unknown, so we replace it with its estimate s^2.

The same formula (standard deviation divided by square root of sample size n) applies to proportions, too.

$$SE_p = \sqrt{\frac{p(1-p)}{n}}$$

$\sqrt{p(1-p)}$ is the standard deviation of the underlying binary variable (yes/no).

As we do not know the true proportion p, we replace it with the sample estimate.

In both cases (mean, proportion), we see that

- the bigger the standard deviation, the bigger the standard error (and by implication the bigger the confidence interval for the true mean and proportion, respectively).
- The bigger the sample size, the smaller the standard error and the confidence interval.

Smaller confidence intervals require larger samples. However, cutting the confidence interval in half requires 4 times as many subjects.