Assume the list of values:

$$x = \{x_1, x_2, x_3, \dots, x_K\},$$

are independent data drawn from a distribution with a theoretical mean μ and theoretical variance σ^2. The sample mean of x is,

$$\bar{x} = \frac{1}{K} \sum_{k=1}^{K} x_k \,.$$

Since each of the values x_k are samples of random variables with a probability distribution, so is the sample mean. We are interested in computing the properties of the distribution of the sample mean of our original data.

First, we compute the expected value, or theoretical mean, of the sample mean,

$$E\big[\bar{x}\big] = \frac{1}{K} \sum_{k=1}^{K} E\big[x_k\big] = \frac{K\mu}{K} = \mu \,.$$

In other words, the sample mean has the same expected value, or theoretical mean, as the original data. This is why we often use the sample mean to estimate the theoretical mean of the data.

Next, we compute the theoretical variance of the sample mean of the data. The variance is the squared deviation of the sample mean of the data from the theoretical mean,

$$\begin{aligned}
\mathrm{Var}\Big(\bar{x}\Big) &= E\Big[\Big(\frac{1}{K} \sum_{k=1}^{K} x_k - \mu\Big)^2\Big] \,, \\
&= E\Big[\Big(\frac{1}{K} \sum_{k=1}^{K} \big(x_k - \mu\big)\Big)^2\Big] \,, \\
&= \frac{1}{K^2}\, E\Big[\Big(\sum_{k=1}^{K} \big(x_k - \mu\big)\Big)^2\Big] \,.
\end{aligned}$$

The expected value in the last line of the expression above is the theoretical variance of the sum of zero-mean random variables $(x_k - \mu)$. Assuming that the variables x_k are independent, the variance of the sum becomes the sum of the variances, so,

$$\begin{aligned}
\mathrm{Var}\left(\bar{x}\right) &= \frac{1}{K^2}\sum_{k=1}^{K} E\left[\left(x_k-\mu\right)^2\right], \\
&= \frac{1}{K^2}\left(\mathrm{Var}\left[x_1\right]+\mathrm{Var}\left[x_2\right]+\mathrm{Var}\left[x_3\right]+\ldots\mathrm{Var}\left[x_K\right]\right), \\
&= \frac{1}{K^2}\left(\sigma^2+\sigma^2+\sigma^2+\ldots+\sigma^2\right), \\
&= \frac{1}{K^2}\left(K\sigma^2\right), \\
&= \frac{1}{K}\sigma^2 .
\end{aligned}$$

where, we use the upper-case Var to refer to the theoretical variance of a random variable; we will use the lower-case var to refer to the sample variance computed from the data. The last expression defines the variance of the sample mean of x in terms of the variance of each observation of x. To know σ^2, we'd need information about the distribution from which each x_k was drawn. We usually do not have this information, so let's instead assume that each x_k is drawn from the same distribution, and approximate σ^2 as the sample variance of x. The advantage of this approach is that we can then estimate σ^2 from the observed data, i.e. $\hat{\sigma}^2 = \mathrm{var}\left[x\right]$. Then,

$$\mathrm{Var}\left(\bar{x}\right) \approx \frac{1}{K}\mathrm{var}\left[x\right].$$

In words, to find the theoretical variance of the sample mean of x, we first compute the sample variance of x, and then divide this sample variance by the number of observations K.

The sample standard deviation of the sample mean of x, which is also called the **standard error of the mean** (sem), is then,

$$\begin{aligned}
\mathrm{sem}\left[x\right] &= \sqrt{\mathrm{Var}\left[\bar{x}\right]} \\
&\approx \sqrt{\frac{1}{K}\mathrm{var}\left[x\right]} \\
&= \frac{\mathrm{std}\left[x\right]}{\sqrt{K}},
\end{aligned}$$

where $\mathrm{std}\left[x\right]$ is the sample standard deviation of x.