Ref: Wikipedia "Bessel's Correction"

If have a set of measurements \( x_1, x_2, x_3, \ldots, x_N \),
then can construct a sample dist'n with mean
\[
\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i
\]
& std. dev.
\[
\sigma^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2
\]

The parent dist'n has mean \( \mu = \lim_{N \to \infty} \left( \frac{1}{N} \sum_{i=1}^{N} x_i \right) \)
& std. dev.
\[
\sigma^2 = \lim_{N \to \infty} \left[ \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2 \right]
\]

Goal is to try to understand why \( s^2 \) has factor \( \frac{1}{N-1} \)
instead of \( \frac{1}{N} \) as in \( \sigma^2 \).

Note: when \( N \) is large, the difference between
\[
\frac{1}{\sqrt{N}} \quad \text{&} \quad \frac{1}{\sqrt{N-1}}
\]
is negligible.
First, one can intuitively understand that, if one was to estimate $\sigma^2$ using $\frac{1}{N} \sum (x_i - \bar{x})^2$, the result would be too small.

The $x_i$ measurements, by construction, are closer to $\bar{x}$ than they are to $\mu$ (mean of parent distribution). Thus, the average $(x_i - \bar{x})^2$ will be a little too small.

Dividing by $N-1$ instead of $N$ compensates for this effect. To see why the correction factor is

$$\frac{N}{N-1}$$

requires a little more work.
Before beginning, establish some notation.

$E(...)$ is the "expectation value" of ...

For example, $E(x_i) = \mu$

We will make use of $E((x_i - \mu)^2) = \text{Var}(x_i)$

where $\text{Var}(x_i)$ is the variance of $x_i$.

$\text{Var}(x_i) = \sigma^2$.

In addition, $E((\bar{x} - \mu)^2) = \text{Var}(\bar{x})$

$\text{Var}(\bar{x})$ is variance of $\bar{x}$

$\text{Var}(\bar{x}) = \frac{\sigma^2}{N}$

This result (discussed on next page) is consistent with the expectation that our determination of $\bar{x}$ should improve as $N$ increases.
Recall from PHYS 231, we demonstrated that if we take 1000 meas of a quantity & find $\bar{x}$ & $s$ & then we take 10,000 meas of the same quantity we find some $\bar{x}$ & $s$. i.e. $\bar{x}$ & $s$ of a distribution determined by experimental setup & not by the no. of measurements.

However, expect our estimate of $\bar{x}$ to improve as $N$ increases. We demonstrated that if we repeat $N$ meas. of $x$ many many times that the dist. of the determined $\bar{x}$'s has width $\sqrt{\frac{\sigma^2}{N}} \Rightarrow \text{Var}(\bar{x}) = \frac{\sigma^2}{N}$

Now, let's determine $E\left(\sum [x_i - \bar{x}]^2\right)$

$= E\left(\sum [(x_i - \mu) - (\bar{x} - \mu)]^2\right)$ (add & subtract $\mu$)

$= E\left(\sum [(x_i - \mu)^2 - 2(x_i - \mu)(\bar{x} - \mu) + (\bar{x} - \mu)^2]\right)$

$= E\left(\sum (x_i - \mu)^2 - 2(\bar{x} - \mu)\sum (x_i - \mu) + (\bar{x} - \mu)^2 \frac{\sum 1}{N}\right)$
\[ E(\sum [x_i - \bar{x}]^2) = E \left( \sum (x_i - \mu)^2 - 2(\bar{x} - \mu) \left[ \sum x_i - \mu \sum 1 \right] + N(\bar{x} - \mu)^2 \right) \]

= \[ E \left( \sum (x_i - \mu)^2 - 2(\bar{x} - \mu) N (\bar{x} - \mu) + N(\bar{x} - \mu)^2 \right) - N(\bar{x} - \mu)^2 \]

= \[ E \left( \sum (x_i - \mu)^2 - N(\bar{x} - \mu)^2 \right) \]

= \[ E \left( \sum (x_i - \mu)^2 \right) - N E \left( (\bar{x} - \mu)^2 \right) \]

= \[ \sum E(\sum (x_i - \mu)^2) - N \sum E((\bar{x} - \mu)^2) \]

\[ \text{Var}(X_i) \quad \text{Var}(\bar{x}) \]

= \[ \sigma^2 \quad \frac{\sigma^2}{N} \]

\[ E \left( \sum (x_i - \bar{x})^2 \right) = \sum \sigma^2 - \sigma^2 = N\sigma^2 - \sigma^2 = \sigma^2 (N-1) \]
\[ E \left( \frac{1}{N} \sum (x_i - \bar{x})^2 \right) = \frac{N-1}{N} \sigma^2 \]

This expression underestimates \( \sigma^2 \) by a factor of \( \frac{N-1}{N} \).

However,

\[ E \left( \frac{1}{N-1} \sum (x_i - \bar{x})^2 \right) = \frac{N-1}{N-1} \sigma^2 = \sigma^2 \]

\( S^2 \) is a good estimator of \( \sigma^2 \).

The standard deviation of sample distribution given by

\[ S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2 \]

is a good estimator of the standard deviation of parent dist'n \( \sigma^2 \).