And we can also make an estimate of the variance, using the formula

\[ \hat{\sigma}^2 = \frac{X_1^2 + X_2^2 + \ldots + X_n^2}{n - 1} - \left( \frac{X_1 + X_2 + \ldots + X_n}{n}\right)^2 \]

(8.20)

The \((n = 1)\)'s in this formula look like typographic errors; it seems they should be \(n\)'s, as in (8.19), because the true variance \(\sigma^2\) is defined by expected values in (8.15). Yet we get a better estimate with \(n - 1\) instead of \(n\) here, because definition (8.20) implies that

\[ E(\hat{\sigma}^2) = \sigma^2. \]

Here's why:

\[
E(\hat{\sigma}^2) = \frac{1}{n - 1} \left[ \sum_{k=1}^{n} \sum_{j=1}^{n} \left( X_k - \frac{1}{n} \sum_{k=1}^{n} X_k \right)^2 \right] \\
= \frac{1}{n - 1} \left[ \sum_{k=1}^{n} \left( \frac{1}{n} \sum_{j=1}^{n} X_j \right)^2 - \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} (E(X_i) - E(X_j))^2 \right] \\
= \frac{1}{n - 1} \left[ \frac{1}{n} \left( nE(X^2) - n \right) \right] \\
= E(X^2) - \frac{1}{n} \left( nE(X^2) + n(n - 1)E(X)^2 \right) \\
= E(X^2) - \frac{1}{n} \left( nE(X^2) + n(n - 1)E(X)^2 \right) \\
= E(X^2) - \frac{1}{n} \left( nE(X^2) + n(n - 1)E(X)^2 \right) \\
= E(X^2) - \sigma^2
\]

(This derivation uses the independence of the observations when it replaces \(E(X_iX_k)\) by \((EX)^2[j = k] + E(X^2)[j = k]\).)

In practice, experimental results about a random variable \(X\) are usually obtained by calculating a sample mean \(\hat{\mu} = \bar{X}\) and a sample standard deviation \(\hat{\sigma} = \sqrt{\hat{\sigma}^2}\), and presenting the answer in the form \(\hat{\mu} \pm \hat{\sigma}/\sqrt{n}\). For example, here are ten rolls of two supposedly fair dice:

\[
\begin{array}{ccccccc}
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\| & \| & \| & \| & \| & \| & \| \\
\end{array}
\]

The sample mean of the spot sum \(S\) is

\[ \hat{\mu} = (7 + 11 + 8 + 5 + 4 + 6 + 10 + 8 + 8 + 7)/10 = 7.4; \]

the sample variance is

\[ (7^2 + 11^2 + 8^2 + 5^2 + 4^2 + 6^2 + 10^2 + 8^2 + 8^2 + 7^2 - 10\hat{\mu}^2)/9 \approx 2.1^2 \]