Some Basic Statistical Measures

- **Individual Measurement:**

  The variable x_i is used to represent the i^{th} measurement in a series of N measurements of the quantity x, where $i = 1, 2, \ldots, N$.

- **Mean:**

$$\bar{x} \;=\; \langle x \rangle \;=\; \frac{1}{N}\,(x_1 + x_2 + \cdots + x_N) \;=\; \frac{1}{N}\sum_{i=1}^{N} x_i \tag{1}$$

  This is usually used to represent the "average" measurement value. Two other possibilities are the *median* and the *mode*, which have different definitions.

- **Individual Deviation from the Mean:**

  The deviation of one measurement is $\delta_{x_i} = x_i - \langle x \rangle$. This quantity can be positive or negative.

- **Sample Variance = Mean Square Deviation:**

$${\sigma_x}^2 \;=\; \frac{1}{N-1}\sum_{i=1}^{N} {\delta_{x_i}}^2 \;=\; \frac{1}{N-1}\sum_{i=1}^{N} (x_i - \langle x \rangle)^2 \;=\; \left(\frac{N}{N-1}\right)\left[\,\langle x^2 \rangle - \langle x \rangle^2\,\right] \tag{2}$$

  This has 1 degree of freedom less than $\langle x \rangle$, hence the $N-1$ factor. The last identity allows for single-pass computation and is related to the *parallel axis theorem* of mechanics.

- **Standard Deviation = Root Mean Square (RMS) Deviation:**

$$\sigma_x \;=\; \sqrt{{\sigma_x}^2} \;=\; \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \langle x \rangle)^2} \;=\; \sqrt{\left(\frac{N}{N-1}\right)[\,\langle x^2 \rangle - \langle x \rangle^2\,]} \tag{3}$$

  This is the amount of "scatter" in a set of measurements, or the amount of "noise" associated with a single measurement. It does not change as the number of measurements is increased.

- **Standard Deviation of the Mean (SDM):**

$$\sigma_{\langle x \rangle} \;=\; \frac{\sigma_x}{\sqrt{N}} \;=\; \sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N} (x_i - \langle x \rangle)^2} \;=\; \sqrt{\left(\frac{1}{N-1}\right)[\,\langle x^2 \rangle - \langle x \rangle^2\,]} \tag{4}$$

  Also known as the **error in the mean** or **standard error**, this represents the uncertainty ("precision") of the average value of a set of measurements. For purely random errors, it decreases as the number of measurements increases, leading to a more certain result. Measured results are commonly reported in the form $\langle x \rangle \pm \sigma_{\langle x \rangle}$.

- **Signal-to-Noise Ratio (S/N = SNR):**

$$S/N \;=\; \langle x \rangle \,/\, \sigma_{\langle x \rangle} \tag{5}$$

  This characterizes the *significance* of a result. $S/N \geq 3$ is a minimum threshhold of reliability, with $S/N \geq 5$ preferrred. In Gaussian statistics, $S/N = 3$ equates to a 99.7% *confidence level*, i.e., the real quantity x has a 99.7% probability of lying between $\langle x \rangle - 3\,\sigma_{\langle x \rangle}$ and $\langle x \rangle + 3\,\sigma_{\langle x \rangle}$, while $S/N = 5$ equates to 99.99994% confidence. For a single measurement, $S/N = x_i/\sigma_x$.