Random errors in physical measurements

Random errors are deviations in the observation which yield results which differ from experiment to experiment. We characterize these deviations by their precision. The precision of a series of measurements thus reflects how closely each measurement is in a series and if they agree with others.

Usually, precision is influenced by physical effects beyond the control of the operator. In other cases, precision is influenced by how consistent someone is in performing a technique. For example, the precision of pipetting a volume of a liquid generally improves with practice.

We can define these random errors as a part of measurements that varies from measurement to measurement. Often this is referred to as noise.The only way to characterize random noise is to repeat a measurement many times, collecting the results in a table and then summarizing them and noticing if any patterns that have developed.

Example:

We shall consider the table of 10 successive measurements of the volume of the pipette which is labeled as 20 ml:

$$a := \begin{bmatrix} 19.988 & 19.986 & 19.980 & 19.973 & 19.982 & 19.975 & 19.986 \end{bmatrix}$$

A dot diagram showing the scatter of these values is presented in the following graph. From this graph, we can readily see how the measurements are distributed. In particular, there is a general location of these observations (in this example we can see that the results are clustered near the value of 19.982 rather than, say 19.3 or 20.3). Also, we note that the observations are spread about the general location in a specific fashion (they extend over about .025 mL).

$$a1 := \begin{bmatrix} 0.1 & 0.1 & 0.1 & 0.2 & 0.3 & 0.4 & 0.1 \end{bmatrix}$$

$$B := \text{join mat rows}\left(\text{sort}(a, \text{"a"}), a1 \right)$$
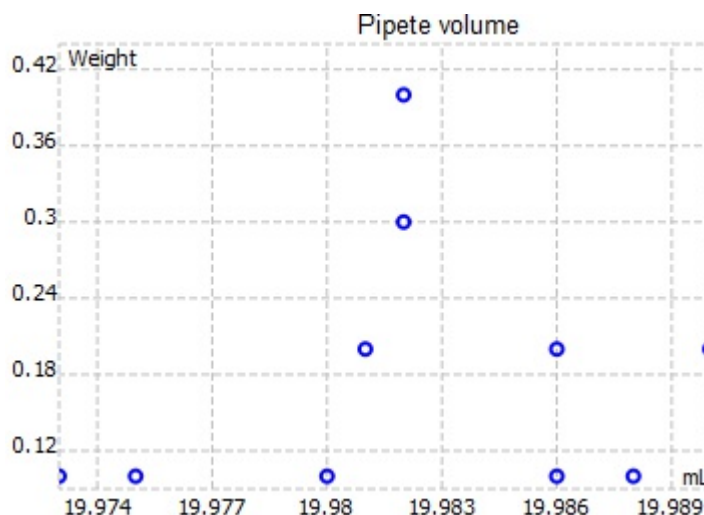
When a larger number of results is available, the dots become harder to distinguish from each other, and we are better able to appreciate the data by constructing a histogram of the results. In the table below we have collected 50 measurements of the same pipette.

| Trial	Vol (mL)	Trial	Vol (mL)	Trial	Vol (mL)
1	19.988	18	19.975	35	19.976
2	19.973	19	19.980	36	19.990
3	19.986	20	19.994	37	19.998
4	19.980	21	19.992	38	19.971
5	19.975	22	19.984	39	19.986
6	19.982	23	19.981	40	19.978
7	19.986	24	19.987	41	19.986
8	19.982	25	19.978	42	19.982
9	19.981	26	19.983	43	19.977
10	19.990	27	19.982	44	19.977
11	19.980	28	19.991	45	19.986
12	19.989	29	19.981	46	19.978
13	19.978	30	19.969	47	19.983
14	19.971	31	19.985	48	19.980
15	19.982	32	19.977	49	19.983
16	19.983	33	19.976	50	19.979
17	19.988	34	19.983		

Table I: Replicate data on the calibration of a 20 ml pipet

In this table the data has been rearranged into groups which we record the number of observations that fall into a series of contiguous .003 mL bins. Figure 4 plots the data of Table II.

Volume Range (mL)	Number in Range	% in Range
19.969 to 19.971	3	6
19.972 to 19.974	1	2
19.975 to 19.977	7	14
19.978 to 19.980	9	18
19.981 to 19.983	13	26
19.984 to 19.986	7	14
19.987 to 19.989	5	10
19.990 to 19.992	4	8
19.993 to 19.995	1	2

Table II: Frequency distribution of data from Table I

$$b := \begin{bmatrix} 19.970 & 19.973 & 19.976 & 19.979 & 19.982 & 19.985 & 19.988 \\ 6 & 2 & 14 & 18 & 26 & 14 & 10 \end{bmatrix}$$

Distribution of the 50 measurements
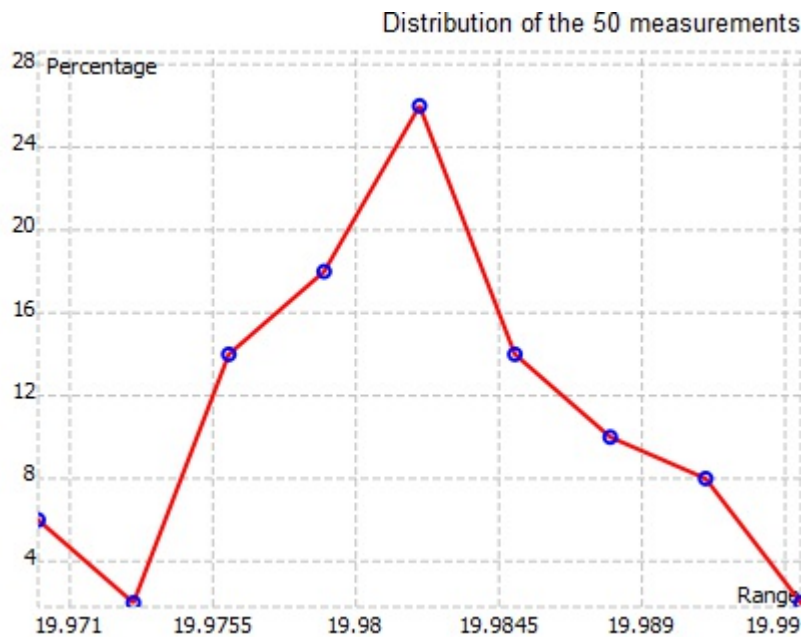
We notice that as the number of measurements grows larger, the graph becomes more and more certain. .

The mean (or average) of the measurements is

$$data := excel\ read\left("Measur\_data.xlsx"\ ,\ "Sheet1"\ ,\ "A1:A50"\ ,\ false\right)$$

$$X_m := average\left(data\right)$$

$$X_m = 19.982$$

There are several ways to express the degree of precision in a series of measurements. We will calculate the standard deviation as the measure of precision. The standard deviation is a measure of the size of the random errors in a set of repeated measurements.

$$S_{data} := devsam\left(data\right)$$

$$S_{data} = 0.006$$

The value of the standard deviation is used to indicate the amount of uncertainty in one's experimental values.

For many types of measurements, the distribution curves are found to be satisfactorily described by the Gaussian function:

$$P(x) = \left(\frac{1}{\sigma \cdot \sqrt[2]{2 \cdot \pi}} \right) \cdot e^{-0.5 \cdot \left(\frac{x - \mu}{\sigma} \right)^2}$$

where σ is the standard deviation and μ is the mean.

$$z := curve2d \left(normaldens \left(x, X_m, S_{data} \right), X_m - 3 \, S_{data}, X_m + 3 \cdot S_{data}, 1000 \right)$$

Normal (Gaussian) distribution
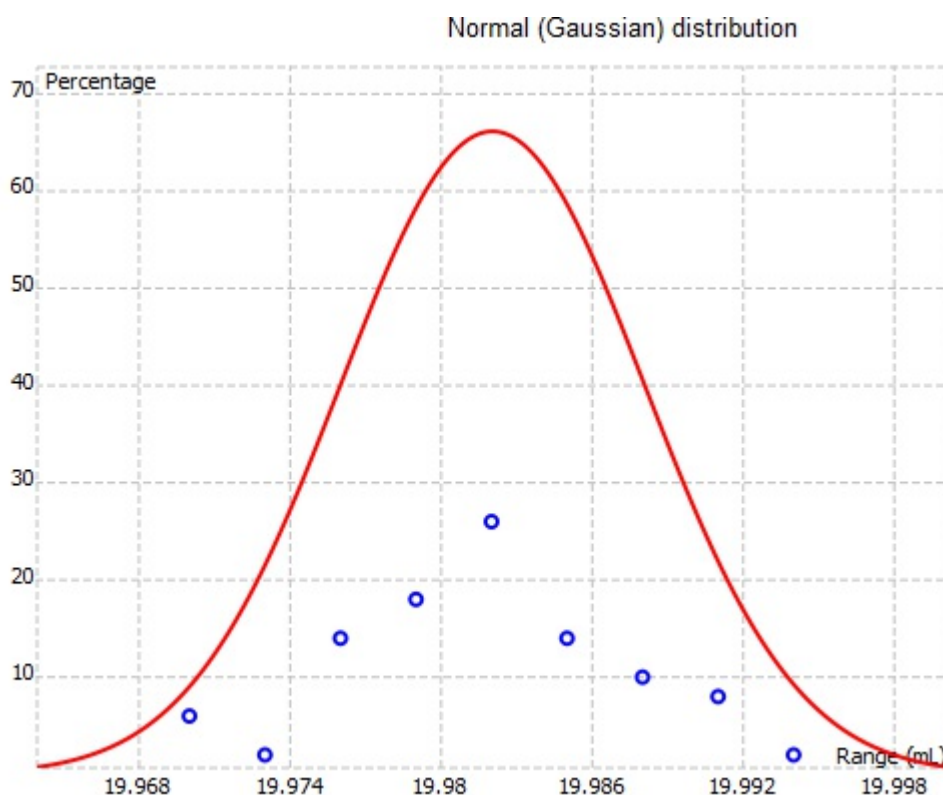
In the graph above we have a overlap of a Gaussian function of μ = 19.982 mL (X_m) and σ = 0.006 mL (S_{data}) and the distribution of the 50 measurements.

In the limit where the number of measurements become infinite, we have a sample of all possible values of the measurement. At this point we realize that the graph represents a probability distribution function, which specifies the probability of obtaining a measurement of a specific value as a function of that value. Probability distribution functions are described in terms of two parameters of the graph of measurements: the mean and standard deviation.