Propagation of Statistical Errors

Ajay K. Prasad
Associate Professor
Department of Mechanical Engineering
University of Delaware

Let us consider a simple experiment in which the quantity to be measured M, depends on only one independent variable x. For example, M could represent the vertical velocity of a rocket as a function of height from the ground, represented by x.

$$M = f(x).$$

From the figure, it is seen that the speed of the rocket is a non-linear function of x; the rocket moves slowly at first but then attains higher and higher velocities as it continues upward.

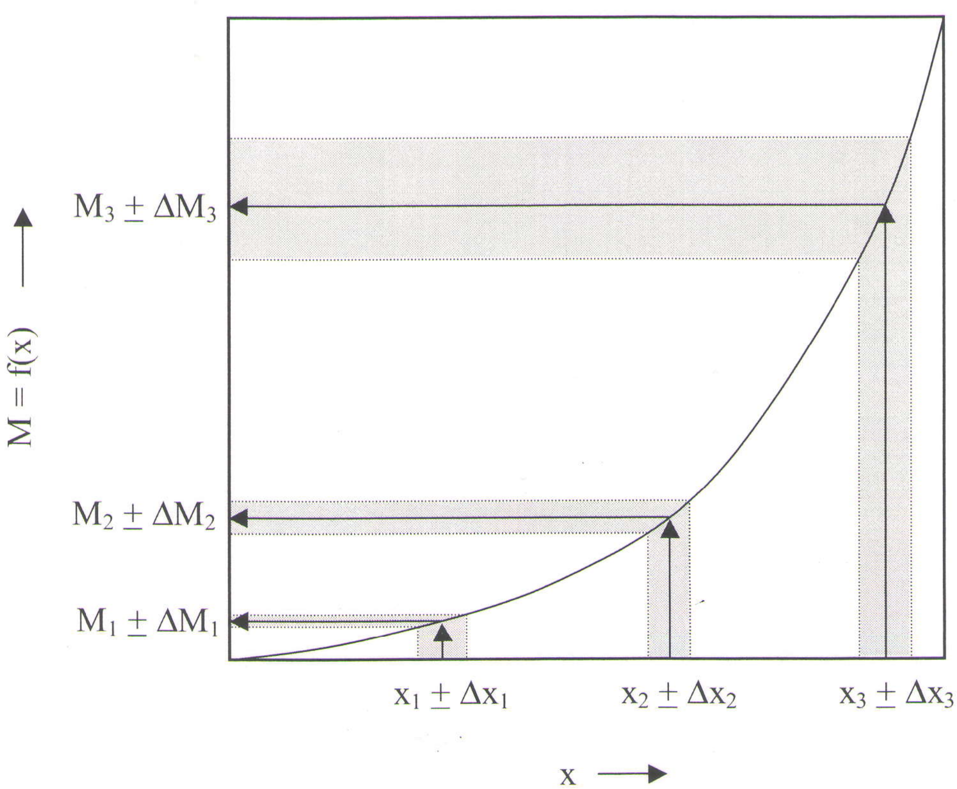

Figure 1: Dependence of error on the slope of the curve

Suppose we wanted to measure the velocity at three x locations, x_1, x_2, x_3. We will call these measurements M_1, M_2, M_3. Let the error in locating x be Δx. Typically, $\Delta x_1 \approx \Delta x_2 \approx \Delta x_3$. But it is very easy to see from the figure that $\Delta M_1 \neq \Delta M_2 \neq \Delta_3$. In fact, it is obvious that $\Delta M_3 > \Delta M_2 > \Delta M_1$.

It is also fairly obvious that the variation of ΔM is governed by the *slope* of the M vs. x curve. When the curve is flat, ΔM is small, i.e., in this region, M is not very sensitive to errors in x. At the other end of the curve where the slope is steep, M is very sensitive to errors in x. Because the slope of the curve is critical to the way the errors relate to one another, we can write:

$$
\begin{aligned}
\Delta M_1 &= \left.\frac{\partial f}{\partial x}\right|_1 \Delta x_1 \\
\Delta M_2 &= \left.\frac{\partial f}{\partial x}\right|_2 \Delta x_2 \\
\Delta M_3 &= \left.\frac{\partial f}{\partial x}\right|_3 \Delta x_3
\end{aligned}
$$

Or, more generally,

$$\Delta M = \frac{\partial f}{\partial x}\Delta x \tag{1}$$

Now, let us consider a more realistic situation where the quantity to be measured depends on more than one independent variables:

$$M = f(x, y, z, \cdots).$$

Extending result (1) to multiple variables, we can write:

$$
\begin{aligned}
\Delta M_x &= \frac{\partial f}{\partial x}\Delta x \\
\Delta M_y &= \frac{\partial f}{\partial y}\Delta y \\
\Delta M_z &= \frac{\partial f}{\partial z}\Delta z
\end{aligned}
$$

and so on.

At this point, we have determined how much error is caused in the measured quantity by errors in each of the contributing variables, but how do we combine them? Is it correct to say

$$\Delta M = \Delta M_x + \Delta M_y + \Delta M_z + \cdots ?$$

This would certainly not work, because by the very definition of random error, one does not know the sign of Δx, Δy, and so on. Even if we assume all of them to be positive, the derivatives may have different signs, leading to some error terms remaining positive and others becoming negative. Consequently, a simple summation might lead to a deceptively small final error by cancellation. A better way might be

$$\Delta M = |\Delta M_x| + |\Delta M_y| + |\Delta M_z| + \cdots$$

i.e., we are now forcing the sum of the errors to be greater than the errors of the contributing variables, which is what we would expect. However, it turns out that this method produces an error that is much too large.

The correct way to combine errors is called the Root Sum Square (RSS) method. Accordingly,

$$\Delta M = \sqrt{(\Delta M_x)^2 + (\Delta M_y)^2 + (\Delta M_z)^2 + \cdots}.$$

Finally,

$$\Delta M = \sqrt{\left(\frac{\partial f}{\partial x}\Delta x\right)^2 + \left(\frac{\partial f}{\partial y}\Delta y\right)^2 + \left(\frac{\partial f}{\partial z}\Delta z\right)^2 + \cdots} \tag{2}$$

Examples:

1. Consider the simplest possible example, of measuring volume flow rate Q by collecting a volume V in time t.

   $$Q = \frac{V}{t}$$

   According to result (2) we can write:

   $$\begin{aligned} \Delta Q &= \sqrt{\left(\frac{\partial Q}{\partial V}\Delta V\right)^2 + \left(\frac{\partial Q}{\partial t}\Delta t\right)^2} \\ &= \sqrt{\left(\frac{1}{t}\Delta V\right)^2 + \left(\frac{V}{t^2}\Delta t\right)^2} \end{aligned}$$

   Expressing the error as a *fractional* error you can easily show that:

   $$\frac{\Delta Q}{Q} = \sqrt{\left(\frac{\Delta V}{V}\right)^2 + \left(\frac{\Delta t}{t}\right)^2}$$

Let us use some numerical values, say, $V = 200$ ml and $\Delta V = \pm 1$ ml. Similarly, let $t = 10$ sec with $\Delta t = \pm 0.1$ sec.

Then,

$$\begin{aligned}
\frac{\Delta Q}{Q} &= \sqrt{\left(\frac{1}{200}\right)^2 + \left(\frac{0.1}{10}\right)^2} \\
\frac{\Delta Q}{Q} &= \frac{1}{100}\sqrt{0.25 + 1} \\
&= 1.1\%
\end{aligned}$$

In order to reduce error, would it make make sense to buy a very expensive device to measure volume (such as an accurate balance) so that $\Delta V = 0.01$ ml? Not really! As can be seen from the above, the total error is dominated by Δt anyway, so any attempt to further reduce ΔV is wasted effort. So, with the given apparatus, how can you reduce overall error? One simple way is to increase the amount of water collected. If you double V, then t will also double (for the same Q) and the overall fractional error will reduce by 1/2.

Moral of the story: Always measure the *largest* quantity that your apparatus will allow for any experimental situation.

2. Consider a venturi tube that is used to measure Q. We know that,

$$Q = A_B \sqrt{\frac{2g(h_A - h_B)}{1 - (A_B/A_A)^2}}$$

where A_A and A_B are the upstream and throat areas, and h_A and h_B are the heights in the corresponding piezometers.

Assuming that A_A, A_B and g are given (no associated errors), we can rewrite as follows:

$$Q = C\sqrt{h_A - h_B}$$

where C is a constant. Using the identical procedure as before, you can show that:

$$\begin{aligned}
\frac{\Delta Q}{Q} &= \sqrt{\left(\frac{\Delta h_A}{2(h_A - h_B)}\right)^2 + \left(\frac{\Delta h_B}{2(h_A - h_B)}\right)^2} \\
&= \frac{1}{2(h_A - h_B)}\sqrt{\Delta h_A^2 + \Delta h_B^2}
\end{aligned}$$