A8 *Weight factors when variances are not equal*

What is the "best" determination of the mean of a number of data x_i with the same expectations μ but with unequal standard deviations σ_i?

The answer is: take a *weighted average*:

$$\langle x \rangle = \frac{1}{w}\sum_{i=1}^{n} w_i x_i; \quad w = \sum_{i=1}^{n} w_i. \tag{A8.1}$$

But the question remains: how should you choose the w's? What criterion for the "best" choice is valid here? The criterion that the estimate of the mean should be *unbiased*, i.e., that the expectation of the mean should be equal to μ, is not useful because this is true for any choice of weight factors. The next obvious criterion is the *minimal variance estimate*: the most sharp and hence most accurate value. So let us determine w_i such that

$$E[(\langle x \rangle - \mu)^2] = E[\langle x - \mu \rangle^2] \; minimal, \tag{A8.2}$$

or

$$\begin{aligned} E[\langle x - \mu \rangle^2] &= E\left[\frac{1}{w^2}\left(\sum_i w_i(x_i - \mu)\right)^2\right] \\ &= \frac{1}{w^2}\sum_{i,j} w_i w_j E[(x_i - \mu)(x_j - \mu)] \\ &= \frac{1}{w^2}\sum_{i,j} w_i w_j \operatorname{cov}(x_i, x_j) = minimal. \end{aligned} \tag{A8.3}$$

Now assume that x_i and x_j are uncorrelated, meaning that in the summation only $j = i$ survives. So we search for the minimum of the quantity $\sum_i w_i^2 \sigma_i^2$ under the condition that $\sum_i w_i$ remains constant. The standard way to solve such an *optimization with boundary condition* problem is Lagrange's method of undetermined multipliers. In this method the boundary condition ($\sum_i w_i$ is constant) is multiplied by an as-yet-undetermined multiplier λ and then added to the function that is to be minimized. The partial derivatives of this total function with respect to each of the variables is then set to zero. The solution of the obtained set of equations still contains the undetermined multiplier, but the latter follows from the boundary condition. This is the way it goes:

$$\frac{\partial}{\partial w_i}\left(\sum_j w_j^2\sigma_j^2 + \lambda \sum_j w_j\right) = 2w_i\sigma_i^2 + \lambda = 0. \tag{A8.4}$$

Therefore

$$w_i \propto \frac{1}{\sigma_i^2}. \tag{A8.5}$$

The conclusion must be that the weight of each data point must be proportional to the inverse variance of that point. This is valid when the deviations of the data points are uncorrelated.

The same conclusion can be reached if it is *assumed* that the distribution of deviations is normal. However, the requirement of minimal variance is much more general and the result applies to any distribution function with finite variance.

How large is the variance in $\langle x \rangle$?

For this the expectation of $(\langle x \rangle - \mu)^2$ must be computed:

$$\sigma_{\langle x \rangle}^2 = E[(\langle x \rangle - \mu)^2] = \frac{1}{w^2}\sum_i w_i^2\sigma_i^2.$$

Here use is made of the fact that x_i and x_j are uncorrelated. For w_i we choose $w_i = 1/\sigma_i^2$ and it follows that

$$\sigma_{\langle x \rangle}^2 = \frac{1}{w^2}\sum_i \frac{1}{\sigma_i^2} = \left(\sum_i \frac{1}{\sigma_i^2}\right)^{-1}. \tag{A8.6}$$