To avoid singularities in the integral, let us take a power of m slightly greater than -1, for example -0.99, and let us limit its domain to 30, getting

$$f_{\circ S}(m) = \frac{0.01 \cdot 30^{0.01}}{m^{0.99}}\,. \tag{6.27}$$

The upper limit becomes

$$m < 0.006\,\mathrm{eV}/c^2 \quad \text{at} \;\; 0.95\,\% \text{ probability}\,. \tag{6.28}$$

Any experienced physicist would find this result ridiculous. The upper limit is about $0.2\,\%$ of the experimental resolution; rather like expecting to resolve objects having dimensions smaller than a micron with a design ruler! Note instead that in the previous examples the limit was always of the order of magnitude of the experimental resolution σ. As $f_{\circ S}(m)$ becomes more and more peaked at zero (power of $x \to 1$) the limit gets smaller and smaller. This means that, asymptotically, the degree of belief that $m = 0$ is so high that whatever you measure you will conclude that $m = 0$: you could use the measurement to calibrate the apparatus! This means that this choice of initial distribution was unreasonable.

Instead, priors motivated by the *positive attitude* of the researchers are much more stable, and even when the observation is "very negative" the result is stable, and one always gets a limit of the order of the experimental resolution. Anyhow, it is also clear that when x is several σ below zero one starts to suspect that "something is wrong with the experiment", which formally corresponds to doubts about the likelihood itself. In this case one needs to change analysis model. An example of remodelling the likelihood is shown in Chapter 11.

We shall come back to this delicate issue in Chapter 13.

6.8 Uncertainty of the instrument scale offset

In our scheme any quantity of influence of which we do not know the exact value is a source of systematic error. It will change the final distribution of μ and hence its uncertainty. We have already discussed the most general case in Sec. 5.1.1. Let us make a simple application making a small variation to the example in Sec. 6.2: the "zero" of the instrument is not known exactly, owing to calibration uncertainty. This can be parametrized assuming that its true value Z is normally distributed around 0 (i.e. the calibration was properly done!) with a standard deviation σ_Z. Since, most probably, the true value of μ is independent of the true value of Z, the initial joint probability density function can be written as the product of the marginal ones:

$$f_\circ(\mu, z) = f_\circ(\mu)\, f_\circ(z) = k\, \frac{1}{\sqrt{2\,\pi}\,\sigma_Z}\, \exp\left[-\frac{z^2}{2\,\sigma_Z^2}\right]. \tag{6.29}$$

Also the likelihood changes with respect to Eq. (6.1):

$$f(x_1 \,|\, \mu, z) = \frac{1}{\sqrt{2\,\pi}\,\sigma_1}\, \exp\left[-\frac{(x_1-\mu-z)^2}{2\,\sigma_1^2}\right]. \tag{6.30}$$

Putting all the pieces together and making use of Eq. (5.3) we finally get

$$f(\mu \,|\, x_1, \ldots, f_\circ(z)) = \frac{\int \frac{1}{\sqrt{2\,\pi}\,\sigma_1}\, \exp\left[-\frac{(x_1-\mu-z)^2}{2\,\sigma_1^2}\right] \frac{1}{\sqrt{2\,\pi}\,\sigma_Z}\, \exp\left[-\frac{z^2}{2\,\sigma_Z^2}\right] \mathrm{d}z}{\iint \frac{1}{\sqrt{2\,\pi}\,\sigma_1}\, \exp\left[-\frac{(x_1-\mu-z)^2}{2\,\sigma_1^2}\right] \frac{1}{\sqrt{2\,\pi}\,\sigma_Z}\, \exp\left[-\frac{z^2}{2\,\sigma_Z^2}\right] \mathrm{d}\mu\, \mathrm{d}z}\,.$$

Integrating we get

$$f(\mu) = f(\mu \,|\, x_1, \ldots, f_\circ(z)) = \frac{1}{\sqrt{2\,\pi}\,\sqrt{\sigma_1^2+\sigma_Z^2}}\, \exp\left[-\frac{(\mu-x_1)^2}{2\,(\sigma_1^2+\sigma_Z^2)}\right]. \tag{6.31}$$

(It may help to know that

$$\int_{-\infty}^{+\infty} \exp\left[b\,x - \frac{x^2}{a^2}\right] \mathrm{d}x = \sqrt{a^2\,\pi}\, \exp\left[\frac{a^2\,b^2}{4}\right] \;.)$$

For an introduction to Bayesian methods, where Gaussian integrals are also discussed, see e.g. Ref. [46]. The result is that $f(\mu)$ is still a Gaussian, but with a larger variance. The global standard uncertainty is the quadratic combination of that due to the statistical fluctuation of the data sample and the uncertainty due to the imperfect knowledge of the *systematic effect*:

$$\sigma_{tot}^2 = \sigma_1^2 + \sigma_Z^2\,. \tag{6.32}$$

This result (a theorem under well stated conditions!) is often used as a 'prescription', although there are still some "old-fashioned" recipes which require different combinations of the contributions to be performed.

It must be noted that in this framework it makes no sense to speak of "statistical" and "systematical" uncertainties, as if they were of a different nature. They have the same probabilistic nature: $\overline{Q}_{n_1}$ is around μ with a standard deviation σ_1, and Z is around 0 with standard deviation σ_Z. What distinguishes the two components is how the knowledge of the uncertainty is gained: in one case (σ_1) from repeated measurements; in the second case (σ_Z) the evaluation was done by someone else (the constructor of the instrument), or in a previous experiment, or guessed from the knowledge of the detector, or by simulation, etc. This is the reason why the ISO Guide [5] prefers the generic names *Type A* and *Type B* for the two kinds of contribution to global uncertainty (see Sec. 8.7). In particular, the name "systematic uncertainty" should be avoided, while it is correct to speak about "uncertainty due to a systematic effect".

6.9 Correction for known systematic errors

It is easy to be convinced that if our prior knowledge about Z was of the kind

$$Z \sim \mathcal{N}(z_\circ, \sigma_Z) \tag{6.33}$$

the result would have been

$$\mu \sim \mathcal{N}\left(x_1 - z_\circ, \sqrt{\sigma_1^2 + \sigma_Z^2}\right)\,, \tag{6.34}$$

i.e. one has first to correct the result for the best value of the systematic error and then include in the global uncertainty a term due to imperfect knowledge about it. This is a well-known and practised procedure, although there are still people who confuse $z_\circ$ with its uncertainty.

6.10 Measuring two quantities with the same instrument having an uncertainty of the scale offset

Let us take an example which is a little more complicated (at least from the mathematical point of view) but conceptually very simple and also very common in laboratory practice. We measure two physical quantities with the same instrument, assumed to have an uncertainty on the "zero", modelled with a normal distribution as in the previous sections. For each of the quantities we collect a sample of data under the same conditions, which means that the unknown offset error does not change from one set of measurements to the other. Calling μ_1 and μ_2 the true values, x_1 and x_2 the sample averages, σ_1 and σ_2 the average's standard deviations, and Z the true value of the "zero", the initial probability density and the likelihood are

$$
\begin{aligned}
f_\circ(\mu_1,\mu_2,z) &= f_\circ(\mu_1)\, f_\circ(\mu_2)\, f_\circ(z) = k\,\frac{1}{\sqrt{2\,\pi}\,\sigma_Z}\,\exp\left[-\frac{z^2}{2\,\sigma_Z^2}\right]\\
f(x_1,x_2\,|\,\mu_1,\mu_2,z) &= \frac{1}{\sqrt{2\,\pi}\,\sigma_1}\,\exp\left[-\frac{(x_1-\mu_1-z)^2}{2\,\sigma_1^2}\right]\\
&\quad\times\frac{1}{\sqrt{2\,\pi}\,\sigma_2}\,\exp\left[-\frac{(x_2-\mu_2-z)^2}{2\,\sigma_2^2}\right]\\
&= \frac{1}{2\,\pi\,\sigma_1\sigma_2}\exp\left[-\frac{1}{2}\left(\frac{(x_1-\mu_1-z)^2}{\sigma_1^2}\right.\right.\\
&\quad\left.\left.+\frac{(x_2-\mu_2-z)^2}{\sigma_2^2}\right)\right]\,. \qquad (6.35)
\end{aligned}
$$

The result of the inference is now the joint probability density function of μ_1 and μ_2:

$$
f(\mu_1,\mu_2\,|\,x_1,x_2,\sigma_1,\sigma_2,f_\circ(z)) = \frac{\int f(x_1,x_2\,|\,\mu_1,\mu_2,z)\,f_\circ(\mu_1,\mu_2,z)\mathrm{d}z}{\int\ldots\,\mathrm{d}\mu_1\,\mathrm{d}\mu_2\,\mathrm{d}z}\,, \qquad (6.36)
$$

where expansion of the functions has been omitted for the sake of clarity. Integrating we get

$$
\begin{aligned}
f(\mu_1,\mu_2) &= \frac{1}{2\,\pi\,\sqrt{\sigma_1^2+\sigma_Z^2}\,\sqrt{\sigma_2^2+\sigma_Z^2}\,\sqrt{1-\rho^2}}\\
&\quad\times\exp\left\{-\frac{1}{2\,(1-\rho^2)}\left[\frac{(\mu_1-x_1)^2}{\sigma_1^2+\sigma_Z^2}\right.\right.\\
&\quad\left.\left.-2\,\rho\,\frac{(\mu_1-x_1)(\mu_2-x_2)}{\sqrt{\sigma_1^2+\sigma_Z^2}\,\sqrt{\sigma_2^2+\sigma_Z^2}}\;+\frac{(\mu_2-x_2)^2}{\sigma_2^2+\sigma_Z^2}\right]\right\}\,, \qquad (6.37)
\end{aligned}
$$

where

$$
\rho = \frac{\sigma_Z^2}{\sqrt{\sigma_1^2+\sigma_Z^2}\,\sqrt{\sigma_2^2+\sigma_Z^2}}\,. \qquad (6.38)
$$

If σ_Z vanishes then Eq. (6.37) has the simpler expression

$$
f(\mu_1,\mu_2)\xrightarrow[\sigma_Z\to 0]{}\frac{1}{\sqrt{2\,\pi}\,\sigma_1}\,\exp\left[-\frac{(\mu_1-x_1)^2}{2\,\sigma_1^2}\right]\frac{1}{\sqrt{2\,\pi}\,\sigma_2}\,\exp\left[-\frac{(\mu_2-x_2)^2}{2\,\sigma_2^2}\right], \qquad (6.39)
$$

i.e. if there is no uncertainty on the offset calibration then the joint density function $f(\mu_1,\mu_2)$ is equal to the product of two independent normal functions, i.e. μ_1 and μ_2 are independent. In the general case we have to conclude the following.

- The effect of the *common uncertainty* σ_Z makes the two values correlated, since they are affected by a common unknown systematic error.
- The joint density function is a bivariate Gaussian distribution of parameters x_1, $\sigma_{\mu_1} = \sqrt{\sigma_1^2 + \sigma_Z^2}$, x_2, $\sigma_{\mu_2} = \sqrt{\sigma_2^2 + \sigma_Z^2}$, and ρ (see example of Fig. 4.2).
- The marginal distributions are still normal:

$$\mu_1 \sim \mathcal{N}\left(x_1, \sqrt{\sigma_1^2 + \sigma_Z^2}\right), \tag{6.40}$$

$$\mu_2 \sim \mathcal{N}\left(x_2, \sqrt{\sigma_2^2 + \sigma_Z^2}\right). \tag{6.41}$$

- The covariance between μ_1 and μ_2 is

$$\begin{aligned} \mathrm{Cov}(\mu_1, \mu_2) &= \rho\, \sigma_{\mu_1} \sigma_{\mu_2} \\ &= \rho \sqrt{\sigma_1^2 + \sigma_Z^2}\, \sqrt{\sigma_2^2 + \sigma_Z^2} = \sigma_Z^2\,. \end{aligned} \tag{6.42}$$

- The correlation coefficient is always non-negative ($\rho \geq 0$), as intuitively expected from the definition of this kind of systematic error. The correlation coefficient vanishes when σ_Z is much smaller than σ_1 and σ_2, tends to 1 if σ_Z dominates (the uncertainties become 100% correlated).
- The distribution of any function $g(\mu_1, \mu_2)$ can be calculated using the standard methods of probability theory. For example, one can demonstrate that the sum $S = \mu_1 + \mu_2$ and the difference $D = \mu_1 - \mu_2$ are also normally distributed (see also the introductory discussion to the central limit theorem and Sec. 8.13 for the calculation of averages and standard deviations):

$$S \sim \mathcal{N}\left(x_1 + x_2, \sqrt{\sigma_1^2 + \sigma_2^2 + (2\,\sigma_Z)^2}\right), \tag{6.43}$$

$$D \sim \mathcal{N}\left(x_1 - x_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right). \tag{6.44}$$

  The result can be interpreted in the following way.

  – The uncertainty on the difference does not depend on the common offset uncertainty: whatever the value of the true “zero” is, it cancels in differences.

– In the sum, instead, the effect of the common uncertainty is somewhat amplified since it enters "in phase" in the global uncertainty of each of the quantities.

6.11 Indirect calibration

Let us use the result of the previous section to solve another typical problem of measurements. Suppose that after (or before, it doesn't matter) we have done the measurements of x_1 and x_2 and we have the final result, summarized in Eq. (6.37), we know the "exact" value of μ_1 (for example we perform the measurement on a reference). Let us call it μ_1°. Will this information provide a better knowledge of μ_2? In principle yes: the difference between x_1 and μ_1° defines the systematic error (the true value of the "zero" Z). This error can then be subtracted from x_2 to get a corrected value. Also the overall uncertainty of μ_2 should change, intuitively it "should" decrease, since we are adding new information. But its value doesn't seem to be obvious, since the logical link between μ_1° and μ_2 is $\mu_1^\circ \to Z \to \mu_2$.

The problem can be solved exactly using the concept of conditional probability density function $f(\mu_2 \,|\, \mu_1^\circ)$ [see Eqs. (4.83)-(4.84)]. We get

$$\mu_{2\,|\,\mu_1^\circ} \sim \mathcal{N}\left(x_2 + \frac{\sigma_Z^2}{\sigma_1^2+\sigma_Z^2}\,(\mu_1^\circ - x_1),\ \sqrt{\sigma_2^2 + \left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_Z^2}\right)^{-1}}\right)\,. \quad (6.45)$$

The best value of μ_2 is shifted by an amount Δ, with respect to the measured value x_2, which is not exactly $x_1 - \mu_1^\circ$, as was naïvely guessed, and the uncertainty depends on σ_2, σ_Z and σ_1. It is easy to be convinced that the exact result is more reasonable than the (suggested) first guess. Let us rewrite Δ in two different ways:

$$\Delta = \frac{\sigma_Z^2}{\sigma_1^2+\sigma_Z^2}\,(\mu_1^\circ - x_1) \quad (6.46)$$

$$= \frac{1}{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_Z^2}}\left[\frac{1}{\sigma_1^2}\cdot(x_1-\mu_1^\circ) + \frac{1}{\sigma_Z^2}\cdot 0\right]\,. \quad (6.47)$$

- Equation (6.46) shows that one has to apply the correction $x_1 - \mu_1^\circ$ only if $\sigma_1 = 0$. If instead $\sigma_Z = 0$ there is no correction to be applied, since the instrument is perfectly calibrated. If $\sigma_1 \approx \sigma_Z$ the correction is half of the measured difference between x_1 and μ_1°.