

1 The χ^2 in the presence of correlations

Suppose to have an ensemble of n measurements having the following structure:

$$m_i \pm \sigma_{i,\text{stat}} \pm \sigma_{i,\text{unc}} \pm \sigma_{i,\text{corr}}^{(1)} \pm \cdots \pm \sigma_{i,\text{corr}}^{(k)}, \quad (1)$$

where m_i , with $i = 1, \dots, n$, is the central value of the i -th measurement, $\sigma_{i,\text{stat}}$ its (uncorrelated) statistical uncertainty, $\sigma_{i,\text{unc}}$ its uncorrelated systematic uncertainty¹, and $\sigma_{i,\text{corr}}^{(l)}$, with $l = 1, \dots, k$, its correlated systematic uncertainties. With this information at hand, one can construct the full covariance matrix V_{ij} as follows (see for example Ref. [3]):

$$V_{ij} = (\sigma_{i,\text{stat}}^2 + \sigma_{i,\text{unc}}^2) \delta_{ij} + \sum_{l=1}^k \sigma_{i,\text{corr}}^{(l)} \sigma_{j,\text{corr}}^{(l)}. \quad (2)$$

This is a clearly symmetric matrix. Given a set of predictions t_i corresponding to the n measurements of the ensemble, the χ^2 takes the form:

$$\chi^2 = \sum_{i,j=1}^n (m_i - t_i) V_{ij}^{-1} (m_j - t_j) = \mathbf{y}^T \cdot \mathbf{V}^{-1} \cdot \mathbf{y}, \quad (3)$$

where in the second equality we have used the matricial notation and defined $y_i = m_i - t_i$. A convenient way to compute the χ^2 relies on the Cholesky decomposition of the covariance matrix \mathbf{V} . In particular, it can be proven that any symmetric and positive definite matrix \mathbf{V} can be decomposed as:

$$\mathbf{V} = \mathbf{L} \cdot \mathbf{L}^T, \quad (4)$$

where \mathbf{L} is a lower triangular matrix whose entries are related recursively to those of \mathbf{V} as follows:

$$\begin{aligned} L_{kk} &= \sqrt{V_{kk} - \sum_{j=1}^{k-1} L_{kj}^2}, \\ L_{ik} &= \frac{1}{L_{kk}} \left(V_{ik} - \sum_{j=1}^{k-1} L_{ij} L_{kj} \right), \quad k < i, \\ L_{ik} &= 0, \quad k > i. \end{aligned} \quad (5)$$

It is then easy to see that the χ^2 can be written as:

$$\chi^2 = |\mathbf{L}^{-1} \cdot \mathbf{y}|^2. \quad (6)$$

But the vector $\mathbf{x} \equiv \mathbf{L}^{-1} \cdot \mathbf{y}$ is the solution of the linear system:

$$\mathbf{L} \cdot \mathbf{x} = \mathbf{y}, \quad (7)$$

that can be efficiently solved by forward substitution, so that:

$$\chi^2 = |\mathbf{x}|^2. \quad (8)$$

Following this procedure, one does not need to compute explicitly the inverse of the covariance matrix \mathbf{V} , simplifying significantly the computation of the χ^2 .

¹There could be more than one uncorrelated systematic uncertainty. In this case, $\sigma_{i,\text{unc}}$ is just the square root of the sum in quadrature of all the uncorrelated systematic uncertainties.

2 Additive and multiplicative uncertainties

The correlated systematic uncertainties $\sigma_{i,\text{corr}}^{(l)}$ may be either *additive* or *multiplicative*. The nature of the single uncertainties is typically provided by the experiments that release the measurements. A typical example of multiplicative uncertainty is the luminosity uncertainty but there can be others.

Now let us express all the correlated systematic uncertainties $\sigma_{i,\text{corr}}^{(l)}$ as relative to the associate central value m_i , so that we define²:

$$\sigma_{i,\text{corr}}^{(l)} \equiv \delta_{i,\text{corr}}^{(l)} m_i \quad (9)$$

and let us also define $s_i^2 \equiv \sigma_{i,\text{stat}}^2 + \sigma_{i,\text{unc}}^2$ so that Eq. (2) can be rewritten as:

$$V_{ij} = s_i^2 \delta_{ij} + \left(\sum_{l=1}^k \delta_{i,\text{corr}}^{(l)} \delta_{j,\text{corr}}^{(l)} \right) m_i m_j. \quad (10)$$

Now we split the correlated systematic uncertainties into k_a additive uncertainties and k_m multiplicative uncertainties, such that $k_a + k_m = k$. This way Eq. (10) takes the form:

$$V_{ij} = s_i^2 \delta_{ij} + \left(\sum_{l=1}^{k_a} \delta_{i,\text{add}}^{(l)} \delta_{j,\text{add}}^{(l)} + \sum_{l=1}^{k_m} \delta_{i,\text{mult}}^{(l)} \delta_{j,\text{mult}}^{(l)} \right) m_i m_j. \quad (11)$$

It is well known that this definition of the covariance matrix is problematic in that it results in the so-called D’Agostini bias of the multiplicative uncertainties [2]. A possible solution to this problem is the so-called t_0 -prescription [1], where the experimental central value m_i in the multiplicative term is replaced by a fixed theoretical predictions $t_i^{(0)}$, typically computed in a previous fit in which the “standard” definition of the covariance matrix in Eq. (2) (often referred to as *experimental* definition) is used. Applying the t_0 prescription, the covariance matrix takes the form:

$$V_{ij} = s_i^2 \delta_{ij} + \sum_{l=1}^{k_a} \delta_{i,\text{add}}^{(l)} \delta_{j,\text{add}}^{(l)} m_i m_j + \sum_{l=1}^{k_m} \delta_{i,\text{mult}}^{(l)} \delta_{j,\text{mult}}^{(l)} t_i^{(0)} t_j^{(0)}. \quad (12)$$

3 Artificial generation of correlated systematics

In order to implement the definition of the χ^2 discussed above, it is necessary to have the experimental information in terms of the correlated systematic uncertainties $\sigma_{i,\text{corr}}^{(l)}$. This is what the experimental collaborations usually release. However, in some cases this information is given in terms of a covariance matrix. Therefore, one needs to find a workaround to generate correlated systematic uncertainties out of a covariance matrix.

Given a $n \times n$ symmetric matrix \mathbf{C} , it will have n orthonormal eigenvectors $\mathbf{x}^{(i)}$, such that $\mathbf{x}^{(i)} \cdot \mathbf{x}^{(j)} = \delta_{ij}$, each of which will have a non-negative eigenvalue λ_i associated:

$$\mathbf{C} \cdot \mathbf{x}^{(i)} = \lambda_i \mathbf{x}^{(i)}, \quad i = 1, \dots, n. \quad (13)$$

²Note that this redefinition does not change the nature of the uncertainties, additive uncertainties remain additive as well as multiplicative uncertainties remain multiplicative.

If we define:

$$\sigma_{i,\text{corr}}^{(l)} = \sqrt{\lambda_l} x_i^{(l)}, \quad i, l = 1, \dots, n, \quad (14)$$

one can show that:

$$\sum_{l=1}^n \sigma_{i,\text{corr}}^{(l)} \sigma_{j,\text{corr}}^{(l)} = C_{ij}. \quad (15)$$

To prove this equality we start from the following matricial relation:

$$\mathbf{C} = \mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^{-1}, \quad (16)$$

where $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues λ_i on the diagonal ($\Lambda_{ij} = \lambda_i \delta_{ij}$), while \mathbf{Q} is a matrix whose columns are the eigenvectors $\mathbf{x}^{(i)}$ ($Q_{ij} = x_i^{(j)}$). In addition, since in this particular case $\mathbf{x}^{(i)} \cdot \mathbf{x}^{(j)} = \delta_{ij}$, this implies that:

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I} \quad \Rightarrow \quad \mathbf{Q}^{-1} = \mathbf{Q}^T, \quad (17)$$

so that:

$$\mathbf{C} = \mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}^T. \quad (18)$$

It follows that:

$$C_{ij} = \sum_{k,l=1}^n Q_{ik} \Lambda_{kl} Q_{jl} = \sum_{k,l=1}^n x_i^{(k)} \lambda_k \delta_{kl} x_j^{(l)} = \sum_{l=1}^n \lambda_l x_i^{(l)} x_j^{(l)} = \sum_{l=1}^n \sigma_{i,\text{corr}}^{(l)} \sigma_{j,\text{corr}}^{(l)}, \quad (19)$$

as required.

The matrix \mathbf{C} can be regarded as the correlated contribution to the full covariance matrix \mathbf{V} . In particular, considering Eqs. (2) and (10), one can write:

$$\mathbf{V} = \mathbf{U} + \mathbf{C}, \quad (20)$$

where \mathbf{U} is a diagonal matrix of uncorrelated uncertainties:

$$U_{ij} = s_i^2 \delta_{ij}. \quad (21)$$

This defines the matrix \mathbf{C} as:

$$\mathbf{C} = \mathbf{V} - \mathbf{U}, \quad (22)$$

such that, given a $n \times n$ covariance matrix \mathbf{V} along with its uncorrelated contribution \mathbf{U} , one can generate a set of n *artificial* correlated systematics according to Eq. (14), where \mathbf{C} is given in Eq. (22), for each of the n measurements. This allows us to implement Eq. (12) for the construction of the covariance matrix.

4 Determining the systematic shifts

In order to visualise the effect of systematic uncertainties, it is instructive to compute the *systematic shift* generated by the systematic uncertainties. To do so, we need to write the χ^2 in terms of the so-called “nuisance parameters” λ_α . One can show that the definition of the χ^2 in Eq. (3) is equivalent to [3]:

$$\chi^2 = \sum_{i=1}^n \frac{1}{s_i^2} \left(m_i - t_i - \sum_{\alpha=1}^k \lambda_\alpha \sigma_{i,\text{corr}}^{(\alpha)} \right)^2 + \sum_{\alpha=1}^k \lambda_\alpha^2. \quad (23)$$

The optimal value of the nuisance parameters can be computed by minimising the χ^2 with respect to them, that is imposing:

$$\frac{\partial \chi^2}{\partial \lambda_\beta} = 0. \quad (24)$$

This yields the system:

$$\sum_{\beta=1}^k A_{\alpha\beta} \lambda_\beta = \rho_\alpha, \quad (25)$$

with:

$$A_{\alpha\beta} = \delta_{\alpha\beta} + \sum_{i=1}^n \frac{\sigma_{i,\text{corr}}^{(\alpha)} \sigma_{i,\text{corr}}^{(\beta)}}{s_i^2} \quad \text{and} \quad \rho_\alpha = \sum_{i=1}^n \frac{m_i - t_i}{s_i^2} \sigma_{i,\text{corr}}^{(\alpha)}, \quad (26)$$

that determines the values of λ_β . The quantity:

$$d_i = \sum_{\alpha=1}^k \lambda_\alpha \sigma_{i,\text{corr}}^{(\alpha)} \quad (27)$$

in Eq. (23) can be interpreted as a shift caused by the correlated systematic uncertainties. Defining the shifted predictions as:

$$\bar{t}_i = t_i + d_i, \quad (28)$$

the χ^2 reads:

$$\chi^2 = \sum_{i=1}^n \left(\frac{m_i - \bar{t}_i}{s_i} \right)^2 + \sum_{\alpha=1}^k \lambda_\alpha^2. \quad (29)$$

Therefore, up to a penalty term given by the sum of the square of the nuisance parameters, the χ^2 takes the form of the uncorrelated definition. In order to achieve a visual assessment of the agreement between data and theory, it appears natural to compare the central experimental values m_i to the shifted theoretical predictions \bar{t}_i in units of the uncorrelated uncertainty s_i .

References

- [1] R. D. Ball *et al.* [NNPDF Collaboration], JHEP **1005** (2010) 075 doi:10.1007/JHEP05(2010)075 [arXiv:0912.2276 [hep-ph]].
- [2] G. D’Agostini, Nucl. Instrum. Meth. A **346** (1994) 306. doi:10.1016/0168-9002(94)90719-6
- [3] R. D. Ball *et al.*, JHEP **1304** (2013) 125 doi:10.1007/JHEP04(2013)125 [arXiv:1211.5142 [hep-ph]].
- [4] R. Boughezal, A. Guffanti, F. Petriello and M. Ubiali, JHEP **1707** (2017) 130 doi:10.1007/JHEP07(2017)130 [arXiv:1705.00343 [hep-ph]].