Statistical analysis of experimental data
Systemactic uncertainties

Aleksander Filip Żarnecki

FACULTY OF PHYSICS
UNIVERSITY OF WARSAW

Lecture 10
December 5, 2024

Systemactic uncertainties

General case

Lecture 08

We introduced χ^2 in a very general form:

$$\chi^2 = \sum_{i=1}^{N} \frac{(y_i - \mu_i)^2}{\sigma_i^2}$$

where different μ_i and σ_i are possible for each of N measurement y_i

It is quite often the case that values of μ_i depend on some controlled variables x_i and a smaller set of model parameters:

$$\mu_i = \mu(x_i; \mathbf{a})$$

we can then use the least-squares method to extract the best estimates of parameters **a** from the collected set of data points (x_i, y_i)

We can look for minimum of χ^2 using different numerical algorithms...

Linear case

The case which is particularly interesting is when the dependence is linear in parameters (!):

$$\mu(x; \mathbf{a}) = \sum_{k=1}^{M} a_k \, f_k(x)$$

where $f_k(x)$ is a set of functions with arbitrary analytical form.

One of the examples is the polynomial series:

$$f_k(x) = x^k \qquad \Rightarrow \quad \mu(x; \mathbf{a}) = \sum_{k=1}^{M} a_k \, x^{k-1}$$

but any set of functions can be used, if they are not linearly dependent.
Set of functions ortogonal for a given set of points x_i should work best...

Parameter fit

We obtain a set of M equations for M parameters a_i. In the matrix form:

$$\mathbb{A} \cdot \mathbf{a} = \mathbf{b}$$

$$\text{where} \quad \mathbb{A}_{lk} = \sum_{i=1}^{N} \frac{f_l(x_i)\; f_k(x_i)}{\sigma_i^2} \quad \text{and} \quad b_l = \sum_{i=1}^{N} \frac{f_l(x_i)\; y_i}{\sigma_i^2}$$

Solution can be obtained by inverting matrix $\mathbb{A}$:

$$\mathbf{a} = \mathbb{A}^{-1} \cdot \mathbf{b}$$

This also gives us the estimate of parameter covariance matrix:

$$\mathbb{C}_{\mathbf{a}} = \left(\sum_{i=1}^{N} \frac{f_l(x_i)\; f_k(x_i)}{\sigma_i^2} \right)^{-1} = \mathbb{A}^{-1}$$

Iterative procedure

Lecture 09

We start from some "initial guess" of parameter values $\mathbf{a_0}$.

Assuming small variations of the model parameters, $\mathbf{a} = \mathbf{a_0} + \delta\mathbf{a}$, we can expand χ^2 in a series:

$$\chi^2(\mathbf{a}) \;=\; \chi^2(\mathbf{a_0}) - 2\,\mathbf{b}\cdot(\mathbf{a}-\mathbf{a_0}) + \ldots$$

where $\mathbf{b}$ is the negative gradient of χ^2:

$$\mathbf{b} \;=\; -\frac{1}{2}\,\nabla\,\chi^2(\mathbf{a_0}) \qquad\qquad b_j \;=\; -\frac{1}{2}\frac{\partial\chi^2}{\partial a_j} \;=\; \sum_{i=1}^{N}\frac{y_i-\mu_i}{\sigma_i^2}\cdot\frac{\partial\mu_i}{\partial a_j}$$

Vector $\mathbf{b}$ defines the direction of steepest χ^2 descent.

One of the possible procedures is to make a step in this direction:

$$\mathbf{a_1} \;=\; \mathbf{a_0} \;+\; \varepsilon\,\mathbf{b}$$

with small $\varepsilon > 0$ and then repeat the whole procedure...

Iterative procedure

We can try to be “smarter”. Expanding χ^2 to quadratic term:

$$\chi^2(\mathbf{a}) \quad = \quad \chi^2(\mathbf{a_0}) - 2\,\mathbf{b}\cdot(\mathbf{a}-\mathbf{a_0}) + (\mathbf{a}-\mathbf{a_0})^{\mathsf{T}}\mathbb{A}(\mathbf{a}-\mathbf{a_0}) + \ldots$$

where $\mathbb{A}$ is the so called Hessian matrix of second derivatives:

$$\mathbb{A}_{jk} \quad = \quad \frac{1}{2}\,\frac{\partial^2\chi^2}{\partial a_j\,\partial a_k}\bigg|_{\mathbf{a}=\mathbf{a_0}} \approx \sum_{i=1}^{N}\frac{1}{\sigma_i^2}\cdot\frac{\partial\mu_i}{\partial a_j}\cdot\frac{\partial\mu_i}{\partial a_k} \qquad \left(\text{neglecting } \frac{\partial^2\mu_i}{\partial a_j\,\partial a_k}\right)$$

In this approximation, we can calculate the expected position of the χ^2 minimum:

$$\nabla\,\chi^2(\mathbf{a}) \quad = \quad -2\,\mathbf{b} \,+\, 2\,\mathbb{A}\,(\mathbf{a}-\mathbf{a_0}) \quad = \; 0$$

$$\Rightarrow \quad \mathbf{a_m} \;=\; \mathbf{a_0} \,+\, \mathbb{A}^{-1}\mathbf{b}$$

and we can try to “jump” directly to the minimum...

Marquardt Minimization

One of the popular approaches, combining the two previously discussed:

$$\mathbf{a_{i+1}} = \mathbf{a_i} + \left(\mathbb{A} + \lambda \cdot \mathbb{I}\right)^{-1} \mathbf{b}$$

where λ is an additional parameter determining the performance of the algorithm:

- for $\lambda \gg 1$ we make a small step along the gradient direction
  which corresponds to the gradient minimization with $\varepsilon \approx \frac{1}{\lambda}$
- for $\lambda \ll 1$ we try to "jump" directly to the minimum position
  Hessian matrix solution is reproduced for $\lambda \to 0$

The key element proposed by D.W.Marquardt (1963) was to use variable λ parameter, adjusting its value to the results of the previous step...

Model constraints

We consider set of N measurement points (x_i, y_i), which can be compared to model predictions depending on parameters **a** in terms of the χ^2 value:

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}))^2}{\sigma_i^2}$$

Best estimate of **a** should correspond to the minimum of $\chi^2(\mathbf{a})$.

However, we now need to look for this minimum taking additional constraints into account:

$$w_k(\mathbf{a}) = 0 \qquad k = 1 \ldots K$$

where number of constraints K should be lower than number of parameters M.

How can we find the best parameter values in this case?

Model reduction

The first approach is to reduce number of model parameters, using constraints to eliminate some of the model variables. $\Rightarrow$ We thus reduce the problem with M model parameters to problem with $M' = M - K$ independent parameters. (method of elements)

Example

We would like to fit polynomial model to a series of measurements where the azimuthal angle $\theta \in [-\pi, +\pi]$ is the controlled variable:

$$\mu(x; \mathbf{a}) = \sum_{k=0}^{M-1} a_k \left(\frac{\theta}{\pi}\right)^k = \sum_k a_k \, x^k$$

where we introduced $x = \frac{\theta}{\pi}$ for simplicity.

And we expect that the distribution should vanish for $\theta \to \pm\pi$:

$$\mu(-1; \mathbf{a}) = \mu(+1; \mathbf{a}) = 0 \qquad K = 2$$

Method of Lagrange Multipliers

The method, invented by J.L.Lagrange in 1788, applies to general minimization problem with additional constraints imposed.

Problem of finding minimum of $\chi^2\,(\mathbf{a})$ with constraints $w_k(\mathbf{a}) = 0$ is equivalent to finding a stationary point (point with all first derivatives at zero) of the Lagrange function:

$$\mathcal{L}(\mathbf{a}, \boldsymbol{\lambda}) \;=\; \chi^2(\mathbf{a}) + \sum_k 2\lambda_k w_k(\mathbf{a})$$

where we introduce additional K parameters λ_k - Lagrange multipliers

Our problem is now reduced to finding parameters $\mathbf{a}$ and $\boldsymbol{\lambda}$ fulfilling

$$\frac{\partial \mathcal{L}}{\partial a_j} \;=\; 0 \qquad \text{and} \qquad \frac{\partial \mathcal{L}}{\partial \lambda_k} \;=\; 0$$

(without any additional constraints)

Method of Lagrange Multipliers

We can write these equations the matrix form:

$$\underbrace{\left(\begin{array}{c|c} \mathbb{A} & \mathbb{D} \\ \hline \mathbb{D}^\mathsf{T} & 0 \end{array}\right)}_{\tilde{\mathbb{A}}} \cdot \left(\begin{array}{c} \mathbf{a} \\ \hline \boldsymbol{\lambda} \end{array}\right) = \left(\begin{array}{c} \mathbf{b} \\ \hline \mathbf{c} \end{array}\right)$$

where: $\mathbb{A}_{jk} = \sum_{i=1}^{N} \frac{f_j(x_i)\, f_k(x_i)}{\sigma_i^2}$, $\mathbb{D}_{jk} = d_{k,\,j}$ and $b_j = \sum_{i=1}^{N} \frac{f_j(x_i)\, y_i}{\sigma_i^2}$

and the problem can be solved by inverting matrix $\tilde{\mathbb{A}}$.

Covariance matrix for **a** can be extracted as: (seems to work for linear problems)

$$(\mathbb{C}_\mathbf{a})_{ij} = (\tilde{\mathbb{A}}^{-1})_{ij} \qquad i,j = 1 \ldots M$$

Systemactic uncertainties

Statistical uncertainties

We have considered numerical results of experiments as random variables.

Probability to obtain given numerical result was described by PDF.
Results of a repeated experiment were considered as independent variables.

Statistical uncertainties

We have considered numerical results of experiments as random variables.

Probability to obtain given numerical result was described by PDF.
Results of a repeated experiment were considered as independent variables.

Uncertainties of the results were related to the fluctuations in the measurement due to:

- actual nature of the physics process studied (lecture 01)
  eg. exponential distribution for decay time measurement
- finite precision of our instruments
  eg. precision with which decay time is measured in the detector
- inhomogeneity of the population studied
  eg. different particles/isotopes in the considered sample

Statistical uncertainties

We have considered numerical results of experiments as random variables.

Probability to obtain given numerical result was described by PDF.
Results of a repeated experiment were considered as independent variables.

Uncertainties of the results were related to the fluctuations in the measurement due to:

- actual nature of the physics process studied (lecture 01)
  eg. exponential distribution for decay time measurement
- finite precision of our instruments
  eg. precision with which decay time is measured in the detector
- inhomogeneity of the population studied
  eg. different particles/isotopes in the considered sample

Uncertainties related to fluctuations of the individual measurement results
are usually referred to as **statistical uncertainties**.

Systematic uncertainties

Particle physics experiments are quite complex, and so is the data analysis.
We frequently use Monte Carlo methods to correct for different effects.

Simplest example is the (differential) cross section measurement:

$$\boldsymbol{\sigma}_i = \frac{N_i}{\varepsilon_i \, A_i \, \mathcal{L}}$$

where: N_i is the measured number of events (in given bin i), ε - event selection efficiency, A - detector acceptance and $\mathcal{L}$ - integrated luminosity.

Systematic uncertainties

Particle physics experiments are quite complex, and so is the data analysis.
We frequently use Monte Carlo methods to correct for different effects.

Simplest example is the (differential) cross section measurement:

$$\boldsymbol{\sigma}_i = \frac{N_i}{\varepsilon_i \, A_i \, \mathcal{L}}$$

where: N_i is the measured number of events (in given bin i), ε - event selection efficiency, A - detector acceptance and $\mathcal{L}$ - integrated luminosity.

Statistical uncertainty of the extracted cross section value is due to the Poisson fluctuations in the number of reconstructed events.

But we also need to take into account that other factors (ε_i, A_i, $\mathcal{L}$) are also known with finite precision $\Rightarrow$ **systematic uncertainties**

Sources of systematic uncertainties

One needs to distinguish between systematic effect/error and uncertainty!

Sources of systematic uncertainties

One needs to distinguish between systematic effect/error and uncertainty!

(Bonamente) The term systematic error designates sources of error that systematically shift the signal of interest either too high or too low. Sources of systematic errors need to be identified to correct the erroneous offset. A typical example is an instrument that is miscalibrated and systematically reports measurements that have an erroneous offset.

Sources of systematic uncertainties

One needs to distinguish between systematic effect/error and uncertainty!

(Bonamente) The term systematic error designates sources of error that systematically shift the signal of interest either too high or too low. Sources of systematic errors need to be identified to correct the erroneous offset. A typical example is an instrument that is miscalibrated and systematically reports measurements that have an erroneous offset.

(Barlow; quotes after J. Orear, *Notes on Statistics for Physicists*)
"Systematic effects" is a general category which includes effects such as background, selection bias, scanning efficiency, energy resolution, angle resolution, variation of counter efficiency with beam position and energy, dead time, etc.

Sources of systematic uncertainties

One needs to distinguish between systematic effect/error and uncertainty!

(Bonamente) The term systematic error designates sources of error that systematically shift the signal of interest either too high or too low. Sources of systematic errors need to be identified to correct the erroneous offset. A typical example is an instrument that is miscalibrated and systematically reports measurements that have an erroneous offset.

(Barlow; quotes after J. Orear, *Notes on Statistics for Physicists*)
"Systematic effects" is a general category which includes effects such as background, selection bias, scanning efficiency, energy resolution, angle resolution, variation of counter efficiency with beam position and energy, dead time, etc.

Systematic effects are not a problem, if we understand them and know how to model them precisely (correct the final result for systematic error).

Sources of systematic uncertainties

Systematic uncertainty is the uncertainty in the estimation of systematics.

(O.Behnke et. al, *Data Analysis in High Enegy Physics*) Systematic uncertainties:
all uncertainties that are not due to statistical fluctuations in real or simulated data samples.

Sources of systematic uncertainties

Systematic uncertainty is the uncertainty in the estimation of systematics.

(O.Behnke et. al, *Data Analysis in High Enegy Physics*)   Systematic uncertainties:
all uncertainties that are not due to statistical fluctuations in real or simulated data samples.

Systematic effects and their uncertainties are often estimated based on separate, independent experiments. This is the case for both the experimental uncertainties (eg. detector calibration, alignment) as well as those related to the theoretical model (eg. value of the coupling parameter, particle masses, assumed cross sections).

Sources of systematic uncertainties

Systematic uncertainty is the uncertainty in the estimation of systematics.

(O.Behnke et. al, *Data Analysis in High Enegy Physics*) Systematic uncertainties: all uncertainties that are not due to statistical fluctuations in real or simulated data samples.

Systematic effects and their uncertainties are often estimated based on separate, independent experiments. This is the case for both the experimental uncertainties (eg. detector calibration, alignment) as well as those related to the theoretical model (eg. value of the coupling parameter, particle masses, assumed cross sections).

The maximum-likelihood approach can be used to estimate the impact of systematic effect and the resulting uncertainty of the measurement.
Likelihood function of the procedure used to constrain the systematic effect should be folded into the likelihood function of the main experiment.

Systematic uncertainties

In our simple example of cross section measurement:

$$\boldsymbol{\sigma}_i = \frac{N_i}{\varepsilon_i \, A_i \, \mathcal{L}}$$

the statistical uncertainty on $\boldsymbol{\sigma}_i$ is due to Poisson fluctuations in N_i:

$$\sigma_{stat} = \frac{\sigma_{N_i}}{\varepsilon_i \, A_i \, \mathcal{L}} = \frac{\sqrt{N_i}}{\varepsilon_i \, A_i \, \mathcal{L}} = \frac{\boldsymbol{\sigma}_i}{\sqrt{N_i}}$$

Systematic uncertainties

In our simple example of cross section measurement:

$$\boldsymbol{\sigma}_i \;=\; \frac{N_i}{\varepsilon_i \, A_i \, \mathcal{L}}$$

the statistical uncertainty on $\boldsymbol{\sigma}_i$ is due to Poisson fluctuations in N_i:

$$\sigma_{stat} \;=\; \frac{\sigma_{N_i}}{\varepsilon_i \, A_i \, \mathcal{L}} \;=\; \frac{\sqrt{N_i}}{\varepsilon_i \, A_i \, \mathcal{L}} \;=\; \frac{\boldsymbol{\sigma}_i}{\sqrt{N_i}}$$

Uncertainty on $\varepsilon_i \; A_i$ can result from many different sources (including eg. energy calibration), but one should also take into account contribution from the finite statistics of the Monte Carlo events: $r_i = \varepsilon_i \; A_i$

$$\sigma_{r_i} \;=\; \sqrt{\frac{r_i(1-r_i)}{N^{MC}}} \qquad \text{(binomial distribution)}$$

where N^{MC} is the total number of (unweighted) Monte Carlo events (before selection)

Systematic uncertainties

The resulting systematic uncertainty on the cross section measurement:

$$\sigma_{sys(MC)} = \boldsymbol{\sigma}_i \cdot \frac{\sigma_{r_i}}{r_i} = \boldsymbol{\sigma}_i \cdot \sqrt{\frac{1-r_i}{r_i\, N^{MC}}} = \boldsymbol{\sigma}_i \cdot \sqrt{\frac{1-r_i}{N_i^{MC}}}$$

where N_i^{MC} is the number of MC events accepted in cross section bin i.
Uncertainties due to MC statistics are not correlated between bins!

Systematic uncertainties

The resulting systematic uncertainty on the cross section measurement:

$$\sigma_{sys(MC)} = \boldsymbol{\sigma}_i \cdot \frac{\sigma_{r_i}}{r_i} = \boldsymbol{\sigma}_i \cdot \sqrt{\frac{1-r_i}{r_i\, N^{MC}}} = \boldsymbol{\sigma}_i \cdot \sqrt{\frac{1-r_i}{N_i^{MC}}}$$

where N_i^{MC} is the number of MC events accepted in cross section bin i.
Uncertainties due to MC statistics are not correlated between bins!

Systematic uncertainty due to integrated luminosity measurement:

$$\sigma_{sys(\mathcal{L})} = \boldsymbol{\sigma}_i \cdot \frac{\sigma_{\mathcal{L}}}{\mathcal{L}}$$

is 100% correlated between different cross section measurements (bins).

Systematic uncertainties

The resulting systematic uncertainty on the cross section measurement:

$$\sigma_{sys(MC)} = \boldsymbol{\sigma}_i \cdot \frac{\sigma_{r_i}}{r_i} = \boldsymbol{\sigma}_i \cdot \sqrt{\frac{1-r_i}{r_i\, N^{MC}}} = \boldsymbol{\sigma}_i \cdot \sqrt{\frac{1-r_i}{N_i^{MC}}}$$

where N_i^{MC} is the number of MC events accepted in cross section bin i.
Uncertainties due to MC statistics are not correlated between bins!

Systematic uncertainty due to integrated luminosity measurement:

$$\sigma_{sys(\mathcal{L})} = \boldsymbol{\sigma}_i \cdot \frac{\sigma_{\mathcal{L}}}{\mathcal{L}}$$

is 100% correlated between different cross section measurements (bins).

In general, arbitrary level of correlation (when more than one effect is taken into account) is possible for systematic uncertainties...

Systemactic uncertainties

Extended model

We considered outcome of our experiment y_i as a random variable with given probability density function (usually assumed to be Gaussian)

$$f(y_i) \;=\; G(y_i;\mu_i,\sigma_i)$$

where, in the general case, the uncertainty of the measurement was given by (the square root of) the variance of the distribution:

$$\sigma^2_{(stat)\,i} \;=\; \mathbb{V}(y_i) \;=\; \langle (y_i-\mu_i)^2 \rangle$$

Extended model

We considered outcome of our experiment y_i as a random variable with given probability density function (usually assumed to be Gaussian)

$$f(y_i) = G(y_i; \mu_i, \sigma_i)$$

where, in the general case, the uncertainty of the measurement was given by (the square root of) the variance of the distribution:

$$\sigma^2_{(stat)\,i} = \mathbb{V}(y_i) = \langle (y_i - \mu_i)^2 \rangle$$

This is how we can define the statistical uncertainty: uncertainty of the measurement when the expected value (and other parameters of pdf) are precisely known:

$$\mu_i = \mu(x_i; \mathbf{a})$$

with controlled variable x_i and all model parameters **a** fixed.

Extended model

To describe systematic effects, we need to introduce additional parameters in the model:

$$\mu_i = \mu(x_i; \mathbf{a}, \mathbf{s})$$

where parameters s_j describe different sources of systematic uncertainty.

Extended model

To describe systematic effects, we need to introduce additional parameters in the model:

$$\mu_i \quad = \quad \mu(x_i; \mathbf{a}, \mathbf{s})$$

where parameters s_j describe different sources of systematic uncertainty.

We usually assume some nominal, expected values of these parameters, $\mathbf{s}_0$.
Uncertainties of these parameters, $\sigma_{\mathbf{s}}$, are then what contributes to the systematic uncertainty of our measurements:

$$\mu_i \ = \ \mu(x_i; \mathbf{a}, \mathbf{s}) \quad = \quad \mu(x_i; \mathbf{a}, \mathbf{s}_0) \ + \sum_j \frac{\partial \mu_i}{\partial s_j} \cdot (s_j - s_{0,\,j})$$

$$\mu_i \quad = \quad \mu_0(x_i; \mathbf{a}) \ + \sum_j \frac{\partial \mu_i}{\partial s_j} \, \sigma_{s_j} \cdot \delta_j \qquad \delta_j = \frac{s_j - s_{0,\,j}}{\sigma_{s_j}}$$

where we introduce variations δ_j scaled to unit normal distribution ($\mu = 0$, $\sigma = 1$)

Extended model

Assuming there is no systematic bias in the measurement (or we already corrected for it), averaging over **s** we should get:

$$\mathbb{E}(\mu_i) \quad = \quad \mu_0(x_i; \mathbf{a})$$

Extended model

Assuming there is no systematic bias in the measurement (or we already corrected for it), averaging over **s** we should get:

$$\mathbb{E}(\mu_i) \quad = \quad \mu_0(x_i; \mathbf{a})$$

what follows is: $y_i - \mu_i = (y_i - \mu_{0,\,i}) + (\mu_{0,\,i} - \mu_i)$

$$\mathbb{V}(y_i) \;=\; \langle (y_i - \mu_i)^2 \rangle \quad = \quad \langle (y_i - \mu_{0,\,i})^2 \rangle \;+\; \langle (\mu_i - \mu_{0,\,i})^2 \rangle$$

Extended model

Assuming there is no systematic bias in the measurement (or we already corrected for it), averaging over **s** we should get:

$$\mathbb{E}(\mu_i) \quad = \quad \mu_0(x_i; \mathbf{a})$$

what follows is: $y_i - \mu_i = (y_i - \mu_{0,\,i}) + (\mu_{0,\,i} - \mu_i)$

$$\mathbb{V}(y_i) \;=\; \langle (y_i - \mu_i)^2 \rangle \;=\; \langle (y_i - \mu_{0,\,i})^2 \rangle \;+\; \langle (\mu_i - \mu_{0,\,i})^2 \rangle$$

$$\sigma^2_{y_i} \quad = \quad \sigma^2_{(stat)\,i} \quad + \quad \sigma^2_{(sys)\,i}$$

Extended model

Assuming there is no systematic bias in the measurement (or we already corrected for it), averaging over **s** we should get:

$$\mathbb{E}(\mu_i) \;=\; \mu_0(x_i; \mathbf{a})$$

what follows is: $y_i - \mu_i = (y_i - \mu_{0,\,i}) + (\mu_{0,\,i} - \mu_i)$

$$\mathbb{V}(y_i) \;=\; \langle (y_i - \mu_i)^2 \rangle \;=\; \langle (y_i - \mu_{0,\,i})^2 \rangle \;+\; \langle (\mu_i - \mu_{0,\,i})^2 \rangle$$

$$\sigma^2_{y_i} \;=\; \sigma^2_{(stat)\,i} \;+\; \sigma^2_{(sys)\,i}$$

$$= \; \sigma^2_{(stat)\,i} \;+\; \sum_j \left(\frac{\partial \mu_i}{\partial s_j} \right)^2 \sigma^2_{s_j}$$

where we assume independent sources of systematic variations.

Extended model

Covariance matrix for the series of measurements y_i:

$$\mathbb{C}_{\mathbf{y}} = \mathbb{E}((y_i - \mu_i)(y_j - \mu_j))$$

Extended model

Covariance matrix for the series of measurements y_i:

$$\mathbb{C}_{\mathbf{y}} = \mathbb{E}((y_i - \mu_i)(y_j - \mu_j)) \qquad y_i - \mu_i = (y_i - \mu_{0,\,i}) + (\mu_{0,\,i} - \mu_i)$$

$$= \mathbb{E}((y_i - \mu_{0,\,i})(y_j - \mu_{0,\,j})) + \mathbb{E}((\mu_i - \mu_{0,\,i})(\mu_j - \mu_{0,\,j}))$$

where mixed terms vanish, as systematic variations and statistical fluctuations are independent

Extended model

Covariance matrix for the series of measurements y_i:

$$\mathbb{C}_{\mathbf{y}} = \mathbb{E}((y_i - \mu_i)(y_j - \mu_j)) \qquad y_i - \mu_i = (y_i - \mu_{0,\,i}) + (\mu_{0,\,i} - \mu_i)$$

$$= \mathbb{E}((y_i - \mu_{0,\,i})(y_j - \mu_{0,\,j})) + \mathbb{E}((\mu_i - \mu_{0,\,i})(\mu_j - \mu_{0,\,j}))$$

$$= \mathbb{C}_{\mathbf{y}}^{(stat)} + \mathbb{C}_{\mathbf{y}}^{(sys)}$$

where covariance matrix for statistical uncertainties is diagonal:

$$\mathbb{C}_{ij}^{(stat)} = \begin{cases} \sigma^2_{(stat)\,i} & \text{for } i = j \\ 0 & i \neq j \end{cases}$$

statistical fluctuations are independent

Extended model

Covariance matrix for the series of measurements y_i:

$$\mathbb{C}_{\mathbf{y}} = \mathbb{E}((y_i - \mu_i)(y_j - \mu_j)) \qquad y_i - \mu_i = (y_i - \mu_{0,\,i}) + (\mu_{0,\,i} - \mu_i)$$

$$= \mathbb{E}((y_i - \mu_{0,\,i})(y_j - \mu_{0,\,j})) + \mathbb{E}((\mu_i - \mu_{0,\,i})(\mu_j - \mu_{0,\,j}))$$

$$= \mathbb{C}_{\mathbf{y}}^{(stat)} + \mathbb{C}_{\mathbf{y}}^{(sys)}$$

where covariance matrix for statistical uncertainties is diagonal:

$$\mathbb{C}_{ij}^{(stat)} = \begin{cases} \sigma^2_{(stat)\,i} & \text{for } i = j \\ 0 & i \neq j \end{cases}$$

statistical fluctuations are independent

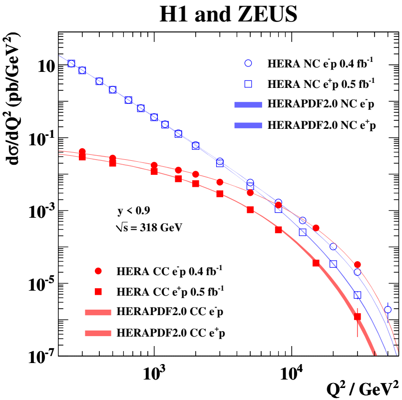
systematic uncertainties result in correlations of expectations:

$$\mathbb{C}_{ij}^{(sys)} = \sum_k \left(\frac{\partial \mu_i}{\partial s_k}\right)\left(\frac{\partial \mu_j}{\partial s_k}\right) \sigma^2_{s_k}$$

We can no longer treat measurements as independent...

Example
SM predictions from HERA

measurement already discussed in lecture 06



NC and CC DIS cross sections comparable for the highest Q^2 values

$$Q^2 \sim M_Z^2,\ M_W^2$$

Combined QCD+EW analysis shows good agreement with SM predictions

Phys. Rev. D 93 (2016) 092002, arXiv:1603.09628

How were systematic uncertainties on the SM predictions calculated?

Example

Let us focus on the "PDF uncertainties", i.e. uncertainties related to our knowledge of the Parton Distribution Functions (PDF) of the proton.

Cross section for NC and CC DIS $e^{\pm}p$ scattering are given in terms of the quark density functions. In the leading order:

$$\frac{d^2\sigma_{CC}^{e^{\pm}p}}{dxdQ^2} = \frac{G_F^2}{4\pi}\left(\frac{M_W^2}{M_W^2+Q^2}\right)^2 \begin{cases} u+c+(1-y)^2(\bar{d}+\bar{s}+\bar{b}) & \text{for } e^-p \\ (1-y)^2(d+s+b)+\bar{u}+\bar{c} & \text{for } e^+p \end{cases}$$

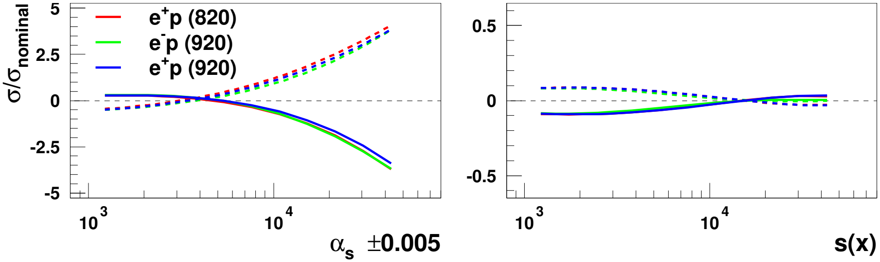
where u, d, s, c, b are quark densities ($\bar{u}, \bar{d} \ldots$ - antiquark) in the proton, extracted by fitting QCD evolution equations to the large set of data from many experiments (not only DIS).

However, one has to take into account uncertainties of the input data, as well as uncertainties related to different assumptions in the fit...

Example

Analysis of HERA data was based on the QCD fit results implemented in EPDFLIB library

It provided not only the nominal parton density values, but also density values corresponding to variations of different "systematic parameters".

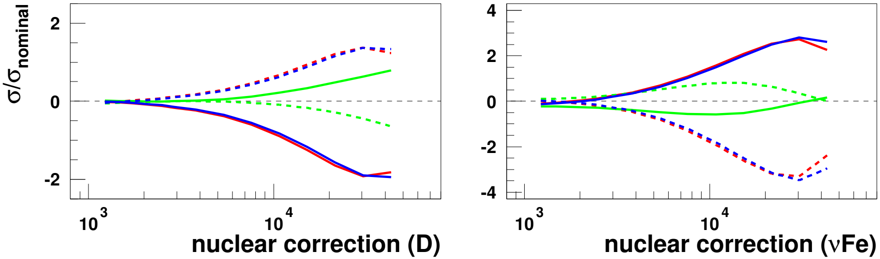


Variations of the three considered NC DIS data sets as a function of Q^2
for up (solid) and down (dashed) variation of systematic parameters (as indicated in labels)

Example

Analysis of HERA data was based on the QCD fit results implemented in EPDFLIB library

It provided not only the nominal parton density values, but also density values corresponding to variations of different “systematic parameters”.

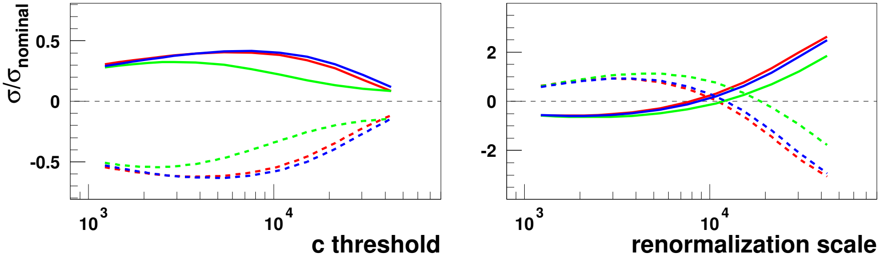


Variations of the three considered NC DIS data sets as a function of Q^2 for up (solid) and down (dashed) variation of systematic parameters (as indicated in labels)

Example

Analysis of HERA data was based on the QCD fit results implemented in EPDFLIB library

It provided not only the nominal parton density values, but also density values corresponding to variations of different "systematic parameters".

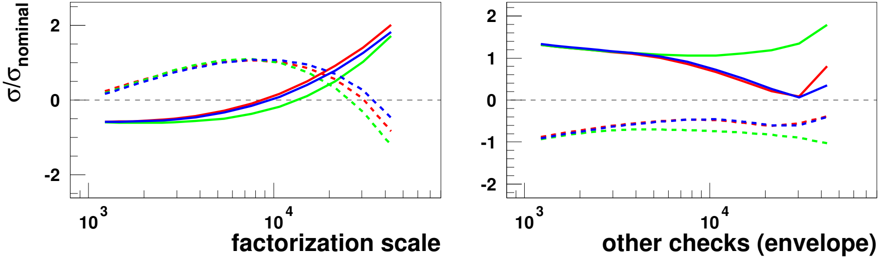


Variations of the three considered NC DIS data sets as a function of Q^2
for up (solid) and down (dashed) variation of systematic parameters (as indicated in labels)

Example

Analysis of HERA data was based on the QCD fit results implemented in EPDFLIB library

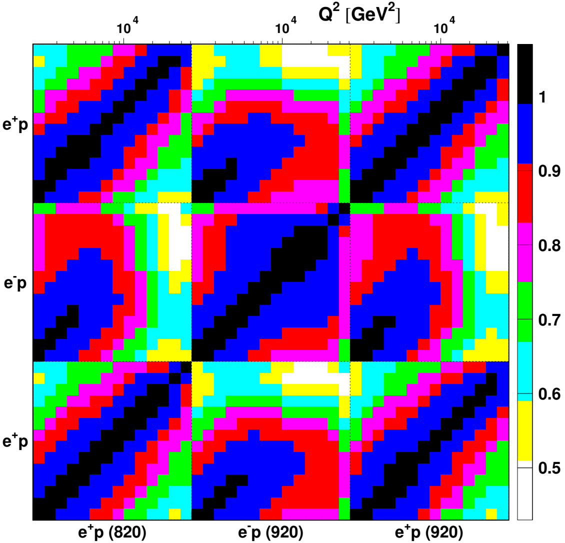
It provided not only the nominal parton density values, but also density values corresponding to variations of different "systematic parameters".



Variations of the three considered NC DIS data sets as a function of Q^2
for up (solid) and down (dashed) variation of systematic parameters (as indicated in labels)

Example

Correlation matrix for the expected high Q^2 NC DIS cross sections:



Must be taken into account when we compare our data to SM predictions

General remarks

One could think that obtaining the proper final result from the analysis (including estimate of the statistical uncertainty) is most important and most difficult. We are almost done...

Estimating systematic uncertainties

General remarks

One could think that obtaining the proper final result from the analysis (including estimate of the statistical uncertainty) is most important and most difficult. We are almost done...

In many cases, proper estimate of systematic uncertainties turn out to be much more difficult and more time consuming than the "nominal study".

This is also because there is no "default solution" to the problem.

One should consider all possible systematic effects, sources of systematic uncertainties, which could affect the measurement. Each parameter you use in your formula or your analysis code should be considered as a potential source of uncertainty.

General remarks

One could think that obtaining the proper final result from the analysis (including estimate of the statistical uncertainty) is most important and most difficult. We are almost done...

In many cases, proper estimate of systematic uncertainties turn out to be much more difficult and more time consuming than the "nominal study".

This is also because there is no "default solution" to the problem.

One should consider all possible systematic effects, sources of systematic uncertainties, which could affect the measurement. Each parameter you use in your formula or your analysis code should be considered as a potential source of uncertainty.

But one should also be careful not to overestimate the uncertainties!
Need to distinguish "systematic variations" and "systematic checks"...

Systematic checks

Usually, there are many parameters in the theoretical model or in the detector descriptions which are known with finite precision. This is the source of systematic uncertainties.
And systematic bias, if our estimates of these parameters are wrong.

Estimating systematic uncertainties

Systematic checks

Usually, there are many parameters in the theoretical model or in the detector descriptions which are known with finite precision. This is the source of systematic uncertainties. And systematic bias, if our estimates of these parameters are wrong.

But there are also many parameters, which we put "by hand", i.e. reconstruction method, selection cuts. There are also many choices, which one can make, i.e. of the proton PDF.

Final result of our analysis should not depend on these choice, if our approach is valid, but some variations can occur.

Systematic checks

Usually, there are many parameters in the theoretical model or in the detector descriptions which are known with finite precision. This is the source of systematic uncertainties. And systematic bias, if our estimates of these parameters are wrong.

But there are also many parameters, which we put “by hand”, i.e. reconstruction method, selection cuts. There are also many choices, which one can make, i.e. of the proton PDF.

Final result of our analysis should not depend on these choice, if our approach is valid, but some variations can occur.

One should be very careful! These variations are often due to the finite MC statistics. One should not include them in the systematic uncertainty estimate.

Otherwise, systematic uncertainties can easily “explode”,
if we use large number of “systematic checks”...

Systemactic uncertainties

Example (Toy model)

An experiment is designed to measure an unknown parameter η.

Two measurements are considered (different experiment configurations) corresponding to two random variables x and y related to the physics parameter η :

$$x_{true} = a + \eta$$
$$y_{true} = a + 2 \cdot \eta$$

where a is the background contribution predicted by theory.
Both variables are measured with the same statistical precision σ_{stat}

Example (Toy model)

An experiment is designed to measure an unknown parameter η.

Two measurements are considered (different experiment configurations) corresponding to two random variables x and y related to the physics parameter η :

$$x_{true} = a + \eta$$
$$y_{true} = a + 2 \cdot \eta$$

where a is the background contribution predicted by theory.
Both variables are measured with the same statistical precision σ_{stat}

We can find the optimum way of extracting η by writing down:

$$\chi^2(\eta) = \left(\frac{x - a - \eta}{\sigma_{stat}}\right)^2 + \left(\frac{y - a - 2\eta}{\sigma_{stat}}\right)^2$$

Example

Looking at the minimum of χ^2 we find:

$$0 = \frac{\partial \chi^2}{\partial \eta} = -\frac{2}{\sigma_{stat}^2}\Big(x - a - \eta + 2(y - a - 2\eta)\Big)$$

Example

Looking at the minimum of χ^2 we find:

$$0 = \frac{\partial\chi^2}{\partial\eta} = -\frac{2}{\sigma^2_{stat}}\Big(x - a - \eta + 2(y - a - 2\eta)\Big)$$

$$\Rightarrow \quad 5\eta = x + 2y - 3a \quad \Rightarrow \quad \eta = \frac{1}{5}\,x + \frac{2}{5}\,y - \frac{3}{5}\,a$$

Example

Looking at the minimum of χ^2 we find:

$$0 = \frac{\partial \chi^2}{\partial \eta} = -\frac{2}{\sigma_{stat}^2} \Big(x - a - \eta + 2(y - a - 2\eta) \Big)$$

$$\Rightarrow \quad 5\eta = x + 2y - 3a \quad \Rightarrow \quad \eta = \frac{1}{5}\, x + \frac{2}{5}\, y - \frac{3}{5}\, a$$

This result is not surprising, if we realize that η can be extracted from x and y independently:

$$\eta(x) = x - a \qquad \text{and} \qquad \eta(y) = \frac{1}{2}(y - a)$$

$$\sigma_{\eta(x)} = \sigma_{stat} \qquad\qquad \sigma_{\eta(y)} = \frac{1}{2}\, \sigma_{stat}$$

and minimum of χ^2 corresponds to the weighted average of the two measurements, with uncertainty: $\sigma_y = \frac{1}{\sqrt{5}}\, \sigma_{stat}$

Example

Let us now include systematic variation Δ_{sys} of the background estimate a, so that the expected results of the measurement are

$$\begin{aligned} \langle x \rangle &= x_{true} + \Delta_{sys} \\ \langle y \rangle &= y_{true} + \Delta_{sys} \end{aligned}$$

Example

Let us now include systematic variation Δ_{sys} of the background estimate a, so that the expected results of the measurement are

$$\langle x \rangle = x_{true} + \Delta_{sys}$$
$$\langle y \rangle = y_{true} + \Delta_{sys}$$

We assume that systematic variation Δ_{sys} has normal distribution with zero mean (unbiased) and width given by σ_{sys}:

$$\Delta_{sys} = \delta \cdot \sigma_{sys}$$

where variation δ has unit normal distribution ($\mu = 0$, $\sigma = 1$)

Example

Let us now include systematic variation Δ_{sys} of the background estimate a, so that the expected results of the measurement are

$$\langle x \rangle = x_{true} + \Delta_{sys}$$
$$\langle y \rangle = y_{true} + \Delta_{sys}$$

We assume that systematic variation Δ_{sys} has normal distribution with zero mean (unbiased) and width given by σ_{sys}:

$$\Delta_{sys} = \delta \cdot \sigma_{sys}$$

where variation δ has unit normal distribution ($\mu = 0$, $\sigma = 1$)

First guess would be to include systematic uncertainty in the final results:

$$\eta = \frac{1}{5}\, x + \frac{2}{5}\, y - \frac{3}{5}\, a \quad \Rightarrow \quad \sigma_y^2 = \frac{1}{25}\sigma_{stat}^2 + \frac{4}{25}\sigma_{stat}^2 + \frac{9}{25}\sigma_{sys}^2$$

Example

Let us now include systematic variation Δ_{sys} of the background estimate a, so that the expected results of the measurement are

$$\langle x \rangle = x_{true} + \Delta_{sys}$$
$$\langle y \rangle = y_{true} + \Delta_{sys}$$

We assume that systematic variation Δ_{sys} has normal distribution with zero mean (unbiased) and width given by σ_{sys}:

$$\Delta_{sys} = \delta \cdot \sigma_{sys}$$

where variation δ has unit normal distribution ($\mu = 0$, $\sigma = 1$)

First guess would be to include systematic uncertainty in the final results:

$$\eta = \frac{1}{5}\,x + \frac{2}{5}\,y - \frac{3}{5}\,a \quad \Rightarrow \quad \sigma_y^2 = \frac{1}{5}\sigma^2_{stat} + \frac{9}{25}\sigma^2_{sys}$$

but is it the optimal procedure?

Example

We should include systematic variation in global likelihood. For Gaussian uncertainties we get:

$$\chi^2(\eta,\delta) \;=\; \left(\frac{x-a-\eta-\delta\sigma_{sys}}{\sigma_{stat}}\right)^2 + \left(\frac{y-a-2\eta-\delta\sigma_{sys}}{\sigma_{stat}}\right)^2 + \delta^2$$

where δ^2 term corresponds to the likelihood of the systematic variation

Example

We should include systematic variation in global likelihood. For Gaussian uncertainties we get:

$$\chi^2(\eta,\delta) = \left(\frac{x-a-\eta-\delta\sigma_{sys}}{\sigma_{stat}}\right)^2 + \left(\frac{y-a-2\eta-\delta\sigma_{sys}}{\sigma_{stat}}\right)^2 + \delta^2$$

We can calculate partial derivatives to get system of equations:

$$\begin{aligned} \frac{\partial\chi^2}{\partial\eta} &: \quad 5\eta \quad + \quad 3\sigma_{sys}\delta = x+2y-3a \\ \frac{\partial\chi^2}{\partial\delta} &: \quad 3\sigma_{sys}\eta + (2\sigma_{sys}^2+\sigma_{stat}^2)\delta = (x+y-2a)\sigma_{sys} \end{aligned}$$

Example

We should include systematic variation in global likelihood. For Gaussian uncertainties we get:

$$\chi^2(\eta,\delta) = \left(\frac{x - a - \eta - \delta\sigma_{sys}}{\sigma_{stat}}\right)^2 + \left(\frac{y - a - 2\eta - \delta\sigma_{sys}}{\sigma_{stat}}\right)^2 + \delta^2$$

We can calculate partial derivatives to get system of equations:

$$\begin{aligned}
\frac{\partial\chi^2}{\partial\eta} &: \quad 5\eta + 3\sigma_{sys}\delta = x + 2y - 3a \\
\frac{\partial\chi^2}{\partial\delta} &: \quad 3\sigma_{sys}\eta + (2\sigma_{sys}^2 + \sigma_{stat}^2)\delta = (x + y - 2a)\sigma_{sys}
\end{aligned}$$

which we can solve to obtain:

$$\eta = \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \quad \text{where} \quad f = \frac{\sigma_{sys}}{\sigma_{stat}}$$

Example

For small systematic uncertainties $(\sigma_{sys} \ll \sigma_{stat}, \ f \ll 1)$

$$\eta \;=\; \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \;\rightarrow\; \frac{2y + x - 3a}{5}$$

and we reproduce previous result.

Example

For small systematic uncertainties $(\sigma_{sys} \ll \sigma_{stat},\ f \ll 1)$

$$\eta \;=\; \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \;\rightarrow\; \frac{2y + x - 3a}{5}$$

and we reproduce previous result.

However, if systematic uncertainties are large $(\sigma_{sys} \gg \sigma_{stat},\ f \gg 1)$

$$\eta \;=\; \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \;\rightarrow\; y - x$$

which also seems natural (systematic uncertainty cancels in $y - x$).

Example

For small systematic uncertainties $(\sigma_{sys} \ll \sigma_{stat}, \ f \ll 1)$

$$\eta \ = \ \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \ \rightarrow \ \frac{2y + x - 3a}{5}$$

and we reproduce previous result.

However, if systematic uncertainties are large $(\sigma_{sys} \gg \sigma_{stat}, \ f \gg 1)$

$$\eta \ = \ \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \ \rightarrow \ y - x$$

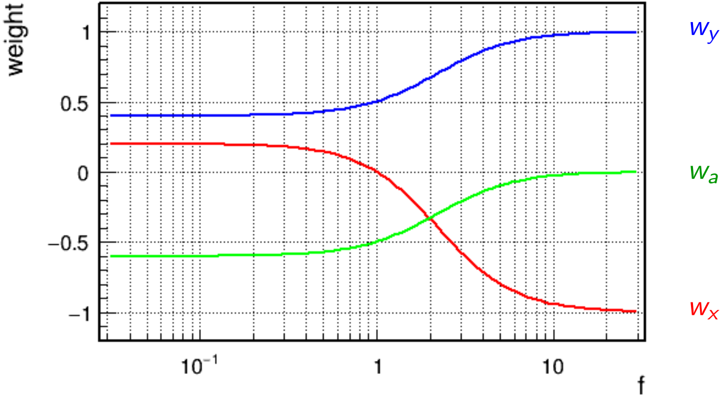
which also seems natural (systematic uncertainty cancels in $y - x$).

It is also interesting to note that for $(\sigma_{sys} = \sigma_{stat}, \ f = 1)$, measurement of x is not used:

$$\eta \ = \ \frac{(2+f^2)y + (1-f^2)x - 3a}{5+f^2} \ = \ \frac{y - a}{2}$$

Example

Weights of the two measurements and background estimate



$$\eta = w_x\, x + w_y\, y + w_a\, a$$

Example

How about uncertainty of the extracted η value?

We can obtain it from the covariance matrix:

$$\mathbb{C}_{(\eta,\delta)} = \left(\frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_l \, \partial a_k} \right)^{-1} = \begin{pmatrix} \frac{5}{\sigma_{stat}^2} & \frac{3\sigma_{sys}}{\sigma_{stat}^2} \\ \frac{3\sigma_{sys}}{\sigma_{stat}^2} & \frac{2\sigma_{sys}^2}{\sigma_{stat}^2} + 1 \end{pmatrix}^{-1}$$

Example

How about uncertainty of the extracted η value?

We can obtain it from the covariance matrix:

$$\mathbb{C}_{(\eta,\delta)} = \left(\frac{1}{2}\frac{\partial^2\chi^2}{\partial a_l\,\partial a_k}\right)^{-1} = \begin{pmatrix} \frac{5}{\sigma_{stat}^2} & \frac{3\sigma_{sys}}{\sigma_{stat}^2} \\ \frac{3\sigma_{sys}}{\sigma_{stat}^2} & \frac{2\sigma_{sys}^2}{\sigma_{stat}^2}+1 \end{pmatrix}^{-1}$$

The resulting uncertainty on η is:

$$\sigma_\eta = \sigma_{stat}\sqrt{\frac{1+2f^2}{5+f^2}}$$

Example

How about uncertainty of the extracted η value?

We can obtain it from the covariance matrix:

$$\mathbb{C}_{(\eta,\delta)} = \left(\frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_l \, \partial a_k} \right)^{-1} = \begin{pmatrix} \frac{5}{\sigma_{stat}^2} & \frac{3\sigma_{sys}}{\sigma_{stat}^2} \\ \frac{3\sigma_{sys}}{\sigma_{stat}^2} & \frac{2\sigma_{sys}^2}{\sigma_{stat}^2} + 1 \end{pmatrix}^{-1}$$

The resulting uncertainty on η is:

$$\sigma_\eta = \sigma_{stat} \sqrt{\frac{1 + 2f^2}{5 + f^2}}$$

$$= \sigma_{stat} \sqrt{\frac{1}{5} + \frac{9f^2}{25 + 5f^2}}$$

Example

How about uncertainty of the extracted η value?

We can obtain it from the covariance matrix:

$$\mathbb{C}_{(\eta,\delta)} = \left(\frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_l \, \partial a_k} \right)^{-1} = \begin{pmatrix} \frac{5}{\sigma_{stat}^2} & \frac{3\sigma_{sys}}{\sigma_{stat}^2} \\ \frac{3\sigma_{sys}}{\sigma_{stat}^2} & \frac{2\sigma_{sys}^2}{\sigma_{stat}^2} + 1 \end{pmatrix}^{-1}$$

The resulting uncertainty on η is:

$$\sigma_\eta = \sigma_{stat} \sqrt{\frac{1+2f^2}{5+f^2}}$$

"simplified approach"

$$= \sigma_{stat} \sqrt{\frac{1}{5} + \frac{9f^2}{25+5f^2}} \leq \sigma_{stat} \sqrt{\frac{1}{5} + \frac{9}{25} f^2}$$

Example

How about uncertainty of the extracted η value?

We can obtain it from the covariance matrix:

$$\mathbb{C}_{(\eta,\delta)} = \left(\frac{1}{2}\frac{\partial^2\chi^2}{\partial a_l\,\partial a_k}\right)^{-1} = \begin{pmatrix} \frac{5}{\sigma_{stat}^2} & \frac{3\sigma_{sys}}{\sigma_{stat}^2} \\ \frac{3\sigma_{sys}}{\sigma_{stat}^2} & \frac{2\sigma_{sys}^2}{\sigma_{stat}^2}+1 \end{pmatrix}^{-1}$$

The resulting uncertainty on η is:

$$\sigma_\eta = \sigma_{stat}\sqrt{\frac{1+2f^2}{5+f^2}} \qquad \text{"simplified approach"}$$

$$= \sigma_{stat}\sqrt{\frac{1}{5}+\frac{9f^2}{25+5f^2}} \quad \leq \quad \sigma_{stat}\sqrt{\frac{1}{5}+\frac{9}{25}f^2}$$

for $f \to \infty$ $\quad \to \sqrt{2}\,\sigma_{stat} \qquad\qquad \to \infty$

General procedure

General procedure for including systematic uncertainties in the analysis is to consider corresponding systematic shifts as additional model parameters

$$\begin{aligned}\mu_i &= \mu(x_i; \mathbf{a}, \mathbf{s}) \\ \chi^2(\mathbf{a}, \mathbf{s}) &= \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}, \mathbf{s}))^2}{\sigma_i^2} + \sum_{k=1}^{K} \frac{(s_k - s_{0,\,k})^2}{\sigma_{s_k}^2}\end{aligned}$$

General procedure

General procedure for including systematic uncertainties in the analysis is to consider corresponding systematic shifts as additional model parameters

$$\mu_i = \mu(x_i; \mathbf{a}, \mathbf{s}) = \mu(x_i; \mathbf{a}')$$

$$\chi^2(\mathbf{a}, \mathbf{s}) = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}, \mathbf{s}))^2}{\sigma_i^2} + \sum_{k=1}^{K} \frac{(s_k - s_{0,\,k})^2}{\sigma_{s_k}^2}$$

$$\chi^2(\mathbf{a}') = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}'))^2}{\sigma_i^2} + \sum_{k=1}^{K} \delta_k^2 \qquad \delta_k = \frac{s_k - s_{0,\,k}}{\sigma_{s_k}}$$

General procedure

General procedure for including systematic uncertainties in the analysis is to consider corresponding systematic shifts as additional model parameters

$$\mu_i = \mu(x_i; \mathbf{a}, \mathbf{s}) = \mu(x_i; \mathbf{a}')$$

$$\chi^2(\mathbf{a}, \mathbf{s}) = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}, \mathbf{s}))^2}{\sigma_i^2} + \sum_{k=1}^{K} \frac{(s_k - s_{0,\,k})^2}{\sigma_{s_k}^2}$$

$$\chi^2(\mathbf{a}') = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}'))^2}{\sigma_i^2} + \sum_{k=1}^{K} \delta_k^2 \qquad \delta_k = \frac{s_k - s_{0,\,k}}{\sigma_{s_k}}$$

If systematic parameters are not independent (are correlated)

$$\chi^2(\mathbf{a}') = \sum_{i=1}^{N} \frac{(y_i - \mu(x_i, \mathbf{a}'))^2}{\sigma_i^2} + \sum_{k,j} (s_k - s_{0,\,k})(s_j - s_{0,\,j})\,(\mathbb{C}_\mathbf{s})^{-1}_{j,k}$$

General procedure

χ^2 minimization procedure is basically unchanged, only the additional terms (systematic constrains) need to be included in calculations (as we did for the parameter constraints).

Negative gradient of χ^2 uncorrelated systematics

$$b_j = -\frac{1}{2}\frac{\partial \chi^2}{\partial a'_j} = \sum_{i=1}^{N} \frac{y_i - \mu_i}{\sigma_i^2} \cdot \frac{\partial \mu_i}{\partial a'_j} - \frac{s_j - s_{0,j}}{\sigma_{s_j}^2}$$

General procedure

χ^2 minimization procedure is basically unchanged, only the additional terms (systematic constrains) need to be included in calculations (as we did for the parameter constraints).

Negative gradient of χ^2 uncorrelated systematics

$$b_j = -\frac{1}{2}\frac{\partial \chi^2}{\partial a'_j} = \sum_{i=1}^{N} \frac{y_i - \mu_i}{\sigma_i^2} \cdot \frac{\partial \mu_i}{\partial a'_j} - \frac{s_j - s_{0,j}}{\sigma_{s_j}^2}$$

Hessian matrix of second derivatives:

$$\mathbb{A}_{jk} = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial a'_j\, \partial a'_k} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \cdot \frac{\partial \mu_i}{\partial a'_j} \cdot \frac{\partial \mu_i}{\partial a'_k} + \frac{\delta_{jk}}{\sigma_{s_k}^2}$$

where systematic shifts **s** are assumed to go first in **a**' (for proper indexing)

General procedure

χ^2 minimization procedure is basically unchanged, only the additional terms (systematic constrains) need to be included in calculations (as we did for the parameter constraints).

Negative gradient of χ^2 general case

$$b_j = -\frac{1}{2}\frac{\partial \chi^2}{\partial a'_j} = \sum_{i=1}^{N} \frac{y_i - \mu_i}{\sigma_i^2} \cdot \frac{\partial \mu_i}{\partial a'_j} - \sum_k (s_k - s_{0,\,k})\,(\mathbb{C}_{\mathbf{s}})^{-1}_{j,k}$$

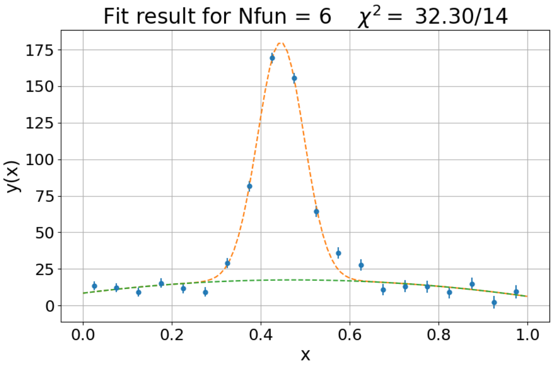
Hessian matrix of second derivatives:

$$\mathbb{A}_{jk} = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial a'_j\,\partial a'_k} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \cdot \frac{\partial \mu_i}{\partial a'_j} \cdot \frac{\partial \mu_i}{\partial a'_k} + (\mathbb{C}_{\mathbf{s}})^{-1}_{j,k}$$

where systematic shifts **s** are assumed to go first in **a**' (for proper indexing)

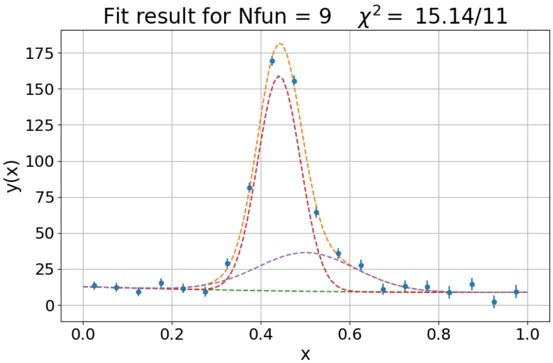
General procedure example

Fitting Gaussian peak on top of background   example from lecture 09



General procedure example

Fitting Gaussian peak on top of background — example from lecture 09

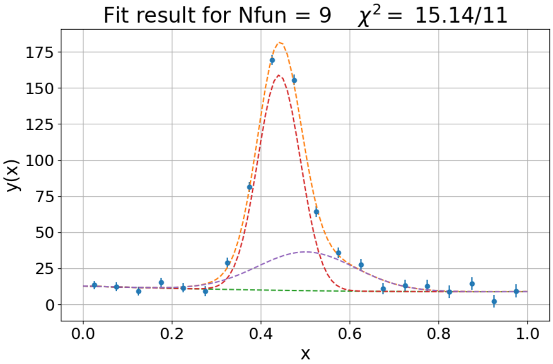


Two peak fit is better, but improvement not very significant, $p = 0.034$ (1.8σ)

General procedure example

Fitting Gaussian peak on top of background

example from lecture 09

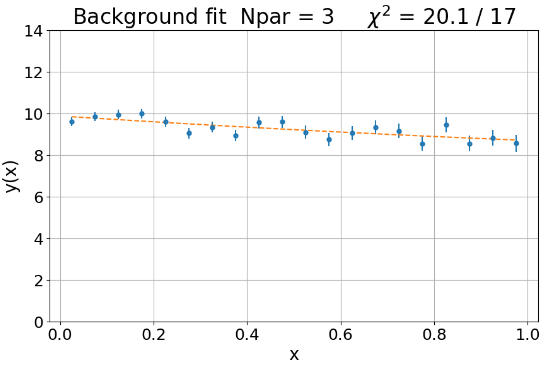


But we also see that background fit changes a lot...

General procedure example

10_bg_fit.ipynb Open in Colab

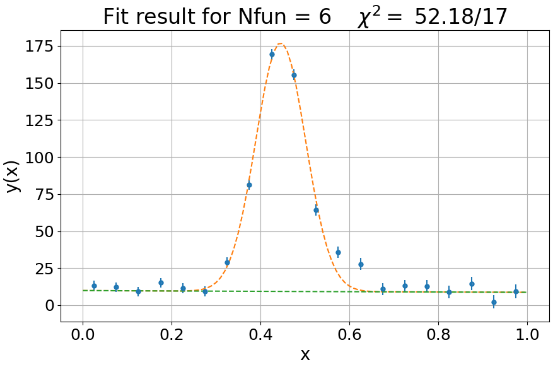
Suppose we can perform an independent background measurement with higher precision and fit parameters of our background model



General procedure example

10_comb_fit.ipynb Open in Colab

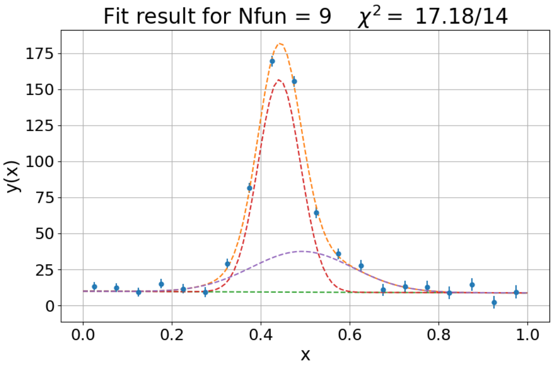
We can now use parameters from the background fit in signal fit



General procedure example

10_comb_fit.ipynb Open in Colab

We can now use parameters from the background fit in signal fit



Second peak significance increase from 1.8σ to 3.1σ ($p = 0.0011$)

Systemactic uncertainties

Problem

In the general case, one can consider a huge number of systematic effects, each will contribute to the final systematic uncertainty.

Number of systematic effects can be larger than the number of relevant model parameters (which we want to extract) or even the number of measurements.

Systematic uncertainties of our measurements are (in most cases) correlated, so one needs to use the full covariance matrix.

Problem

In the general case, one can consider a huge number of systematic effects, each will contribute to the final systematic uncertainty.

Number of systematic effects can be larger than the number of relevant model parameters (which we want to extract) or even the number of measurements.

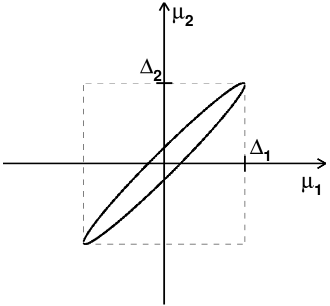
Systematic uncertainties of our measurements are (in most cases) correlated, so one needs to use the full covariance matrix.

Is there a way to simplify the problem?

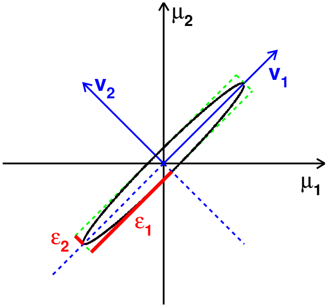
Is there a way to reduce the number of systematic variations to consider?
This is also important when we want to model the experiment (eg. with Monte Carlo methods)

Eigenvectors

Correlations between variables can be removed througth 'rotation' in the variable space.



$\Rightarrow$



This is a problem of finding "eigenvectors" of the covariance matrix. Directions such that:

$$\mathbb{C}_{\mathbf{s}} \cdot \mathbf{v} \quad = \quad \sigma_{\mathbf{v}}^2 \cdot \mathbf{v}$$

Eigenvectors

Eigenvectors of the covariance matrix of systematic parameters define "uncorrelated directions" in the space of systematic parameter variations.

Variations along these directions are independent (uncorrelated).
We can redefine our systematic variables to remove correlations...

Eigenvectors

Eigenvectors of the covariance matrix of systematic parameters define "uncorrelated directions" in the space of systematic parameter variations.

Variations along these directions are independent (uncorrelated).
We can redefine our systematic variables to remove correlations...

Eigenvalues

Eigenvalues give us the size of variations along given eigenvector ($\sigma^2_{\mathbf{v}}$)
$\Rightarrow$ we can tell which variations are most relevant

By identifying variations which give leading contributions to the covariance matrix, we can limit number of variations considered in our problem.

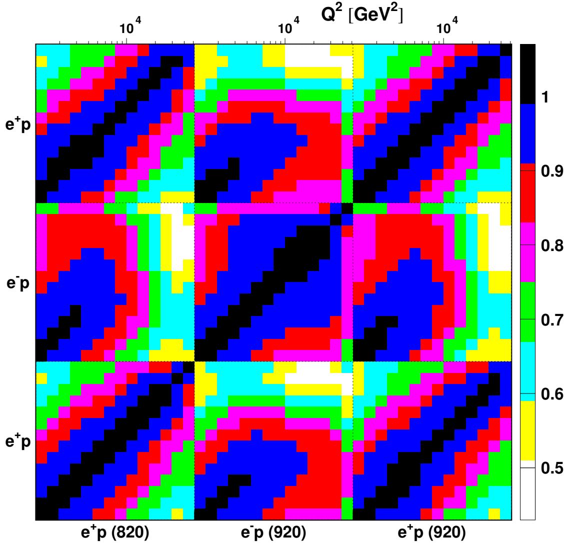
Variations corresponding to eigenvectors with very small eigenvalues can be safely ignored...

In the following, we assume eigenvectors are ordered from highest to lowest eigenvalue.

Eigenvectors

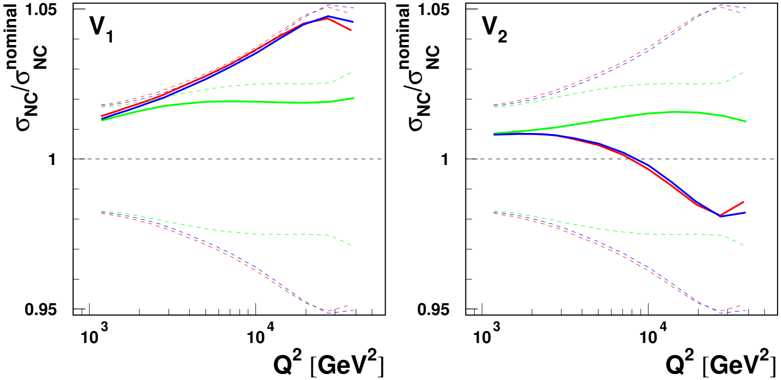
Let us consider uncertainties of the high Q^2 NC DIS cross sections again.

Correlation matrix:



Eigenvectors

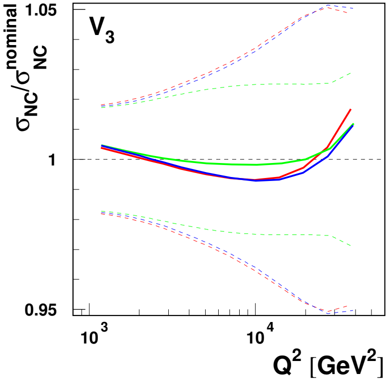
Systematic variations corresponding to eigenvectors of correlation matrix:

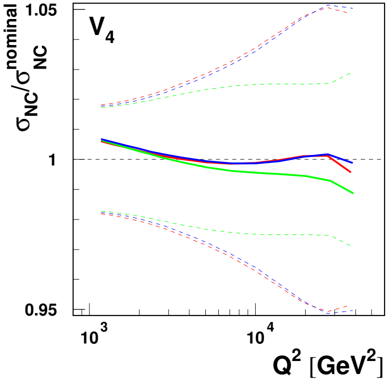


Dominant contribution from the first eigenvector... (variations relative to the nominal SM exp.)

Eigenvectors

Systematic variations corresponding to eigenvectors of correlation matrix:

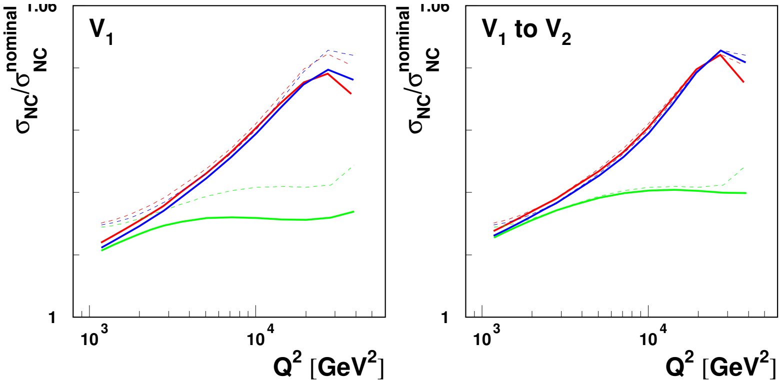




Dominant contribution from the first eigenvector... (variations relative to the nominal SM exp.)

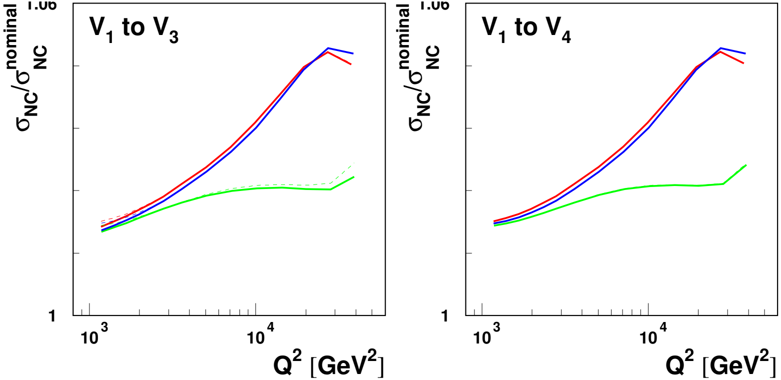
Eigenvectors

Variations corresponding to the sum of eigenvector contributions:



Eigenvectors

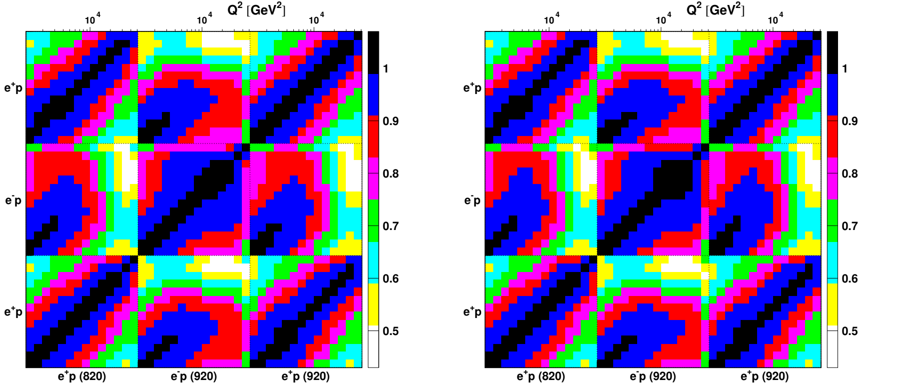
Variations corresponding to the sum of eigenvector contributions:



Four first eigenvectors perfectly reproduce total systematic uncertainty

Eigenvectors

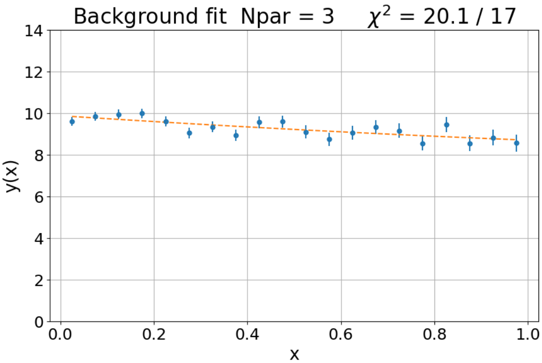
Correlation matrix comparison: Full matrix Four eigenvectors



Correlations between PDF variations also very well reproduced by the first four eigenvectors

Eigenvector example

Previous example of systematic uncertainties: fit to background only measurements

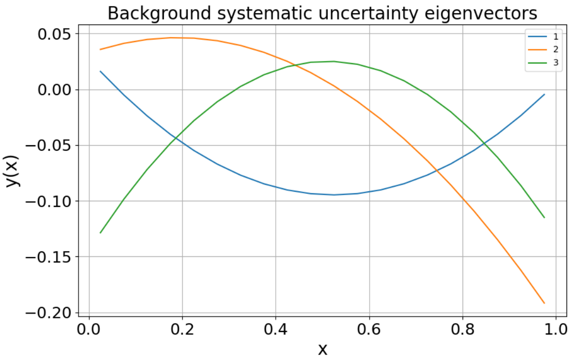


Fitted background expectations are correlated between points. 20×20 correlation matrix...

Eigenvector example

10_bg_fit_ev.ipynb Open in Colab

Previous example of systematic uncertainties: fit to background only measurements

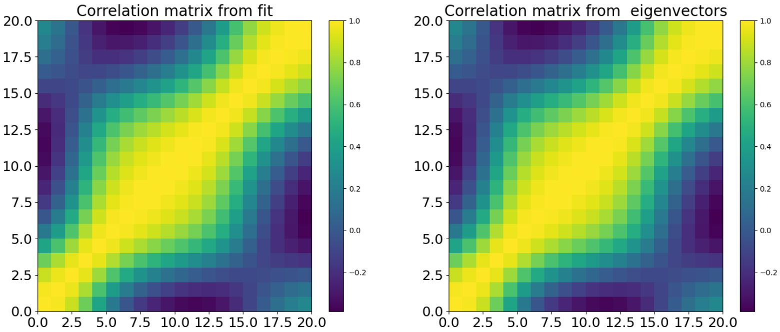


Background uncertainties can be presented as independent (!) variations along 3 eigenvectors

Eigenvector example

10_bg_fit_ev.ipynb Open in Colab

Previous example of systematic uncertainties: fit to background only measurements



Although variations along considered eigenvectors are independent (by definition), systematic correlations between measurements well reproduced...

Systemactic uncertainties

Homework

Solutions to be uploaded by December 18.

Modification of the "toy model" presented today.
Electronic scale is used to weigh a cake with mass M.
To do the actual measurement, we place a cake on a dish with mass m, so we measure:

$$x = m + M$$

We then repeat the measurement but with two (identical) dishes under the cake:

$$y = 2m + M$$

The statistical uncertainty of each measurement is $\sigma_x = \sigma_y = \sigma$,
weight m of the dish is known from independent measurement with uncertainty of $\sigma_m = f \cdot \sigma$.

1. Try to guess the formula for M for $f \to 0$ and $f \to \infty$
2. Find the optimal (most precise) estimate of M (for arbitrary f). Compare to guess 1!
3. Find the uncertainty of the M estimate. Discuss the result.