united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

ICTP 40th Anniversary

SMR.1555 - 39

Workshop on
Nuclear Reaction Data and Nuclear Reactors:
Physics, Design and Safety

16 February - 12 March 2004

Nulcear Data Evaluation

Mihaela SIN
University of Bucharest
Nuclear Physics Department
Magurele MG-11
76900 Bucharest
ROMANIA

NUCLEAR DATA EVALUATION

Mihaela SIN

Nuclear Physics Department, University of Bucharest

msin@brahms.fizica.unibuc.ro, msin@pcnet.ro

Main topics:

- definitions and terminology
- classic and Bayesian statistics
- covariance matrix formalism
- generalized least squares method

Nuclear data

- quantitative results of any scientific investigation of the nuclear properties of matter;
- describe properties of atomic nuclei and the fundamental physical relationships governing their interactions;
- characterize physical processes underlying all nuclear technologies.

Main activities - basic nuclear data production
- nuclear data evaluation
- processing, validation and benchmarking
- applications

- nuclear data evaluation has as study object the basic data and as purpose to recommend the best data to be used in applications; therefore it can not be treated independently, but in connection with the other activities.

Evaluation

- the recommendation of the “best estimate” values and their uncertainties, based on a critical review of all the available information for a particular nuclide (experimental measurements and uncertainties, theoretical predictions) and on the use of statistical procedures (average, fit, inter- and extrapolations, etc.)

The result of a measurement is only and approximation or an estimation of the specific quantity subject to measurement, and thus the result is complete only when accompanied by a quantitative statement of its uncertainty.

EXFOR, reports, articles…

evaluated data files: ENDF, ENSDF…

Evaluated data files

- the evaluated data files represent the quantitative link between two huge fields of research: fundamental nuclear physics and nuclear applications.

- dark age – eye guide curve among experimental data;
- Renaissance – statistical methods to analyze the experimental data and the theoretical constraints;
- modern times - information about uncertainties have been included in evaluated data libraries.

Probability and statistics, expectations and moments….

- probability theory is a branch of mathematical sciences that provides a model for describing the process of observation;
- most observations of natural phenomena do not lead to uniquely predictable results;
-a natural phenomenon would be completely described if sufficient information were available to determine the underlying probability distribution;
-in practice the features of the probability distribution underlying the physical phenomenon under consideration must be estimated;
- such estimations form the study object of statistics.

Mean $$m = E[x] = \int x\,P(x)\,\mathrm{d}x$$

Variance $$s^2 = E[(x-m)^2] = \int (x-m)^2\,P(x)\,\mathrm{d}x$$

Covariance $$\mathrm{cov}(x,y) = E[(x-m_x)(y-m_y)]$$

Correlation $$k = \frac{\mathrm{cov}(x,y)}{s_x s_y}$$

Definitions and terminology ISO Guide (1)

Uncertainty – the parameter associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand (represented by an estimated standard deviation, termed standard uncertainty equal to the positive square root of the variance).

Error (of a measurement) **–** result of a measurement minus the value of the measurand (the deviation of the result of a particular measurement from the unknown true value of the measurand).

- random error – result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions.

- systematic error – mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus the value of the measurand.

Definitions and terminology ISO Guide (2)

Accuracy (of measurement) – closeness of the agreement between the result of measurement and the value of the measurand.

Repeatability (of results of measurements) – closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement.

Reproducibility (of results of measurements) – closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement.

Sources of measurement uncertainty in ISO Guide

1. Incomplete definition of the measurand;
2. Imperfect realization of the definition of the measurand;
3. Non-representative sampling – the sample measured may not represent the defined measurand;
4. Inadequate knowledge of the effects of environmental conditions on the measurement, or imperfect measurement of environmental conditions;
5. Personal bias in reading analogue instruments;
6. Finite instrument resolution or discrimination threshold;
7. Inexact values of measurement standards and reference materials;
8. Inexact values of constants and other parameters obtained from external sources and used in the data reduction algorithm;
9. Approximations and assumptions incorporated in the measurement method
10. Variations in repeated observations of the measurand under apparently identical conditions

1-9 related to systematic effects and 10 to statistical effects

- the separation is not so sharp; if the various effects change during the time of measurement, without any possibility of monitoring them, they contribute to the random error.

Classification of components of uncertainty in ISO Guide

Two categories according to the method used to estimate their numerical values:

A. Those which are evaluated by statistical methods
B. Those which are evaluated by other means

The nature of an uncertainty component is conditioned by the use made of the corresponding quantity, that is, on how quantity appears in the mathematical model that describe the measurement process.

Alternative nomenclature:

- component of uncertainty arising from a random (systematic) effect
- a random (systematic) effect is one that gives rise to a possible a random (systematic) error in the current measurement process.

Type A evaluation of standard uncertainty: may be based on any valid statistical method for treating data.

Type B evaluation of standard uncertainty: is used on scientific judgment using all the relevant information available.

Aspects of evaluation (1)

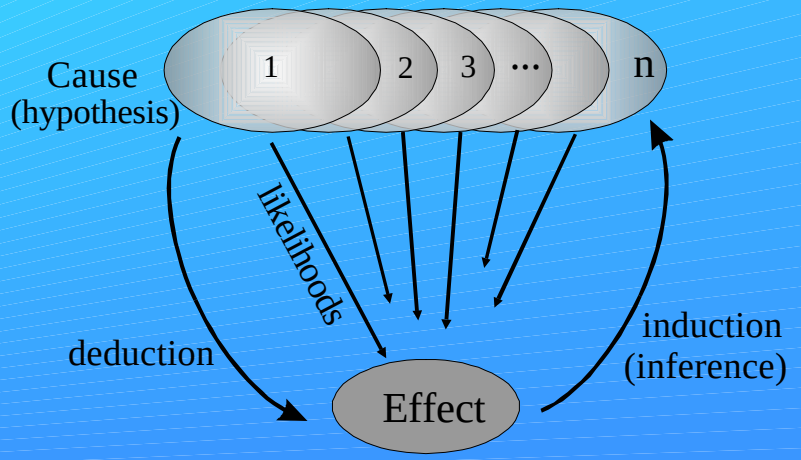
Analysis of the experimental data

Scientific experiments are usually describable by a statistical model, statistical elements being introduced by uncontrollable, seemingly random instrumental effects, by unknown errors and often by theory itself.

- **direct measurement** - of the value *x* of the physical quantity *X ;* effects' prediction given the causes; direct probability, deduction;

- **indirect measurement** - to derive the value *x* of the physical quantity *X* related to the directly measurable quantity *Y* by *Y=Y(X);* causes' prediction given the effects; inverse probability; inference;

- **test of hypothesis** - to confirm or to infirm a theoretical model.



Evaluation up-dating

Aspects of evaluation (2)

(i) Nuclear data evaluation consists in finding the closest values to the true values starting from experimental and theoretical basic data;

(ii) These data are limited and affected by uncertainties;

(iii) Evaluators have to take decisions based on incomplete information.

Interpretation of probability

1. Probability as a relative frequency

2.Probability as a a degree of rational expectation on a numerical scale ranging from 0 (impossibility) to 1 (certainty); degree of plausibility

Bayesian statistics (1)

Bayes' theorem (1763)

Direct consequence of the basic sum and product rules of probability theory:

$$P(A|B)+P(\overline{A}|B)=1$$

$$P(AB|C)=P(A|BC)P(B|C)=P(B|AC)P(A|C)$$

Relates direct probabilities (of effects given the causes) and inverse probabilities (of causes given the effects):

$$P(A|BC)=\frac{P(B|AC)P(A|C)}{P(B|C)}$$

Laplace (1812) generalized it to the case of several distinct , mutually exclusive alternatives A_i:

$$P(A|BC)=\frac{P(B|A_iC)P(A_i|C)}{\sum_i p(B|A_iC)\,p(A_i|C)} \qquad p(A|BC)dA=\frac{p(B|AC)\,p(A|C)dA}{\int p(B|AC)\,p(A|C)dA}$$

Results considered as the very cornerstone of data evaluation

Bayesian statistics (2)

We are interested in the quantity *A* (physical quantity, condition, hypothesis)
Knowledge about *A* is summarized in the *a priori* probability (prior) *P(A|C)*, the probability for *A* to have a certain value in the circumstances *C*.

A new set of experimental information about B which depends on A becomes available. The information is included in the likelihood function *P(B|AC)*, the probability to obtain a certain value *B*, for given *A* and *C*.

Our up-dated information about A is given by the *a posteriori* probability calculated as the product of likelihood function and *a priori* probability, representing the impact of the new information on what we already knew about A.

$$P(A|BC) \propto P(B|AC)P(A|C)$$

Bayesian statistics (3)

Recommended values

The users are not interested in *a posteriori* distribution but in a recommended value and its uncertainty. Automatically we calculate the expectation value and the standard deviation. WHY?

In decision theory is a penalty for bad estimates described by a loss function; it vanishes for the true value and is positive everywhere else.
Usually in the vicinity of the true value the loss function is taken as quadratic in error. The estimation for A which minimizes the loss is the expectation value.
The penalty corresponding to the expectation value is just the variance.

Under quadratic loss the mean and the square root of the variance (standard uncertainty) are the optimal estimates of the (unknown) true value and its uncertainty. These are therefore the numerical values to be given by experimentalists in their documentation or to be put in the data file by evaluators.

Bayesian statistics (4)

A priori distributions

The arbitrariness, subjectivity of the priors, has led many statisticians, for more than a century, to repudiate the Bayesian approach to parameter estimation and to seek alternative methods to circumvent priors.

H. Jeffreys (1939) invoked invariance arguments to find priors which avoided ambiguities.
E.T.Jaynes (1968) applied group theory and information theory to the problem of priors.

Least informative priors (group theory)

He demonstrated for simple but practically important cases, that even if one is completely ignorant about the numerical value of the estimated parameters, the symmetry of the problem, the invariance under a group of transformations, determines the prior unambiguously.

-location parameter (ex. the mean of a Gaussian) – invariant under a shift of location

$$p(m)dm = p(m+c)d(m+c) \qquad p(m)dm \propto dm \quad -\infty < m < \infty$$

-scale parameter (ex. the standard deviation of a Gaussian) – invariance under rescaling

$$p(s)ds = p(cs)d(cs) \qquad p(s)ds \propto \frac{ds}{s} \quad 0 < s < \infty$$

Valid for any rate constants which multiplies time intervals in a problem.

Bayesian statistics (5)

Example: the determination of the decay constant *?* of a short-lived isotope from decays registered at $t_1 \ldots t_n$

? – *A;* decays registered at $t_1 \ldots t_n$ - *B;*

The applicability of the exponential decay law, purity of the sample, reliability of instruments - C

The statistical model – the sampling distribution; the probability with which one may reasonably expect the various alternatives if one sample once, given the parameters of the model (given *?*, the probability that one particular decay is registered in a particular time interval):

$$p(t_i \mid \lambda)dt_i = \exp(-\lambda t_i)\,\lambda\, dt_i \qquad 0 < t_i < \infty$$

According to the product rule, the joint probability of observing the mutually independent data:

$$p(t_1,\ldots t_n \mid \lambda)dt_1 \ldots dt_n = \exp\left(-\lambda \sum_{i=1}^{n} t_i\right)\lambda^n\, dt_1 \ldots dt_n$$

The likelihood function does not depend on all the individual sample values; given *n*, the sample average $\bar{t}$ carries all the information contained in the data:

Bayesian statistics (6)

Example (2)

The likelihood function: $p(n\bar{t} \mid \lambda) = \exp(-\lambda n\bar{t})\lambda^n$

The least informative *a priori* distribution for the scale parameter *?*:

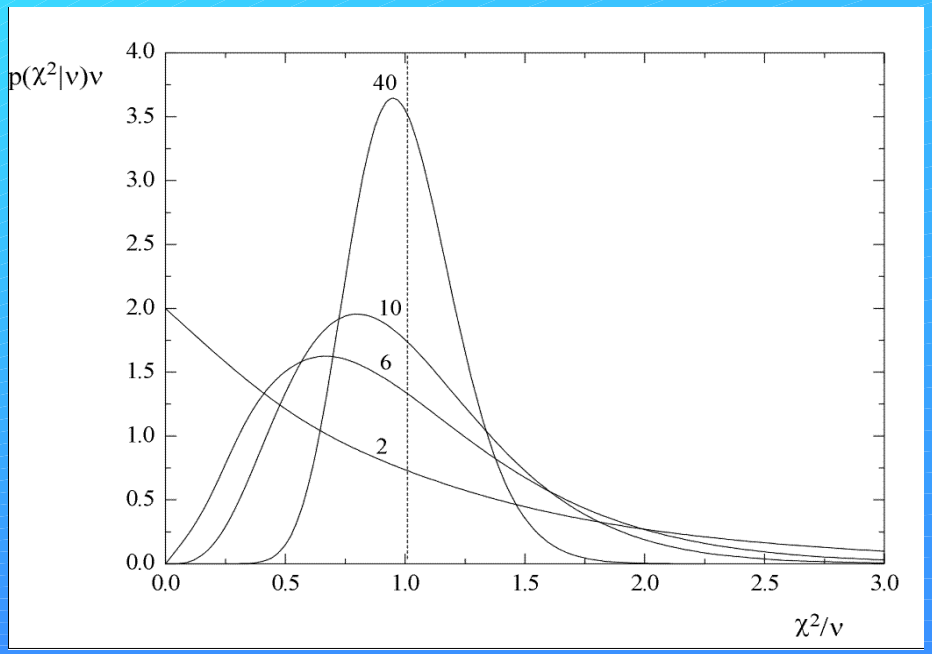
$$p(\lambda)d\lambda \propto \frac{d\lambda}{\lambda} \quad 0 < \lambda < \infty$$

The normalized *a posteriori* distribution:

$$p(\lambda \mid n\bar{t})d\lambda = \Gamma(n)^{-1} \exp(-\lambda n\bar{t})\lambda^{n-1} d\lambda \qquad 0 < \lambda < \infty$$

This chi-square distribution with *? = 2n* degrees of freedom represents the complete information about *?* which can be obtained from the data and the assumed prior.

As the sample size increases the *a posteriori* distribution gets narrower; the more data available, the better defined is *?*.



Bayesian statistics (7)

Example (3)

Recommended values

- the expectation value

$$\langle l \rangle = \int_0^\infty l \; p(l \mid n\bar{t})\, dl = \frac{1}{\bar{t}}$$

- the standard deviation

$$\Delta l = \left[\int_0^\infty \left(l - \langle l \rangle \right)^2 p(l \mid n\bar{t})\, dl \right]^{1/2} = \frac{1}{\bar{t}\sqrt{n}}$$

Bayesian statistics (8)

A priori distributions

Assignment of probabilities by entropy maximization (information theory)

Jaynes (1968-80) showed how probability can be assigned in a well defined way if at least vague information is available about average quantities. Ex., we don't know p(x), but we have global information in the form of expectation values for several known functions:

$$\langle f_k \rangle = \int f_k(x)\, p(x)\, dx \qquad k = 1,2,\ldots, K$$

The key concept: **information entropy** (C.E.Shannon 1948) as the unique measure of the indeterminacy or missing information implied by a given probability distribution:

$$S = -\int p(x)\, \ln[p(x)]\, dx$$

What is p(x) which satisfies the K equations without implying other information or assumptions? The answer is given by the **principle of maximal entropy.**

Bayesian statistics (9)

A priori distributions

Assignment of probabilities by entropy maximization (information theory)

To obtain the probability compatible with the given information, it must be varied in such a way that its entropy is maximized, subject to K constraints.
This variational problem can be solved using Lagrange multipliers technique.

Conclusion:
If only the first moments of a distribution are given, the optimal probability distribution for further inference is Gaussian, no matter what the unknown true distribution may be.

Bayesian statistics (10)

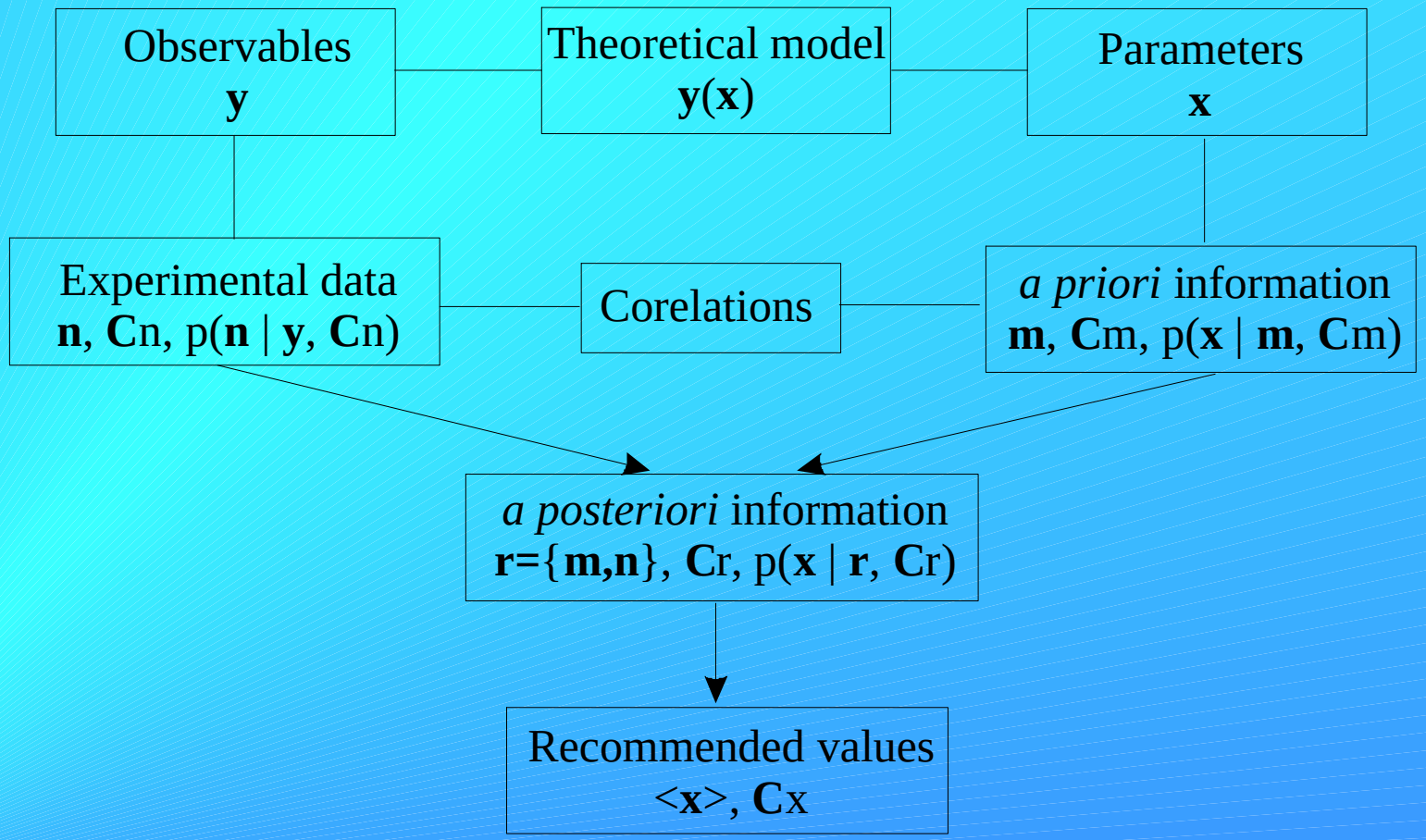
Maximum likelihood approximation

The prior distribution looses importance when new data are abundant, therefore is reasonable to use a constant *a priori* probability. The *a posteriori* probability density becomes equal to the likelihood function.

Maximum likelihood function method consists of recommending that parameter value which maximizes the likelihood function.

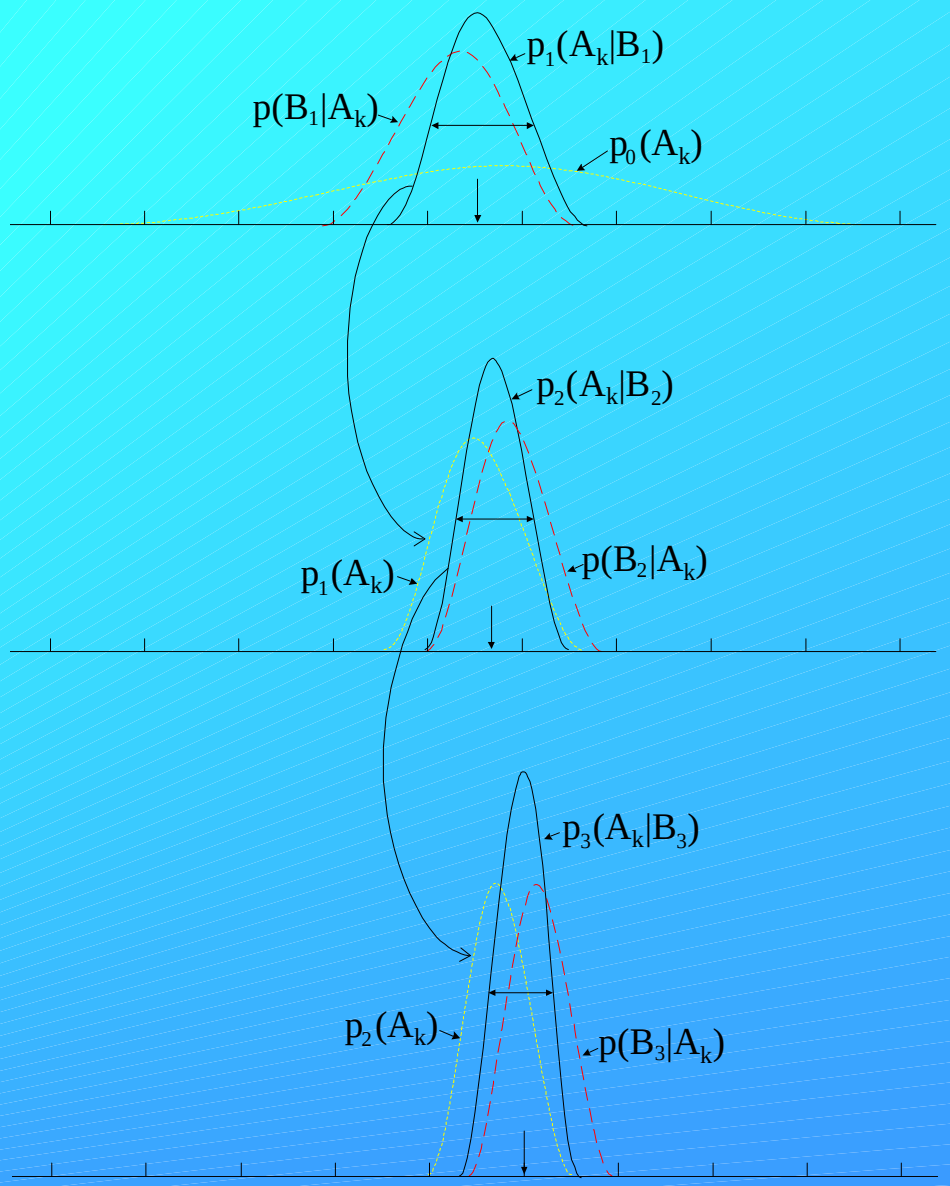
Bayesian statistics (11)

Least-squares approximation



Bayesian statistics (12)

Learning from experience
Evaluation up-dating



Bayesian statistics (13)

Bayes' theorem → a posteriori distribution

Decision theory → how to recommend values and uncertainties

Group theory → use of invariances for probability assignment

Information theory → probability assignment by entropy maximisation

Bayesian approach lead to a concise and mathematically simple treatment parameter estimation and data adjustment in the general framework of inductive inference of learning from real (error-affected, incomplete) observations.

“Although this Guide provides a framework for assessing uncertainty, it cannot substitute for critical thinking, intellectual honesty and professional skill. The evaluation of uncertainty is neither a routine task nor a purely mathematical one; it depends on detailed knowledge of the nature of the measurand and of the measurement. The quality and utility of the uncertainty quoted for the result of a measurement therefore depend on the understanding critical analysis and integrity of those who contribute to the assignment of the value.”

ISO Guide

value

uncertainty

Least-squares method

Covariance matrix formalism

Covariance matrix formalism (1)

Table of uncertainties

x	Components of uncertainty				Total uncertainty
x_1	e_{11}	…	e_{1l} …	e_{1L}	E_{X1}
x_i	e_{in}	…	e_{il} …	e_{iL}	E_{Xi}
x_n	e_{1n}	…	e_{1l} …	e_{nL}	E_{Xn}

- an experiment is described by a set of *n* measurable quantities **x** (x_1, x_2,…,x_n) called ***experimental parameters***;

-$K_{ij,l}$ correlation parameter relating e_{il} and e_{jl}

- element of the covariance matrix

$$C_{x_{ij}} = \sum_{l=1}^{L} K_{ij_l} e_{i_l} e_{j_l}, \quad (i, j = 1, n)$$

- total uncertainty of parameter x_i

$$E_{x_i}^2 = C_{x_{ii}} = \sum_{l=1} K_{ii_l} e_{i_l} e_{i_l} = \sum_{l=1}^{L} e_{i_l}^2, (i = 1, n)$$

- correlation coefficient for the total uncertainties of x_i and x_j

$$K_{x_{ij}} = \frac{C_{x_{ij}}}{E_{x_i} E_{x_j}}, (i, j = 1, n)$$

Covariance matrix formalism (2)

Propagation of uncertainty
- covariance matrix of the derived quantities

- outcome of the experiment - scalar $y = y(\{x_i\}),\ i = 1, n$

- sensitivity matrix **S** (n x n)-dimensional, diagonal $S_i = \frac{\partial y}{\partial x_i}, \quad (i = 1, n)$

- correlation matrix **K** (n x n)-dimensional,

- matrix of parameters' uncertainty $\mathbf{E}_x$ (n x 1)-dimensional

$$E_y^2 = (\mathbf{S} \cdot \mathbf{E}_x)^T \cdot \mathbf{K} \cdot (\mathbf{S} \cdot \mathbf{E}_x)$$

- covariance matrix of parameters $\mathbf{C}_x$ (n x n)-dimensional

$$E_y^2 = (\mathbf{S} \cdot \mathbf{I})^T \cdot \mathbf{C}_x \cdot (\mathbf{S} \cdot \mathbf{I})$$

Covariance matrix formalism (3)

Propagation of uncertainty

Example: scalar output of an experiment depends on two independent experimental parameters x_1, x_2, with total uncertainties E_{x1}, E_{x2}

$$E_y^2 = (\mathbf{S}\cdot\mathbf{E}_x)^T \cdot \mathbf{K} \cdot (\mathbf{S}\cdot\mathbf{E}_x)$$

$$E_y^2 = \left[\begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}\cdot\begin{pmatrix} E_{x_1} \\ E_{x_2} \end{pmatrix}\right]^T \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \left[\begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}\cdot\begin{pmatrix} E_{x_1} \\ E_{x_2} \end{pmatrix}\right] =$$

$$= \begin{pmatrix} S_1E_{x_1} & S_2E_{x_2} \end{pmatrix}\cdot\begin{pmatrix} S_1E_{x_1} \\ S_2E_{x_2} \end{pmatrix} = \left(S_1E_{x_1}\right)^2 + \left(S_2E_{x_2}\right)^2$$

$$s_y^2 = \sum_{i=1}^{2}\left(\frac{\partial y}{\partial x_i}\right)^2 s_{x_i}^2$$

Covariance matrix formalism (4)

Propagation of uncertainty

The derived quantity is the vector **y**=**y**(**x**), $\{y_k\}$ $k=1,m$

m sensitivity matrices $\mathbf{S}_k$; $S_{k_i} = \frac{\partial y_k}{\partial x_i}, \quad (i=1,n;\ k=1,m)$

$$C_{y_{kl}} = (\mathbf{S}_k \cdot \mathbf{E}_x)^T \cdot \mathbf{K}_x \cdot (\mathbf{S}_l \cdot \mathbf{E}_x), \quad (k,l=1,m)$$

$$C_{y_{kl}} = (\mathbf{S}_k \cdot \mathbf{I})^T \cdot \mathbf{C}_x \cdot (\mathbf{S}_l \cdot \mathbf{I})$$

$$E_{y_k} = (C_{y_{kk}})^{1/2}, \quad (k=1,m)$$

$$K_{y_{kl}} = \frac{C_{y_{kl}}}{E_{y_k} E_{y_l}}, \quad (k,l=1,m)$$

Covariance matrix formalism (5)

Propagation of uncertainty

Transformation matrix **T**

$$T_{ik} = \frac{\partial y_i}{\partial x_k}, \quad (k = 1, n;\ i = 1, m)$$

$$C_{y_{kl}} = \sum_{i=1}^{n} \sum_{j=1}^{n} T_{ik} C_{x_{ij}} T_{jl}, \quad (k, l = 1, m;\ i,\ j = 1, n)$$

$$\mathbf{C}_y = \mathbf{T}^T \cdot \mathbf{C}_x \cdot \mathbf{T}$$

Covariance matrix formalism (6)

1. Find the experimental parameters
2. Find the uncertainty components
3. Establish the correlations between the uncertainty components of the parameters
4. Calculate the covariance matrix for the parameters
5. Calculate the correlations between the total uncertainties of the parameters
6. Find the relationship between the parameters and the derived quantities
7. Calculate the sensitivity matrix for each derived quantity (or the transformation matrix)
8. Calculate the covariance matrix for the derived quantities
9. Calculate the uncertainties and the correlations for the derived quantities

Covariance matrix formalism (7)

Example (Mannhart)

Mr.A has to determine 2 markings on a length scale of distances from a fixed 0 point: y1=35 mm, y2=60 mm using 3 gauge blocks :

gauge block	length (mm)	s (µm)	var(µm)^2
L1	50	0.05	0.0025
L2	15	0.03	0.0009
L3	10	0.02	0.0004

y1=L1-L2
y2=L1+L3

Mr.A states his final results as:

y1=35 mm var(y1)=var(L1)+var(L2)=0.0034 µm^2

y2=60 mm var(y2)=var(L1)+var(L3)=0.0029 µm^2

Mr.B and Mr.C have to establish the distance between the two marks.

Mr. B y3= L2+L3=25mm var(y3)=var(L2)+var(L3)=0.0013 µm^2

Mr. C y3= y2 - y1=25mm var(y3)=var(y2) + var(y1)=0.0063 µm^2

Covariance matrix formalism (8)

The whole information requested from Mr.A : $\mathbf{y}$, $\mathbf{C}_y$

$$C_{y_{kl}} = (\mathbf{S}_{yi} \cdot \mathbf{I})^T \cdot \mathbf{C}_x \cdot (\mathbf{S}_{yj} \cdot \mathbf{I}) \quad i, j = 1,2$$

$$C_{y11} = \begin{pmatrix} \frac{\partial y_1}{\partial L_1} & \frac{\partial y_1}{\partial L_2} & \frac{\partial y_1}{\partial L_3} \end{pmatrix} \cdot \begin{pmatrix} E_{l1}^2 & 0 & 0 \\ 0 & E_{l2}^2 & 0 \\ 0 & 0 & E_{l3}^2 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial y_1}{\partial L_1} \\ \frac{\partial y_1}{\partial L_1} \\ \frac{\partial y_1}{\partial L_1} \end{pmatrix}$$

$$\mathbf{C}_y = \begin{pmatrix} E_{L1}^2 + E_{L2}^2 & E_{L1}^2 \\ E_{L1}^2 & E_{L1}^2 + E_{L2}^2 \end{pmatrix} = \begin{pmatrix} 0.0034 & 0.0025 \\ 0.0025 & 0.0029 \end{pmatrix}$$

Least-squares method (1)

y set of experimental values for **Y** set of physical quantities

f**(x)** is the function relating the measurable quantities **Y** to the quantities of interest **X** of values **x**

$$f_k(x) = \sum_{i=1}^{n} A_{ki}\ x_i, \quad (k = 1, m)$$

$\mathbf{f}(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$ **A** design matrix

$\mathbf{y} \approx \mathbf{A} \cdot \mathbf{x}$

Classical least-squares method is applied for linear functions, when *a priori* information is not available or not taken into account.

$$p(\mathbf{x} \mid \mathbf{y}, \mathbf{C}_y)\mathrm{d}\mathbf{x} \propto \exp\left[-\frac{1}{2}(\mathbf{y} - f(\mathbf{x}))^{\mathrm{T}} \cdot \mathbf{C}_{\mathbf{y}}^{-1} \cdot (\mathbf{y} - f(\mathbf{x}))\right]\mathrm{d}\mathbf{x}$$

Least-square method (2)

Generalized least-squares method includes *a priori information* - essentially Bayesian parameter estimation under quadratic loss in saddle point approximation, for the important case that only *a priori* values and covariance matrices are given.

$$p(\mathbf{x}\,|\,\mathbf{x}_0, \mathbf{C}_{x_0}, \mathbf{y}, \mathbf{C}_y)\mathrm{d}\mathbf{x} \propto$$

$$\exp\left[-\frac{1}{2}(\mathbf{x}_0-\mathbf{x})^{\mathrm{T}}\cdot\mathbf{C}_{\mathbf{x}_0}^{-1}\cdot(\mathbf{x}_0-\mathbf{x})-\frac{1}{2}(\mathbf{y}-f(\mathbf{x}))^{\mathrm{T}}\cdot\mathbf{C}_{\mathbf{y}}^{-1}\cdot(\mathbf{y}-f(\mathbf{x}))\right]\mathrm{d}\mathbf{x}$$

$$\mathbf{x}=\mathbf{x}_0+\mathbf{C}_{\mathbf{x}_0}\cdot\mathbf{A}^{\mathrm{T}}\cdot(\mathbf{Q}+\mathbf{C}_{\mathbf{y}})^{-1}\cdot(\mathbf{y}-f(\mathbf{x}))$$

$$\mathrm{C}_{\mathrm{x}}=\mathbf{C}_{\mathbf{x}_0}-\mathbf{C}_{\mathbf{x}_0}\cdot\mathbf{A}^{\mathrm{T}}\cdot(\mathbf{Q}+\mathbf{C}_{\mathbf{y}})^{-1}\cdot\mathbf{A}\cdot\mathbf{C}_{\mathbf{x}_0}$$

$$\mathbf{Q}=\mathbf{A}\cdot\mathbf{C}_{x_0}\cdot\mathbf{A}^{\mathrm{T}}$$

$$(c^2)_{\min}=(\mathbf{y}-f(\mathbf{x}))^{\mathrm{T}}\cdot(\mathbf{Q}+\mathbf{C}_{\mathbf{y}})^{-1}\cdot(\mathbf{y}-f(\mathbf{x}))$$

Least-squares method (3)

Example: average of correlated data

Quantity Y is measured in two experiments using the same technique.

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \mathbf{C}_y = \begin{pmatrix} E_{y_1}^2 & K_{12} E_{y_1} E_{y_2} \\ K_{12} E_{y_1} E_{y_2} & E_{y_2}^2 \end{pmatrix} \qquad \begin{matrix} y_1 \approx x \\ y_2 \approx x \end{matrix} \qquad \mathbf{A} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\mathbf{C}_x = (\mathbf{A}^{\mathrm{T}} \cdot \mathbf{C}_y^{-1} \cdot \mathbf{A})^{-1} \qquad \mathbf{x} = \mathbf{C}_x \cdot \mathbf{A}^{\mathrm{T}} \cdot \mathbf{C}_y^{-1} \cdot \mathbf{y}$$

$$E_x^{-2} = C_x^{-1} = \frac{E_{y1}^{-2} + E_{y2}^{-2} - 2K_{12}E_{y1}^{-1}E_{y2}^{-1}}{1 - K_{12}^2} \qquad E_x^{-2} = E_{y1}^{-2} + E_{y2}^{-2}$$

$$x = \frac{y_1(E_{y1}^{-2} - K_{12}E_{y1}^{-1}E_{y2}^{-1}) + y_2(E_{y2}^{-2} - K_{12}E_{y1}^{-1}E_{y2}^{-1})}{(E_{y1}^{-2} - K_{12}E_{y1}^{-1}E_{y2}^{-1}) + (E_{y2}^{-2} - K_{12}E_{y1}^{-1}E_{y2}^{-1})}$$

$$x = \frac{y_1 E_{y1}^{-2} + y_2 E_{y2}^{-2}}{E_{y1}^{-2} + E_{y2}^{-2}}$$