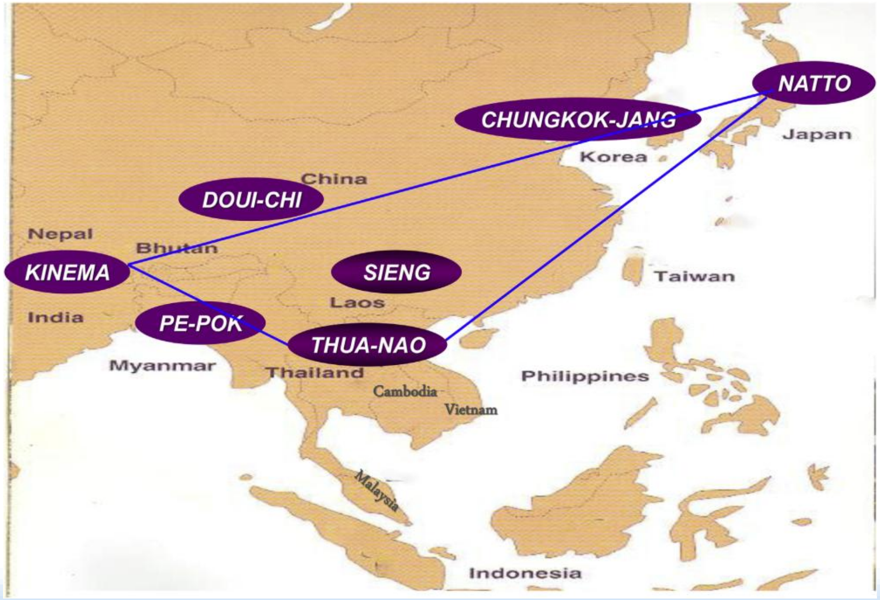
International Atomic Energy Agency

Data Reduction and Uncertainty Propagation

Naohiko Otsuka

IAEA Nuclear Data Section

Natto (Fermented Soya Beans, *Bekang* in Mizo)



Jyoti Prakash Tamang,
J. Ethnic Foods 2 (2015) 8

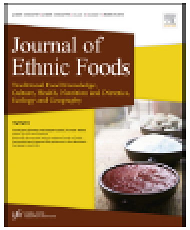
More Details about Bekang, Natto etc.

Journal of Ethnic Foods 2 (2015) 8–17

Contents lists available at ScienceDirect

Journal of Ethnic Foods

journal homepage: http://journalofethnicfoods.net



Review article

Naturally fermented ethnic soybean foods of India

Jyoti Prakash Tamang*

Department of Microbiology, School of Life Sciences, Sikkim University, Tadong, Sikkim, India

Open access article

Mathematical Details about Uncertainty Propagation

This presentation introduces several uncertainty propagation formulae without their proofs. See my recent article for their proofs:

http://www-nds.iaea.org/nrdc/india/ws2017/aizawl2017/otuka.pdf

Radiation Physics and Chemistry xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Radiation Physics and Chemistry

journal homepage: www.elsevier.com/locate/radphyschem

ELSEVIER

Uncertainty propagation in activation cross section measurements

N. Otuka[a,b,*], B. Lalremruata[c], M.U. Khandaker[d], A.R. Usman[d,e], L.R.M. Punte[c]

[a] Nuclear Data Section, Division of Physical and Chemical Sciences, Department of Nuclear Sciences and Applications, International Atomic Energy Agency, A-1400 Wien, Austria
[b] Nishina Center for Accelerator-Based Science, RIKEN, Wako, Saitama 351-0198, Japan
[c] Department of Physics, Mizoram University, Tanhril, Aizawl 796004, India
[d] Department of Physics, University of Malaya, 50603 Kuala Lumpur, Malaysia
[e] Department of Physics, Umaru Musa Yar'adua University, Katsina, Nigeria

Visit our Recent Article for Mathematical Details

$\int_{-\infty}^{+\infty} dx\, p(x_1, x_2, \ldots) = 1$, where $dx = dx_1 dx_2 \ldots$. The mean value (best estimate) x_{k0}, covariance $\mathrm{Cov}(x_k, x_l)$, correlation coefficient $\mathrm{Cor}(x_k, x_l)$, variance $\mathrm{Var}(x_k)$, and standard deviation Δx_k are defined by

$$x_{k0} = \int dx\, x_k p(x_1, x_2, ..), \tag{1}$$

$$\mathrm{Cov}(x_k, x_l) = \int dx (x_k - x_{k0})(x_l - x_{l0}) p(x_1, x_2, ..), \tag{2}$$

$$\mathrm{Cor}(x_k, x_l) = \mathrm{Cov}(x_k, x_l)/(\Delta x_k \Delta x_l), \tag{3}$$

$$\mathrm{Var}(x_k) = \int dx (x_k - x_{k0})^2 p(x_1, x_2, ..) = \mathrm{Cov}(x_k, x_k), \tag{4}$$

$$\Delta x_k = \sqrt{\mathrm{Var}(x_k)}, \tag{5}$$

respectively. By definition, $0 \le \mathrm{Cor}(x_k, x_l) \le 1$ and especially =1 when k=l. In nuclear data, one standard deviation of the parameter is usually treated as its *uncertainty*.[1] If x_1 is independent from the other parameters, we can decompose the probability distribution as

$$p(x_1, x_2, x_3, \ldots) = P(x_1) Q(x_2, x_3, \ldots), \tag{6}$$

and $\mathrm{Cov}(x_1, x_k) = 0$ ($k \neq 1$) according to the definition of the covariance.

If a set of quantities of interest $\{y_i\}$ are related to the parameters $\{x_k\}$ by $y_i = y_i(x_1, x_2, ..)$ and the relation can be linearized by expansion around the mean values of the parameters as

$$y_i = y_{i0} + \sum_k a_{ik}(x_k - x_{k0}) \tag{7}$$

with $y_{i0} = y_i(x_{10}, x_{20}, ..)$ and $a_{ik} = (\partial y_i/\partial x_k)_{x_k = x_{k0}}$ (sensitivity coefficient), the variance and covariance of x_k are propagated to those of y_i by

$$\mathrm{Var}(y_i) = \mathrm{Var}\left(\sum_k a_{ik} x_k\right) = \sum_k a_{ik}^2 \mathrm{Var}(x_k) + 2 \sum_{k>l} a_{ik} \mathrm{Cov}(x_k, x_l) a_{il}, \tag{8}$$

$$\mathrm{Cov}(y_i, y_j) = \mathrm{Cov}\left(\sum_k a_{ik} x_k, \sum_l a_{jl} x_l\right) = \sum_k \sum_l a_{ik} \mathrm{Cov}(x_k, x_l) a_{jl}. \tag{9}$$

Usually not all combinations of x_k and x_l have correlation but correlate each other within their n subsets such as $(x_1, x_2, \ldots, x_{M1})$, $(x_{M1+1}, \ldots, x_{M2})$, … . In such a case, the covariance terms in Eqs. (13) and (14) can be decomposed to

$$\sum_{k>l} g_{ik} \mathrm{cov}(x_k, x_l) g_{il} = \sum_{i=1}^{n} \sum_{k=M_{i-1}+1}^{M_i} \sum_{l=k+1}^{M_i} g_{ik} \mathrm{cov}(x_k, x_l) g_{il}, \tag{15}$$

$$\sum_{k,l} g_{ik} \mathrm{cov}(x_k, x_l) g_{jl} = \sum_{i=1}^{n} \sum_{k,l=M_{i-1}+1}^{M_i} g_{ik} \mathrm{cov}(x_k, x_l) g_{jl} \tag{16}$$

with $M_0 = 0$. For example, we expect that the number of counts C_i (always independent from other parameters), number of atoms in the samples per area n_i, and number of the incident particles Φ_i acting as six parameters $\{x_i\}$ (i=1,6) describing the cross sections σ_i=$C_i/(n_i\Phi_i)$ (i=1,2) has the following fractional covariances:

$$\begin{pmatrix} \mathrm{var}(C_1) & & & & & \\ 0 & \mathrm{var}(C_2) & & & & \\ 0 & 0 & \mathrm{var}(n_1) & & & \\ 0 & 0 & \mathrm{cov}(n_1, n_2) & \mathrm{var}(n_2) & & \\ 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{17}$$

if the uncertainty in Φ_i is negligible.

When y_i cannot be expressed by Eq. (10) and there is no correlation in parameters $\{x_k\}$, Eq. (8) can be rewritten as

$$(\Delta y_i/y_{i0})^2 = \sum_k s_{ik}^2 (\Delta x_k/x_{k0})^2, \tag{18}$$

where

$$s_{ik} = (x_{k0}/y_{i0})(\partial y_i/\partial x_k)_{x_k = x_{k0}} = (x_{k0}/y_{i0}) a_{ik} \tag{19}$$

is the relative sensitivity coefficient. Eq. (18) shows that we should distinguish the following two statements: "Uncertainty in y_i *due to* the uncertainty in x_k." (*i.e.*, $s_{ik}(\Delta x_k/x_{k0})$), and "Uncertainty *in* x_k." (*i.e.*,

Data Reduction

Nuclear reaction quantity q (e.g., cross section) is always derived from primary observables x, y, z, … (data reduction) by a function f:

Example: Activation cross section (σ_1, σ_2,..) may be derived from …

- measured counting rate A_1, A_2,…
- detector efficiency ε_1, ε_2,…
- number of sample atoms N_1, N_2,…
- beam flux density ϕ_1, ϕ_2,…

$\rightarrow \sigma_i = f(A_i, \varepsilon_i, N_i, \phi_i) = A_i/(\varepsilon_i \cdot N_i \cdot \phi_i)$

Basic Data Reduction

*** Addition $z = x + y$ or subtraction $z = x - y$**

Example:

Background correction to raw count: $N' = N - B$

*** Multiplication $z = x \cdot y$ or division $z = x/y$**

Example:

Efficiency correction to raw count: $N' = N / \varepsilon$

Real data reduction may be a combination of these operations,

e.g., $N' = (N-B)/\varepsilon$

or more complicated, e.g., $N[1-\exp(-\lambda t)]/\varepsilon$

Data Reduction and Uncertainty Propagation

Cross sections at n energies σ_i (i=1,n) derived from primary observables A_i, ε_i, N_i, ϕ_i (i=1,n)

Step 1: Measurements of primary observables

Determination of means and covariances for each primary observable $\langle A_i \rangle$, ΔA_i, $\langle \varepsilon_i \rangle$, $\Delta \varepsilon_i$ …

Step 2: Data reduction to cross sections

$\langle A_i \rangle$, ΔA_i, $\langle \varepsilon_i \rangle$, $\Delta \varepsilon_i$ … → $\langle \sigma_i \rangle$ and $\Delta \sigma_i$

Step 2-1: Propagation of mean value

$\langle \sigma_i \rangle = \langle A_i \rangle / (\langle \varepsilon_i \rangle \cdot \langle N_i \rangle \cdot \langle \phi_i \rangle)$

Step 2-2: Propagation of standard deviation

~~$\Delta \sigma_i = \Delta A_i / (\Delta \varepsilon_i \cdot \Delta N_i \cdot \Delta \phi_i)$~~

No!

Uncertainty Propagation (Linear Combination)

$p = \Sigma_{i=1,n}\, a_i \cdot x_i = a_1 x_1 + a_2 x_2 + \ldots$ (x_i is a random variable).

Mean:

$$\langle p \rangle = \Sigma_{i=1,n}\, a_i \langle x_i \rangle$$

Variance:

$$Var(p) = \langle p^2 \rangle - \langle p \rangle^2$$

$$= \langle (\Sigma_{i=1,n}\, a_i x_i)^2 \rangle - \langle \Sigma_{i=1,n}\, a_i x_i \rangle^2$$

$$= \Sigma_{i=1,n}\, a_i^2\, Var(x_i) + 2\, \Sigma_{i=1,n;\, j=1,n; i<j}\, a_i\, a_j\, Cov(x_i, x_j)$$

$$= \Sigma_{i=1,n}\, a_i^2\, (\Delta x_i)^2 + 2\, \Sigma_{i=1,n;\, j=1,n; i<j}\, a_i\, a_j\, Cor(x_i, x_j) \Delta x_i \Delta x_j$$

Uncertainty Propagation $(x,y) \rightarrow z$ for $z=x+y$

$z=x+y$ (e.g., background subtraction $N'=N-B$)

$$\langle z \rangle = \langle x \rangle + \langle y \rangle$$

$$\mathrm{Var}(z) = \mathrm{Var}(x) + \mathrm{Var}(y) + 2\mathrm{Cov}(x,y)$$

$$(\Delta z)^2 = (\Delta x)^2 + (\Delta y)^2 + 2\mathrm{Cor}(x,y)\Delta x \Delta y$$

If x and y are independent (i.e., $c_{xy}=0$),

$(\Delta z)^2=(\Delta x)^2+(\Delta y)^2$ (**quadrature sum rule**)

Uncertainty Propagation (x,y)→z for z=x+y

$z=x+y$

$\langle z \rangle = \langle x \rangle + \langle y \rangle$

$Var(z) = Var(x) + Var(y) + 2Cov(x,y)$

$(\Delta z)^2 = (\Delta x)^2 + (\Delta y)^2 + 2\,Cor(x,y)\Delta x \Delta y$

$=(\Delta x)^2+(\Delta y)^2$

(if x and y are independent, $Cor(x,y)=0$)

$=(\Delta x)^2+(\Delta y)^2+2\Delta x\Delta y = (\Delta x+\Delta y)^2$

(if x and y are fully correlated, $Cor(x,y)=+1$)

$=(\Delta x)^2+(\Delta y)^2-2\Delta x\Delta y = (\Delta x-\Delta y)^2$

(if x and y are fully anti-correlated, $Cor(x,y)=-1$)

Which case gives the largest and smallest uncertainties?

Uncertainty Propagation (General Function)

$p=p(x_1, x_2, \ldots x_n)$: function of n random variables x_n

1^{st} order expansion of p around $\langle p\rangle = p(\langle x_1, \langle x_2\rangle, \ldots \langle x_n\rangle)$:

$p - \langle p\rangle \sim \Sigma_{i=1,n} = (\partial p/\partial x_i)_{xi=\langle xi\rangle} (x_i - \langle x_i\rangle)$

If we set $p' = p - \langle p\rangle$ and $x_i' = x_i - \langle x_i\rangle$, we obtain the linear combination

$p' = \Sigma_{i=1,n} = (\partial p/\partial x_i)_{xi=\langle xi\rangle} \cdot x_i'$

This is a linear combination, therefore

$Var(p)=Var(p') \sim \Sigma_{i=1,n} (\partial p/\partial x_i)_{xi=\langle xi\rangle} Var(x_i)$

$+2\Sigma_{i=1,n;\ j=1,n;i<j} (\partial p/\partial x_i)_{xi=\langle xi\rangle} (\partial p/\partial x_j)_{xj=\langle xj\rangle} Cov(x_i, x_j)$

Limitation of Linear Approximation

$p=p(x_1, x_2, \ldots x_n)$: function of n random variables x_n

1st order expansion of f around $\langle p\rangle = p(\langle x_1\rangle, \langle x_2\rangle, \ldots \langle x_n\rangle)$:

$$p - \langle p\rangle \sim \Sigma_{i=1,n} = (\partial p/\partial x_i)_{xi=\langle xi\rangle} (x_i - \langle x_i\rangle)$$

This linear approximation is valid when $x_i - \langle x_i\rangle \ll \langle x_i\rangle$, namely valid only when the uncertainty is small enough than its mean value.

Uncertainty Propagation (General Function)

$p=p(x_{1,}x_2,\ldots x_n)$: function of n random variables

$$Var(p)=Var(p') \sim \Sigma_{i=1,n} (\partial p/\partial x_i)_{xi=<xi>} Var(x_i)$$
$$+2\Sigma_{i=1,n;\ j=1,n;i<j} (\partial p/\partial x_i)_{xi=<xi>} (\partial p/\partial x_j)_{xj=<xj>} Cov(x_i,x_j)$$

This relation can be easily extended to covariance between two functions $p=p(x_{1,}x_2,\ldots x_n)$, $q=q(y_{1,}y_2,\ldots y_m)$:

$$Cov(p,q) \sim \Sigma_{i=1,n;\ j=1,m} (\partial p/\partial x_i)_{xi=<xi>}(\partial q/\partial y_j)_{yj=<yj>} Cov(x_i,x_j)$$

Special Case: Product/Quotient Function

For two quantities

$$\begin{cases} p=(x_{1,p}x_{2,p},\ldots x_{m,p})/(x_{m+1,p}x_{m+2,p},\ldots x_{n,p}) \\ q=(x_{1,q}x_{2,q},\ldots x_{m,q})/(x_{m+1,q}x_{m+2,q},\ldots x_{n,q}) \end{cases}$$

fractional covariance $cov(p,q)=Cov(p,q)/(p \cdot q)$ is

$$cov(p,q) \sim \Sigma_{k=1,n}\, cov(x_{k,p},x_{k,p})$$

$$\sim \Sigma_{k=1,n}\, Cor(x_{k,p},x_{k,p})(\Delta x_{k,p}/\langle x_{k,p}\rangle)\,(\Delta x_{k,q}/\langle x_{k,q}\rangle)$$

If $p=q$,

$$cov(p,p)=var(p,p) \sim \Sigma_{k=1,n}\,(\Delta x_{k,p}/\langle x_{k,p}\rangle)^2$$

Namely $(\Delta p/\langle p\rangle)^2 = \Sigma_{k=1,n}(\Delta x_{k,p}/\langle x_{k,p}\rangle)^2$

(*Quadrature Sum Rule*)

Example: Activation Cross Section

Activation cross sections at two energies (σ_p and σ_q) derived from

- counts A (corrected for decay)
- by the same detector (ε) and sample (thickness N)
- under flux ϕ_p and ϕ_q:

  $\sigma_p = A_p / (\varepsilon N \phi_p)$ and $\sigma_q = (A_q / \varepsilon N \phi_q)$

Fractional variance (uncertainty):

$$var(\sigma_p) = (\Delta\sigma_p/\sigma_p)^2 = (\Delta A_p/A_p)^2 + (\Delta\varepsilon_p/\varepsilon_p)^2 + (\Delta N_p/N_p)^2 + (\Delta\phi_p/\phi_p)^2$$

Fractional covariance:

$$cov(\sigma_p, \sigma_q) = (\Delta\varepsilon_p/\varepsilon_p)^2 + (\Delta N_p/N_p)^2 \quad (p \neq q)$$

Summary

- **Uncertainty propagation depends on combination of random variables**
  - linear combination (e.g., $z=x+y$)
  - non-linear combination (e.g., $z=x/y$) - Taylor expansion around mean value
  - Product/quotient combination (e.g., activation formula)

- **Quadrature sum rule for the fractional uncertainty**

  The formula $(\Delta y/\langle y\rangle)^2=\Sigma_{i=1,n}(\Delta x_i/\langle x_i\rangle)^2$ is applicable when
  - y is a product/quotient combination $(x_1 x_2 \ldots . x_m)/(x_{m+1} x_{m+2} \ldots x_n)$
  - Δx_i is enough smaller than $\langle x_i\rangle$ (for 1st order approximation)

Visit our Recent Article for Mathematical Details

http://www-nds.iaea.org/nrdc/india/ws2017/aizawl2017/otuka.pdf

Radiation Physics and Chemistry xxx (xxxx) xxx–xxx

ELSEVIER

Contents lists available at ScienceDirect

Radiation Physics and Chemistry

journal homepage: www.elsevier.com/locate/radphyschem

Uncertainty propagation in activation cross section measurements

N. Otuka[a,b,*], B. Lalremruata[c], M.U. Khandaker[d], A.R. Usman[d,e], L.R.M. Punte[c]

[a] *Nuclear Data Section, Division of Physical and Chemical Sciences, Department of Nuclear Sciences and Applications, International Atomic Energy Agency, A-1400 Wien, Austria*
[b] *Nishina Center for Accelerator-Based Science, RIKEN, Wako, Saitama 351-0198, Japan*
[c] *Department of Physics, Mizoram University, Tanhril, Aizawl 796004, India*
[d] *Department of Physics, University of Malaya, 50603 Kuala Lumpur, Malaysia*
[e] *Department of Physics, Umaru Musa Yar'adua University, Katsina, Nigeria*

Also Mannhart's small guide and Smith's big guide!!
http://www-nds.iaea.org/nrdc/india/ws2017/aizawl2017/