

Taking Into Account Interval (and Fuzzy) Uncertainty Can Lead to More Adequate Statistical Estimates

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu

[Data Processing...](#)

[Usually Linearization...](#)

[Least Squares...](#)

[Case When We Do...](#)

[Case of Fuzzy Uncertainty](#)

[Case When We Need...](#)

[Combining...](#)

[From Formulas to...](#)

[But Where Do We Get...](#)

[Home Page](#)

[Title Page](#)

[«](#)

[»](#)

[◀](#)

[▶](#)

[Page 1 of 31](#)

[Go Back](#)

[Full Screen](#)

[Close](#)

[Quit](#)

1. Data Processing: General Introduction

- Some quantities, we can directly measure.
- For example, we can directly measure the distance between two points.
- However, many other quantities we cannot measure directly.
- For example, we cannot directly measure the spatial coordinates.
- To estimate such quantities X_i , we measure them *indirectly*:
 - we measure easier-to-measure quantities Y_1, \dots, Y_m
 - which are connected to X_j in a known way:
 $Y_i = f_i(X_1, \dots, X_n)$ for known functions f_i .

2. Sometimes, Measurement Results Also Depend on Additional Factors of No Interest to Us

- Sometimes, the measurement results also depend on auxiliary factors of no direct interest to us.
- For example, the time delays used to measure distances depend:
 - not only on the distance,
 - but also on the amount of H_2O in the troposphere.
- In such situations, we can add these auxiliary quantities to the list X_j of the unknowns.
- We may also use the result Y_i of additional measurements of these auxiliary quantities.

3. Data Processing (cont-d)

- *Example:*

- we want to measure coordinates X_j of an object;
- we measure the distance Y_i between this object and objects with accurately known coordinates $X_j^{(i)}$:

$$Y_i = \sqrt{\sum_{j=1}^3 (X_j - X_j^{(i)})^2}.$$

- *General case:*

- we know the results \tilde{Y}_i of measuring Y_i ;
- we want to estimate the desired quantities X_j .

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[Page 4 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

4. Usually Linearization Is Possible

- In most practical situations, we know the approximate values $X_j^{(0)}$ of the desired quantities X_j .
- These approximation are usually reasonably good, in the sense that the difference $x_j \stackrel{\text{def}}{=} X_j - X_j^{(0)}$ are small.
- In terms of x_j , we have

$$Y_i = f(X_1^{(0)} + x_1, \dots, X_n^{(0)} + x_n).$$

- We can safely ignore terms quadratic in x_j .
- Indeed, even if the estimation accuracy is 10% (0.1), its square is 1% \ll 10%.
- We can thus expand the dependence of Y_i on x_j in Taylor series and keep only linear terms:

$$Y_i = Y_i^{(0)} + \sum_{j=1}^n a_{ij} \cdot x_j, \quad Y_i^{(0)} \stackrel{\text{def}}{=} f_i(X_1^{(0)}, \dots, X_n^{(0)}), \quad a_{ij} \stackrel{\text{def}}{=} \frac{\partial f_i}{\partial X_j}.$$

[Data Processing: ...](#)[Usually Linearization ...](#)[Least Squares ...](#)[Case When We Do ...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need ...](#)[Combining ...](#)[From Formulas to ...](#)[But Where Do We Get ...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 5 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

5. Least Squares

- Thus, to find the unknowns x_j , we need to solve a system of approximate linear equations

$$\sum_{j=1}^n a_{ij} \cdot x_j \approx y_i, \text{ where } y_i \stackrel{\text{def}}{=} \tilde{Y}_i - Y_i^{(0)}.$$

- Usually, it is assumed that each measurement error is:
 - normally distributed
 - with 0 mean (and known st. dev. σ_i).
- The distribution is indeed often normal:
 - the measurement error is a joint result of many independent factors,
 - and the distribution of the sum of many small independent errors is close to Gaussian;
 - this is known as the Central Limit Theorem.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 6 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

6. Least Squares (cont-d)

- 0 mean also makes sense:
 - we calibrate the measuring instrument by comparing it with a more accurate,
 - so if there was a bias (non-zero mean), we delete it by re-calibrating the scale.
- It is also assumed that measurement errors of different measurements are independent.
- In this case, for each possible combination $x = (x_1, \dots, x_n)$, the probability of observing y_1, \dots, y_m is:

$$\prod_{i=1}^m \left(\frac{1}{\sqrt{2\pi} \cdot \sigma_i} \cdot \exp \left(-\frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j \right)^2}{2\sigma_i^2} \right) \right).$$

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 7 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

7. Least Squares (final)

- It is reasonable to select x_j for which this probability is the largest, i.e., equivalently, for which

$$\sum_{i=1}^n \frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j\right)^2}{\sigma_i^2} \rightarrow \min .$$

- The set S of all possible combinations x is:

$$S = \left\{ x : \sum_{i=1}^n \frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j\right)^2}{\sigma_i^2} \leq \chi_{m-n,\alpha}^2 \right\} .$$

- If $S = \emptyset$, this means that some measurements are outliers.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[Page 8 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

8. Simple Example

- Suppose that we have m measurements y_1, \dots, y_m of the same quantity x_1 , with 0 mean and st. dev. σ_i .
- Then, the least squares estimate for x_1 is

$$\hat{x}_1 = \frac{\sum_{i=1}^m \sigma_i^{-2} \cdot y_i}{\sum_{i=1}^m \sigma_i^{-2}}.$$

- The accuracy of this estimate is $\sigma^2[x_1] = \frac{1}{\sum_{i=1}^m \sigma_i^{-2}}$.
- In particular, for $\sigma_1 = \dots = \sigma_m = \sigma$, we get

$$\hat{x}_1 = \frac{y_1 + \dots + y_m}{m}, \text{ with } \sigma[x_1] = \frac{\sigma}{\sqrt{m}}.$$

[Data Processing: ...](#)[Usually Linearization ...](#)[Least Squares ...](#)[Case When We Do ...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need ...](#)[Combining ...](#)[From Formulas to ...](#)[But Where Do We Get ...](#)[Home Page](#)[Title Page](#)[Page 9 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

9. Least Squares Approach Is Not Always Applicable

- There are cases when this Least Squares approach is not applicable.
- The first case is when we use the most accurate measuring instruments.
- In this case, we have no more accurate instrument to calibrate.
- So, we do not know the mean, we do not know the distribution.
- The second case is when we have many measurements.
- If we simply measure the same quantity m times, we get an estimate (average) with accuracy $\frac{\sigma}{\sqrt{m}}$.
- So, if we use GPS with 1 m accuracy million times, we can 1 mm accuracy, then microns etc.

Data Processing: ...

Usually Linearization ...

Least Squares ...

Case When We Do ...

Case of Fuzzy Uncertainty

Case When We Need ...

Combining ...

From Formulas to ...

But Where Do We Get ...

Home Page

Title Page

◀

▶

◀

▶

Page 10 of 31

Go Back

Full Screen

Close

Quit

10. Least Squares Approach Is Not Always Applicable (cont-d)

- This makes no physical sense.
- When we calibrate, we guarantee that the systematic error (mean) is much smaller than the random error.
- However:
 - when we repeat measurements – and take the average – we decrease random error,
 - however, the systematic error does not decrease,
 - so, systematic error becomes larger than the remaining random error.
- Let us consider these two cases one by one.

Data Processing: ...

Usually Linearization ...

Least Squares ...

Case When We Do ...

Case of Fuzzy Uncertainty

Case When We Need ...

Combining ...

From Formulas to ...

But Where Do We Get ...

Home Page

Title Page

◀

▶

◀

▶

Page 11 of 31

Go Back

Full Screen

Close

Quit

11. Case When We Do Not Know the Distributions: Enter Interval and Fuzzy Uncertainties

- Let us first consider the case when we do not know the distribution of the measurement error.
- In some such cases, we know the upper bound Δ_i on the i -th measurement error.
- Thus, based on the measured values y_i , we can conclude that the actual value of $s_i \stackrel{\text{def}}{=} \sum_{j=1}^n a_{ij} \cdot x_j$ is in the interval

$$y_i \stackrel{\text{def}}{=} [y_i - \Delta_i, y_i + \Delta_i].$$

- In other cases, we do not have a guaranteed bound Δ_i .

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 12 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

12. Case of Fuzzy Uncertainty

- Instead, for each level of certainty p , we have a corresponding bound $\Delta_i(p)$.
- Thus, with certainty p , we can conclude that

$$s_i \in \mathbf{y}_i(p) \stackrel{\text{def}}{=} [y_i - \Delta_i(p), y_i + \Delta_i(p)].$$

- To get higher p , we need to enlarge the interval.
- Thus, we have a nested family of intervals.
- Describing such a family is equivalent to describing a fuzzy set with α -cuts $\mathbf{y}_i(1 - \alpha)$.

[Data Processing: ...](#)[Usually Linearization ...](#)[Least Squares ...](#)[Case When We Do ...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need ...](#)[Combining ...](#)[From Formulas to ...](#)[But Where Do We Get ...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 13 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

13. Case of Interval Uncertainty (cont-d)

- For different $y_i \in \mathbf{y}_i$, we get different values x_j .
- The largest possible value \bar{x}_j can be obtained by solving the following linear programming problem:

$$x_j \rightarrow \max \text{ under constraints } y_i - \Delta_i \leq \sum_{k=1}^n a_{ik} \cdot x_k \leq y_i + \Delta_i.$$

- The smallest possible value \underline{x}_j can be obtained by minimizing x_j under the same constraints.
- There exist efficient algorithms for solving linear programming problems.
- In general, the set S of possible values x is a polyhedron determined by the above inequalities.
- In the fuzzy case, we repeat the same computation for each p , and get bounds $\underline{x}_j(p)$ and $\bar{x}_j(p)$ for each p .

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[◀◀](#)[▶▶](#)[◀](#)[▶](#)[Page 14 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

14. Simple Example

- Suppose that we have m measurements y_1, \dots, y_m of the same quantity x_1 , with bounds Δ_i .
- Then, based on each measurement i , we can conclude that

$$x_1 \in [y_i - \Delta_i, y_i + \Delta_i].$$

- Thus, based on all m measurements, we can conclude that x_1 belongs to the intersection of these m intervals:

$$\bigcap_{i=1}^m [y_i - \Delta_i, y_i + \Delta_i] = \left[\max_{1 \leq i \leq n} (y_i - \Delta_i), \min_{1 \leq i \leq n} (y_i + \Delta_i) \right].$$

- The more measurements, the narrower the resulting interval.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 15 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

15. In General, How Do We Describe the Set S of Possible Values of x ?

- In the first approximation, we find the intervals $[\underline{x}_j, \overline{x}_j]$.
- Then, we can conclude that $x = (x_1, \dots, x_n)$ belongs to the box

$$[\underline{x}_1, \overline{x}_1] \times \dots \times [\underline{x}_n, \overline{x}_n].$$

- Often, not all combinations from the box are possible.
- To get a better description of the set S , we can also find max and min of the values

$$\sum_{i=1}^n \alpha_i \cdot x_i, \text{ with } \alpha_i \in \{-1, 1\}.$$

- For example, for $n = 2$ (e.g., for localizing a point in the plane), we also find the bounds on

$$s_1 \stackrel{\text{def}}{=} x_1 + x_2 \text{ and } s_2 \stackrel{\text{def}}{=} x_1 - x_2.$$

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[Page 16 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

16. Describing the Set S (cont-d)

- Using all these bounds leads to a better description of the set S .
- For example, for $n = 2$, we have bounds

$$\underline{x}_1 \leq x_1 \leq \bar{x}_1, \quad \underline{x}_2 \leq x_2 \leq \bar{x}_2,$$

$$\underline{s}_1 \leq x_1 + x_2 \leq \bar{s}_1, \quad \underline{s}_2 \leq x_1 - x_2 \leq \bar{s}_2.$$

- If this description is not enough, we take values

$$\sum_{i=1}^n \alpha_i \cdot x_i, \text{ with } \alpha_i \in \{-1, 0, 1\} \text{ or, more generally, with:}$$

$$\alpha_i \in \left\{ -1, -1 + \frac{2}{M}, -1 + \frac{4}{M}, \dots, 1 - \frac{2}{M}, 1 \right\} \text{ for } M = 1, 2, \dots$$

17. Additional Constraints

- In some practical situations, we also have additional constraints.
- For example, we can have bounds on the amount of water in the troposphere.
- From the computational viewpoint, dealing with these additional constraints is easy:
 - we simply add these additional constraints

$$\underline{x}_k \leq x_k \leq \bar{x}_k$$

- to the list of constraints under which we optimize x_j .

[Data Processing: ...](#)[Usually Linearization ...](#)[Least Squares ...](#)[Case When We Do ...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need ...](#)[Combining ...](#)[From Formulas to ...](#)[But Where Do We Get ...](#)[Home Page](#)[Title Page](#)[Page 18 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

18. Case When We Need to Take into Account Systematic Error

- In the traditional approach, we assume that $y_i = \sum_{j=1}^n a_{ij} \cdot x_j + e_i$, where the meas. error e_i has 0 mean.
- Sometimes:
 - in addition to the random error $e_i^r \stackrel{\text{def}}{=} e_i - E[e_i]$ with 0 mean,
 - we also have a systematic error $e_i^s \stackrel{\text{def}}{=} E[e_i]$:

$$y_i = \sum_{j=1}^n a_{ij} \cdot x_j + e_i^r + e_i^s.$$

- Sometimes, we know the upper bound Δ_i : $|e_i^s| \leq \Delta_i$.
- In other cases, we have different bounds $\Delta_i(p)$ corresponding to different degree of confidence p .
- What can we then say about x_j ?

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[◀◀](#)[▶▶](#)[◀](#)[▶](#)[Page 19 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

19. Combining Probabilistic and Interval (or Fuzzy) Uncertainty: Main Idea

- If we knew the values e_i^s , then we would conclude that for $e_i^r = y_i - \sum_{j=1}^n a_{ij} \cdot x_j - e_i^s$, we have

$$\sum_{i=1}^m \frac{(e_i^r)^2}{\sigma_i^2} = \sum_{i=1}^m \frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j - e_i^s \right)^2}{\sigma_i^2} \leq \chi_{m-n, \alpha}^2.$$

- In practice, we do not know the values e_i^s , we only know that these values are in the interval $[-\Delta_i, \Delta_i]$.
- Thus, we know that the above inequality holds for some

$$e_i^s \in [-\Delta_i, \Delta_i].$$

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 20 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

20. Main Idea (cont-d)

- The above condition is equivalent to $v(x) \leq \chi_{m-n,\alpha}^2$, where

$$v(x) \stackrel{\text{def}}{=} \min_{e_i^s \in [-\Delta_i, \Delta_i]} \sum_{i=1}^m \frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j - e_i^s \right)^2}{\sigma_i^2}.$$

- So, the set S_α of all combinations $X = (x_1, \dots, x_n)$ which are possible with confidence α is:

$$S_\alpha = \{x : v(x) \leq \chi_{m-n,\alpha}^2\}.$$

- The range of possible values of x_j can be obtained by maximizing and minimizing x_j under the constraint

$$v(x) \leq \chi_{m-n,\alpha}^2.$$

- In the fuzzy case, we have to repeat the computations for every p .

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[Page 21 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

21. How to Check Consistency

- We want to make sure that the measurements are consistent – i.e., that there are no outliers.
- This means that we want to check that there exists some $x = (x_1, \dots, x_n)$ for which $v(x) \leq \chi_{m-n, \alpha}^2$.
- This condition is equivalent to

$$v \stackrel{\text{def}}{=} \min_x v(x) = \min_x \min_{e_i^s \in [-\Delta_i, \Delta_i]} \sum_{i=1}^m \frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j - e_i^s \right)^2}{\sigma_i^2} \leq \chi_{m-n, \alpha}^2.$$

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[◀◀](#)[▶▶](#)[◀](#)[▶](#)[Page 22 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

22. This Is Indeed a Generalization of Probabilistic and Interval Approaches

- In the case when $\Delta_i = 0$ for all i , i.e., when there is no interval uncertainty, we get the usual Least Squares.
- Vice versa, for very small σ_i , we get the case of pure interval uncertainty.
- In this case, the above formulas tend to the set of all the values for which
$$\left| y_i - \sum_{j=1}^n a_{ij} \cdot x_j \right| \leq \Delta_i.$$
- E.g., for m repeated measurements of the same quantity, we get the intersection of the corr. intervals.
- So, the new idea is indeed a generalization of the known probabilistic and interval approaches.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 23 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

23. From Formulas to Computations

- The expression $\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j - e_i^s\right)^2$ is a convex function of x_j .
- The domain of possible values of $e^s = (e_1^s, \dots, e_m^s)$ is also convex: it is a box

$$[-\Delta_1, \Delta_1] \times \dots \times [-\Delta_m, \Delta_m].$$

- There exist efficient algorithms for computing minima of convex functions over convex domains.
- These algorithms also compute locations where these minima are attained.
- Thus, for every x , we can efficiently compute $v(x)$ and thus, efficiently check whether $v(x) \leq \chi_{m-n,\alpha}^2$.
- Similarly, we can efficiently compute v and thus, check whether $v \leq \chi_{m-n,\alpha}^2$ - i.e., whether we have outliers.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[Page 24 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

24. From Formulas to Computations (cont-d)

- The set S_α is convex.
- We can approximate the set S_α by:
 - taking a grid G ,
 - checking, for each $x \in G$, whether $v(x) \leq \chi_{m-n,\alpha}^2$,
and
 - taking the convex hull of “possible” points.
- We can also efficiently find the minimum \underline{x}_j of x_j over
$$x \in S_\alpha.$$
- By computing min of $-x_j$, we can also find the maximum \bar{x}_j .

[Data Processing: ...](#)[Usually Linearization ...](#)[Least Squares ...](#)[Case When We Do ...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need ...](#)[Combining ...](#)[From Formulas to ...](#)[But Where Do We Get ...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 25 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

25. But Where Do We Get the Bounds on Systematic Errors?

- The above algorithms require that we have some bounds on the systematic error component.
- But where can we get these bounds?
- Let's recall that we get σ_i from calibration.
- In the process of calibration:
 - we also get an estimate for the bias, and
 - we use this estimate to re-calibrate our instrument
 - so that its bias will be 0.
- If we could estimate the bias more accurately, we would have eliminated it too.
- So, where do the bounds Δ_i come from?

Data Processing: ...

Usually Linearization ...

Least Squares ...

Case When We Do ...

Case of Fuzzy Uncertainty

Case When We Need ...

Combining ...

From Formulas to ...

But Where Do We Get ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 26 of 31

Go Back

Full Screen

Close

Quit

26. Where Do We Get the Bounds (cont-d)

- The answer is simple:
 - after calibration, we get an estimate for the bias,
 - but this numerical estimate is only approximate.
- From the same calibration experiment, we can extract:
 - not only this estimate b ,
 - but also the confidence interval $[\underline{b}, \bar{b}]$ which contains b with given confidence.
- After we use b to re-scale, the remaining bias is – with given confidence – in the interval

$$[\underline{b} - b, \bar{b} - b].$$

- This is where the corresponding bound Δ_i comes from: it is simply $\Delta_i = \max(\bar{b} - b, b - \underline{b})$.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[◀◀](#)[▶▶](#)[◀](#)[▶](#)[Page 27 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

27. Relation to Uniform Distributions: Caution Is Needed

- Usually, in probability theory:
 - if we do not know the exact distribution,
 - then out of possible distributions, we select the one with the largest entropy

$$- \int \rho(x) \cdot \ln(\rho(x)) dx.$$

- In particular:
 - if we only know that the random variable is located somewhere on the interval $[-\Delta_i, \Delta_i]$,
 - Maximum Entropy approach leads to a uniform distribution on this interval.

[Data Processing: ...](#)[Usually Linearization ...](#)[Least Squares ...](#)[Case When We Do ...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need ...](#)[Combining ...](#)[From Formulas to ...](#)[But Where Do We Get ...](#)[Home Page](#)[Title Page](#)[<<](#)[>>](#)[<](#)[>](#)[Page 28 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)

28. Relation to Uniform Distributions (cont-d)

- If η is distributed with pdf $\rho(x)$, then the sum of η and an m -D uniform distribution has the density

$$\rho'(x) = \max_{e_i^s \in [-\Delta_i, \Delta_i]} \rho(x - e^s).$$

- The maximum likelihood method $\rho'(x) \rightarrow \max$ is equivalent to $-\ln(\rho'(x)) \rightarrow \min$, where:

$$-\ln(\rho'(x)) = \min_{e_i^s \in [-\Delta_i, \Delta_i]} (-\ln(\rho(x - e^s))).$$

- For the normal distribution,

$$-\ln(\rho(x)) = \text{const} + \frac{1}{2} \cdot \sum_{i=1}^m \frac{(e_i^r)^2}{\sigma_i^2}.$$

29. Relation to Uniform Distributions (cont-d)

- Thus, maximum likelihood $\rho'(x) \rightarrow \max$ leads to

$$\min_{e_i^s \in [-\Delta_i, \Delta_i]} \sum_{i=1}^m \frac{\left(y_i - \sum_{j=1}^n a_{ij} \cdot x_j - e_i^s \right)^2}{\sigma_i^2} \rightarrow \min.$$

- The minimized expression is exactly our $v(x)$.
- Does this means that we can safely assume that the systematic error is uniformly distributed on $[-\Delta_i, \Delta_i]$.
- This is, e.g., what ISO suggests.
- Our answer is: not always.

30. Caution Is Needed

- Indeed, for the sum $s = x_1 + \dots + x_m$ of m such errors with $\Delta_i = \Delta$ all we can say is that $s \in [-m \cdot \Delta, m \cdot \Delta]$.
- However, for large m ,
 - due to the Central Limit Theorem,
 - the sum s is practically normally distributed, with 0 mean and st. dev. $\sim \sqrt{m} \cdot \sigma$.

- So, with very high confidence, we can conclude that

$$|s| \leq \text{const} \cdot (\sqrt{m} \cdot \sigma).$$

- For large m , this bound is much smaller than $m \cdot \sigma$ and is, thus, a severe underestimation of the possible error.
- *Conclusion:* in some calculations, we can use MaxEnt and uniform distributions, but not always.
- In other words, we must be cautious.

[Data Processing...](#)[Usually Linearization...](#)[Least Squares...](#)[Case When We Do...](#)[Case of Fuzzy Uncertainty](#)[Case When We Need...](#)[Combining...](#)[From Formulas to...](#)[But Where Do We Get...](#)[Home Page](#)[Title Page](#)[Page 31 of 31](#)[Go Back](#)[Full Screen](#)[Close](#)[Quit](#)