

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

Ligang Sun, Hani Dbouk, Ingo Neumann, Steffen Schön, and Vladik Kreinovich

**Abstract** Traditional statistical data processing techniques (such as Least Squares) assume that we know the probability distributions of measurement errors. Often, we do not have full information about these distributions. In some cases, all we know is the bound of the measurement error; in such cases, we can use known interval data processing techniques. Sometimes, this bound is fuzzy; in such cases, we can use known fuzzy data processing techniques.

However, in many practical situations, we know the probability distribution of the random component of the measurement error and we know the upper bound – numerical or fuzzy – on the measurement error’s systematic component. For such situations, no general data processing technique is currently known. In this paper, we describe general data processing techniques for such situations, and we show that taking into account interval and fuzzy uncertainty can lead to more adequate statistical estimates.

## 1 Formulation of the Problem: Traditional Statistical Approach to Data Processing Is Not Always Applicable

**Data processing: a brief reminder.** Some quantities, we can directly measure. For example, on the Earth, we can usually directly measure the distance between the two nearby points. However, many other quantities  $X_j$  we cannot measure directly.

---

Ligang Sun, Hani Dbouk, Ingo Neumann, Steffen Schön  
Leibniz Universität Hannover, 30167 Hannover, Germany,  
e-mail: ligang.sun@gih.uni-hannover.de, dbouk@mbox.ife.uni-hannover.de,  
neumann@gih.uni-hannover.de, schoen@ife.uni-hannover.de

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

For example, we cannot directly measure the spatial coordinates. To estimate such quantities  $X_j$ , we measure them *indirectly*, i.e.:

- 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  $f_i(X_1, \dots, X_n)$ ,

and then we reconstruct the values  $X_j$  of the desired quantities from the measurement results:

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

This reconstruction is what is often understood by data processing.

**Example.** Suppose that we want to measure coordinates  $X_j$  of an object. For this purpose, we measure the distance  $Y_i$  between this object and objects with known

coordinates  $X_j^{(i)}$ :  $Y_i = \sqrt{\sum_{j=1}^3 (X_j - X_j^{(i)})^2}$ , and then reconstruct the coordinates based on the measured values of these distances. This is how, e.g., GPS works – after estimating the clock offsets between the receiver’s clock and the satellites’ clocks, we use the correspondingly corrected travel times to estimate the distances  $Y_i$  from our location  $X_j$  to satellites whose positions  $X_j^{(i)}$  are known with high accuracy.

**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 and on the sensors’ time offset; see, e.g., [16].

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.

**Usually, linearization is possible.** In most practical situations, we know the approximate values  $X_j^{(0)}$  of the desired quantities  $X_j$ .

For example, in geodesy, we want to find the coordinates  $X_j$  of different locations. We do not know the exact values of these coordinates, but we usually know the approximate location  $X_j^{(0)}$  that was obtained by previous measurements. Our goal is then to use the measurement results  $\tilde{Y}_i$  to come up with more accurate estimates for  $X_j$ .

These approximations 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  $X_i = X_i^{(0)} + x_i$  and thus,

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

For a good approximation, we can safely ignore terms quadratic in  $x_j$ . Indeed, even if the estimation accuracy is 10% (0.1), its square is 1%, which is much smaller than 10%.

We can thus expand the dependence of  $Y_i$  on  $x_j$  in Taylor series and keep only linear terms:

$$Y_i = Y_i^{(0)} + \sum_{j=1}^n a_{ij} \cdot x_j,$$

where

$$Y_i^{(0)} \stackrel{\text{def}}{=} f_i(X_1^{(0)}, \dots, X_n^{(0)})$$

and

$$a_{ij} \stackrel{\text{def}}{=} \frac{\partial f_i}{\partial X_j} \Big|_{X_1=X_1^{(0)}, \dots, X_n=X_n^{(0)}}.$$

By moving the value  $Y_i^{(0)}$  to the other side of this formula, we conclude that

$$Y_i - Y_i^{(0)} = \sum_{j=1}^n a_{ij} \cdot x_j.$$

Here, we know the values  $a_{ij}$  – they are obtained by differentiating the known functions  $f_i(X_1, \dots, X_n)$ . We also know the value  $Y_i^{(0)}$  – we compute each of these values by applying the known function  $f_i(X_1, \dots, X_n)$  to the known approximate values  $X_j^{(0)}$  values that we knew before the measurements. We do not, however, know the exact value  $Y_j$  of the corresponding quantity. Instead, as a result of measuring this quantity, we get the measurement result  $\tilde{Y}_j \approx Y_j$ . Since  $\tilde{Y}_j \approx Y_j$ , the known difference  $y_i \stackrel{\text{def}}{=} \tilde{Y}_i - Y_i^{(0)}$  is approximately equal to  $Y_i - Y_i^{(0)}$ , and thus, approximately equal to the sum  $\sum_{j=1}^n a_{ij} \cdot x_j$ .

Thus, to find the unknowns  $x_j$ , we need to solve a system of approximate linear equations  $\sum_{j=1}^n a_{ij} \cdot x_j \approx y_i$ , with known values  $y_i$  and  $a_{ij}$ .

**The Least Squares approach.** Usually, it is assumed that each measurement error is normally distributed with 0 mean (and known standard deviation  $\sigma_i$ ).

The distribution is indeed often normal; see, e.g., [10, 11]. Indeed, the measurement error is usually a joint result of many independent factors, and the distribution of the sum of many small independent errors is close to Gaussian (this result is known as the *Central Limit Theorem*; see, e.g., [13]).

The assumption that the mean value of the measurement error is 0 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, under the Gaussian assumption, for each possible combination  $x = (x_1, \dots, x_n)$ , the probability of observing  $y_1, \dots, y_m$  is equal to:

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

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

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

(This natural idea is known as the *Maximum Likelihood approach*.) The set  $S_\gamma$  of all possible combinations  $x$  – known as the *confidence set* – has the following form, where  $\chi_{m-n,\gamma}^2$  is the value of the chi-square statistic corresponding to the confidence  $1 - \gamma$  (i.e., to the probability  $\gamma$  of the false alarm; see, e.g., [13]):

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

*Comment.* If this set  $S_\gamma$  is empty, this means that some measurements are outliers.

**A simple example.** Suppose that we have  $m$  measurements  $y_1, \dots, y_m$  of the same quantity  $x_1$ , with 0 mean and standard deviation  $\sigma_i$ . 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 (standard deviation) 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

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

**The Least Squares approach is not always applicable.** While in many practical situations, the Least Squares approach has been very successful, there are cases when the Least Squares approach is not applicable.

The first case is when we use the most accurate measuring instruments. In this case, we don't have any more accurate instrument that we could use for calibration. So, we do not know the mean, and we do not know the distribution. What we may know in such situations is the upper bound on the measurement error; this bound may be a number or it may even be an expert estimate described by using natural language words like "small".

The second case is when:

- we have a good approximation to the probability distribution of the measurement error,
- we have calibrated the measuring instrument so that the remaining bias is statistically indistinguishable from 0 – and
- with thus calibrated measuring instrument, we perform a large number of measurements.

At first glance, this may seem a perfect case for applying the Least Squares techniques. However, if we simply measure the same quantity  $m$  times, we get an estimate (average) with accuracy  $\frac{\sigma}{\sqrt{m}}$ . So, if we, e.g., use GPS with 1 m accuracy million times, we can get 1 mm accuracy, then microns etc. This makes no physical sense. The explanation for this is simple. When we calibrate, we guarantee that the systematic error (i.e., the mean value of the measurement error) is much smaller than the random error. However, when we repeat measurements and take the average of the measurement results, we decrease the random error, while the systematic error does not decrease. So, the systematic error becomes larger than the remaining random error.

**What we do in this paper.** In this paper, we consider these two cases one by one, and we show that in both cases, interval and fuzzy approaches can help make statistical estimates more adequate.

## 2 Case 1, When We Do Not Know the Distributions: Enter Interval and Fuzzy Uncertainties

**What do we know: a question.** Let us first consider the case when we do not know the distribution of the measurement error. As we have mentioned, in this case, we know either the numerical guaranteed upper bound on the measurement error, or at least bounds which are valid with some confidence. Let us consider these two types of situations one by one.

**Situations when we know guaranteed upper bounds on the measurement errors: enter interval uncertainty.** In some situations, we know the upper bound  $\Delta_i$  on the  $i$ -th measurement error. Thus, based on the measured values  $y_i$ , we can con-

clude that the actual value of the quantity  $s_i \stackrel{\text{def}}{=} \sum_{j=1}^n a_{ij} \cdot x_j$  (which is approximately equal to  $y_i$ ) is in the interval  $\mathbf{y}_i \stackrel{\text{def}}{=} [y_i - \Delta_i, y_i + \Delta_i]$ ; see, e.g., [3, 7].

**Situations when we only have imprecise expert estimates of the upper bounds on the measurement errors: center fuzzy uncertainty.** Let us now consider the situations when we do not have guaranteed bounds  $\Delta_i$ , we only have expert estimates of these bounds. These estimates come with different levels of certainty.

So, 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 nested family of intervals is equivalent to describing a fuzzy set with  $\alpha$ -cuts  $\mathbf{y}_i(1 - \alpha)$ ; see, e.g., [5, 9, 17].

**How to process interval uncertainty.** 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 (see, e.g., [6]), we can use them. In general, the set  $S$  of possible values  $x$  is a polyhedron determined by the above inequalities.

**A simple example.** Suppose that we have  $m$  measurements  $y_1, \dots, y_m$  of the same quantity  $x_1$ , with bounds  $\Delta_i$ . 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.

*Comment.* If the intersection is empty – or, more generally, if there are no values  $x_j$  for which  $\sum_{j=1}^n a_{ij} \cdot x_j \in \mathbf{y}_i$  for all  $i$  – this means that some of the measurement results are actually outliers; see, e.g., [14].

**How to process fuzzy uncertainty.** In the fuzzy case, we need to repeat the same interval-related computation for each  $p$ , and get bounds  $\underline{x}_j(p)$  and  $\bar{x}_j(p)$  for each  $p$ . The resulting nested intervals form a fuzzy set of possible values of  $x_j$ .

**In general, how do we describe the set  $S$  of possible values of  $x$ ?** In the first approximation, we find the intervals  $[\underline{x}_j, \bar{x}_j]$ . Then, we can conclude that  $x = (x_1, \dots, x_n)$  belongs to the box  $[\underline{x}_1, \bar{x}_1] \times \dots \times [\underline{x}_n, \bar{x}_n]$ .

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

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

For example, for  $n = 2$  (e.g., for localizing a point in the plane), we also find the bounds on  $s_1 \stackrel{\text{def}}{=} x_1 + x_2$  and  $s_2 \stackrel{\text{def}}{=} x_1 - x_2$ . Using all these bounds leads to a better description of the set  $S$ .

For example, for  $n = 2$ , we have bounds

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

If this description is not enough, we take values  $\sum_{i=1}^n \beta_i \cdot x_i$ , with  $\beta_i \in \{-1, 0, 1\}$  or, more generally, with:

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

**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  $x_j$ .

*Comment.* Alternatively, we can use zonotopes to describe the set of all possible vectors  $x = (x_1, \dots, x_n)$ ; see, e.g., [15].

### 3 Case 2, When We Know (a Good Approximation to) the Probability Distribution of the Measurement Error and We Know an Upper Bound on the Systematic Error

**Reminder.** In the traditional approach, we assume that  $y_i = \sum_{j=1}^n a_{ij} \cdot x_j + e_i$ , where the measurement 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  $e_i^s \stackrel{\text{def}}{=} E[e_i]$ :

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

**What do we know about the systematic error: interval and fuzzy cases.** Sometimes, we know the upper bound  $\Delta_i$  on the systematic error:  $|e_i^s| \leq \Delta_i$ . In other cases,

we have different bounds  $\Delta_i(p)$  corresponding to different degrees of confidence  $p$ . Based on all this information, what can we then say about  $x_j$ ?

**Our 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

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

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 values  $e_1^s, \dots, e_m^s$  for which  $e_i^s \in [-\Delta_i, \Delta_i]$ .

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

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

So, the set  $S_\gamma$  of all combinations  $X = (x_1, \dots, x_n)$  which are possible with confidence  $1 - \gamma$  has the following form:  $S_\gamma = \{x : v(x) \leq \chi_{m-n,\gamma}^2\}$ .

The range of possible values of  $x_j$  can be obtained by maximizing and minimizing  $x_j$  under the constraint  $v(x) \leq \chi_{m-n,\gamma}^2$ . (In the fuzzy case, we have to repeat the computations for every  $p$ .)

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

This condition is equivalent to

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

**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  $\sigma_i$ , we get the case of pure interval uncertainty. In this case, the above formulas tend to the set of all the values for which

$$\left| y_i - \sum_{j=1}^n a_{ij} \cdot x_j \right| \leq \Delta_i.$$



For example, for  $m$  repeated measurements of the same quantity, we get the intersection of the corresponding intervals. So, the new idea is indeed a generalization of the known probabilistic and interval approaches.

**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 the 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; see, e.g., [1, 8]. 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  $v(x) \leq \chi_{m-n,\gamma}^2$ .

Similarly, we can efficiently compute  $v$  and thus, check whether  $v \leq \chi_{m-n,\gamma}^2$  (i.e., whether the measurement results are consistent or we have outliers).

The set  $S_\gamma$  is convex. We can approximate the set  $S_\gamma$  by:

- taking a grid  $G$ ,
- checking, for each  $x \in G$ , whether  $v(x) \leq \chi_{m-n,\gamma}^2$ , and
- taking the convex hull of “possible” points.

We can also efficiently find the minimum  $\underline{x}_j$  of  $x_j$  over  $x \in S_\gamma$ . By computing the minimum of the linear function  $-x_j$ , we can thus efficiently compute the largest possible values  $\bar{x}_j$  of  $x_j$  over  $x \in S_\gamma$ .

## 4 Discussion

**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?

To answer this question, let’s recall that we get  $\sigma_i$  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  $\Delta_i$  come from?

The answer is simple: after the 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 the numerical estimate  $b$  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  $\Delta_i$  comes from: it is simply the largest possible value from this interval, i.e.,

$$\Delta_i = \max(\bar{b} - b, b - \underline{b}).$$

**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  $-\int \rho(x) \cdot \ln(\rho(x)) dx$ , where  $\rho(x)$  is the corresponding probability density function; see, e.g., [4].

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

If a random variable  $\eta$  (corresponding to random error component) is distributed with the probability density function  $\rho(x)$ , then the sum of  $\eta$  and an  $m$ -dimensional uniform distribution has the density  $\rho'(x) = \max_{e_i^s \in [-\Delta_i, \Delta_i]} \rho(x - e^s)$ . For this distribution, the maximum likelihood method  $\rho'(x) \rightarrow \max$  is equivalent to minimizing  $-\ln(\rho'(x)) \rightarrow \min$ , where  $-\ln(\rho'(x)) = \min_{e_i^s \in [-\Delta_i, \Delta_i]} (-\ln(\rho(x - e^s)))$ .

In particular, for the normal distribution with 0 mean,

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

Thus, the maximum likelihood approach  $\rho'(x) \rightarrow \max$  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 mean that we can safely assume that the systematic error is uniformly distributed on  $[-\Delta_i, \Delta_i]$ ? This is, e.g., what International Organization for Standardization (ISO) suggests; see [2, 12]. Our answer is: not always.

Indeed, e.g., for the sum  $s = x_1 + \dots + x_m$  of  $m$  such errors with  $\Delta_i = \Delta$  all we can say is that  $s$  belongs to the interval  $[-m \cdot \Delta, m \cdot \Delta]$ . All the values from this interval are clearly possible.

However, if we assume uniform distributions, then, for large  $m$ , due to the Central Limit Theorem, the sum  $s$  is practically normally distributed, with 0 mean and standard deviation proportional to  $\sqrt{m} \cdot \sigma$ .

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.

Our conclusion is that in some calculations, we can use MaxEnt and uniform distributions, but not always. In other words, we must be cautious.

## Acknowledgments

This work was performed when Vladik was a visiting researcher with the Geodetic Institute of the Leibniz University of Hannover, a visit supported by the German Science Foundation. This work was also supported in part by NSF grant HRD-1242122.

## References

1. D. P. Bertsekas, *Convex Optimization Algorithms*, Athena Scientific, Belmont, Massachusetts, 2015.
2. International Organization for Standardization (ISO), *ISO/IEC Guide 98-3:2008, Uncertainty of Measurement – Part 3: Guide to the Expression of Uncertainty in Measurement (GUM:1995)*, 2008.
3. L. Jaulin, M. Kiefer, O. Dicrit, and E. Walter, *Applied Interval Analysis*, Springer, London, 2001.
4. E. T. Jaynes and G. L. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.
5. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
6. D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, Springer, Cham, Switzerland, 2016.
7. R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.
8. Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Springer Verlag, Berlin, Heidelberg, New York, 2013.
9. H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
10. P. V. Novitskii and I. A. Zograph, *Estimating the Measurement Errors*, Energoatomizdat, Leningrad, 1991 (in Russian).
11. A. I. Orlov, “How often are the observations normal?”, *Industrial Laboratory*, 1991, Vol. 57, No. 7, pp. 770–772.
12. S. G. Rabinovich, *Measurement Errors and Uncertainty: Theory and Practice*, Springer Verlag, Berlin, 2005.
13. D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.
14. S. Schön, “Interval-based reliability and integrity measures”, *Proceedings of the 8th European Space Agency (ESA) Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing NAVITEC’2016*, Noordwijk, Netherlands, December 14–16, 2016.
15. S. Schön and H. Kutterer, “Using zonotopes for overestimation-free interval least-square: some geodetic applications”, *Reliable Computing*, 2005, Vol. 11, pp. 137–155.
16. G. Seeber, *Satellite Geodesy*, de Gruyter, Berlin, Göttingen, 2003.
17. L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.