6-2008

Trade-Off Between Sample Size and Accuracy: Case of Measurements under Interval Uncertainty

Hung T. Nguyen

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Scott Ferson

Comments:
Technical Report: UTEP-CS-07-47c
Published in *International Journal of Approximate Reasoning*, 2009, Vol. 50, No. 8, pp. 1164-1176.

Trade-Off Between Sample Size and Accuracy: Case of Measurements under Interval Uncertainty ⋆

Hung T. Nguyen [a], Olga Kosheleva [b],
Vladik Kreinovich [c,*], and Scott Ferson [d]

[a] *Department of Mathematical Sciences, New Mexico State University Las Cruces, NM 88003, USA*

[b] *Department of Teacher Education, University of Texas, El Paso, TX 79968, USA*

[c] *Department of Computer Science, University of Texas, El Paso, TX 79968, USA*

[d] *Applied Biomathematics, 100 North Country Road, Setauket, NY 11733, USA*

Abstract

In many practical situations, we are not satisfied with the accuracy of the existing measurements. There are two possible ways to improve the measurement accuracy:

- first, instead of a *single* measurement, we can make *repeated* measurements; the additional information coming from these additional measurements can improve the accuracy of the result of this series of measurements;

- second, we can replace the *current* measuring instrument with a *more accurate* one; correspondingly, we can use a more accurate (and more expensive) measurement procedure provided by a measuring lab – e.g., a procedure that includes the use of a higher quality reagent.

In general, we can combine these two ways, and make *repeated* measurements with a *more accurate* measuring instrument. What is the appropriate trade-off between sample size and accuracy? This is the general problem that we address in this paper.

Key words: measurements, interval uncertainty, accuracy, sample size, trade-off

1 General formulation of the problem

We often need more accurate measurement procedures. Measurements are never 100% accurate, there is always a measurement inaccuracy.

Manufacturers of a measuring instrument usually provide the information about the accuracy of the corresponding measurements. In some practical situations, however, we want to know the value of the measured quantity with the accuracy which is higher than the guaranteed accuracy of a single measurement.

Comment. Measurements are provided either by a *measuring instrument* or, in situations like measuring level of pollutants in a given water sample, by a *measuring lab*. Most problems related to measurement accuracy are the same, whether we have an automatic device (measuring instrument) or operator-supervised procedure (measuring lab). In view of this similarity, in the following text, we will consider the term "measuring instrument" in the general sense, so that the measuring lab is viewed as a particular case of such (general) measuring instrument.

Two ways to improve the measurement accuracy: increasing sample size and improving accuracy. There are two possible ways to improve the measurement accuracy:

- first, instead of a *single* measurement, we can make *repeated* measurements; the additional information coming from these additional measurements can improve the accuracy of the result of this series of measurements;
- second, we can replace the *current* measuring instrument with a *more accurate* one; correspondingly, we can use a more accurate (and more expensive)

⋆ This work was supported in part by NSF grants HRD-0734825, EAR-0225670, and EIA-0080940, by Texas Department of Transportation grant No. 0-5453, by the Max Planck Institut für Mathematik, and by the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant 2006-08. The authors are thankful to all the participants of the International Workshop on Interval and Probabilistic Uncertainty and Non-Classical Logic UncLog'08, JAIST, Japan, March 25–28, 2008, for valuable discussions, and to the anonymous referees for important suggestions.
* Corresponding author
Email addresses: hunguyen@nmsu.edu (Hung T. Nguyen), olgak@utep.edu (Olga Kosheleva), vladik@utep.edu (Vladik Kreinovich), scott@ramas.com (Scott Ferson).

measurement procedure provided by a measuring lab – e.g., the procedure that includes the use of a higher quality reagent.

In general, we can combine these two ways, and make *repeated* measurements with a *more accurate* measuring instrument.

Problem: finding the best trade-off between sample size and accuracy. What guidance shall we give to an engineer in this situation? Shall she make repeated measurements with the original instrument? shall she instead purchase a more accurate measuring instrument and make repeated measurements with this new instrument? How more accurate? how many measurement should we perform? In other words, what is the appropriate trade-off between sample size and accuracy?

This is the general problem that we address in this paper.

2 In different practical situations, this general problem can take different forms

There are two different situations which, crudely speaking, correspond to engineering and to science.

In most practical situations – in *engineering*, *ecology*, etc. – we know what accuracy we want to achieve. In *engineering*, this accuracy comes, e.g., from the tolerance with which we need to guarantee some parameters of the manufactured object. To make sure that these parameters fit into the tolerance intervals, we must measure them with the accuracy that is as good as the tolerance. For example, if we want to guarantee, e.g., the resistance of a certain wire does not deviate from its nominal value by more than 3%, then we must measure this resistance with an accuracy of at least 3% (or better).

In *ecological* measurements, we want to make sure that the measured quantity does not exceed the required limit. For example, if we want to guarantee that the concentration of a pollutant does not exceed 0.1 units, then we must be able to measure this concentration with an accuracy somewhat higher than 0.1. In such situations, our objective is to minimize the cost of achieving this accuracy.

In *science*, we often face a different objective:

- we have a certain amount of funding allocated for measuring the value of a certain quantity;

- within the given funding limits, we would like to determine the value of the measured quantity as accurately as possible.

In other words:

- In engineering situations, we have a fixed accuracy, and we want to minimize the measurement cost.
- In scientific situations, we have a fixed cost, and we want to maximally improve the measurement accuracy.

3 A realistic formulation of the trade-off problem

Traditional engineering approach. The traditional engineering approach to solving the above problem is based on the following assumptions – often made when processing uncertainty in engineering:

- that all the measurement uncertainties ("measurement errors") are normally (Gaussian) distributed, with known standard deviations σ;
- that the measurement uncertainties corresponding to different measurements are independent random variables; and
- that the mean value Δ_s of the measurement uncertainty is 0.

Under these assumptions, if we repeat a measurement n times and compute the arithmetic average of n results, then this average approximates the actual value with a standard deviation $\frac{\sigma}{\sqrt{n}}$. So, under the above assumptions, by selecting appropriate large number of iterations n, we can make measurement uncertainties as small as we want.

This approach – and more general statistical approach – has been actively used in many applications to science in engineering problems; see, e.g., [5,6,15,19].

Limitations of the traditional approach. In practice, the distributions are often Gaussian and independent; however, the mean Δ_s (sometimes called "systematic error" in engineering practice) is not necessarily 0. Let us show this if we do not take this bias $\Delta_s \neq 0$ into account, we will underestimate the resulting measurement inaccuracy.

Indeed, suppose that we have a measuring instrument about which we know that its measurement uncertainty cannot exceed 0.1: $|\Delta x| \leq 0.1$. This means, e.g., that if, as a result of the measurement, we got the value $\widetilde{x} = 1.0$, then the actual (unknown) value x $(= \widetilde{x} - \Delta x)$ of the measured quantity can take any value from the interval $[1.0 - 0.1, 1.0 + 0.1] = [0.9, 1.1]$.

If the bias component of the measurement uncertainty is 0, then we can repeat the measurement many times and, as a result, get more and more accurate estimates of x. However, if – as is often the case – we do not have any information about the bias, it is quite possible that the bias is actually equal to 0.07 (and the un-biased component $\Delta x - \Delta_s$ of the measurement uncertainty is negligible in comparison with this bias). In this case, the measured value 1.0 means that the actual value of the measured quantity was $x = 1.0 - 0.07 = 0.93$. In this situation, we can repeat the measurement many times, and every time, the measurement result will be equal to $\approx x + \Delta_s = 0.93 + 0.07 = 1.0$. The average of these values will still be approximately equal to 1.0 – so, no matter how many times we repeat the measurement, we will get the exact same measurement uncertainty 0.07.

In other words, when we are looking for a trade-off between sample size and accuracy, the traditional engineering assumptions can result in misleading conclusions.

A more realistic description of measurement uncertainty. We do not know the actual value of the bias Δ_s – if we knew this value, we could simply re-calibrate the measuring instrument and thus eliminate this bias.

What we do know are the bounds on the bias. Specifically, in measurement standards (see, e.g., [17]), we are usually provided with the upper bound Δ on the bias – i.e., with a value Δ for which $|\Delta_s| \le \Delta$. In other words, the only information that we have about the measurement bias Δ_s is that it belongs to the *interval* $[-\Delta, \Delta]$.

Resulting formulas for the measurement accuracy. Under these assumptions, what is the guaranteed accuracy of a single measurement made by the measuring instrument?

Formally, a normally distributed random variable can take any value from $-\infty$ to $+\infty$. In reality, when the value is too far away from the average, its probability is practically negligible. In practice, it is usually assumed that the values which differ from the average a by more than $k_0 \cdot \sigma$ are impossible – where the value k_0 is determined by how confident we want to be:

- 95% confidence corresponds to $k_0 = 2$,
- 99.9% corresponds to $k_0 = 3$, and
- confidence $100\% - 10^{-6}\%$ corresponds to $k_0 = 6$.

Thus, with selected confidence, we know that the measurement uncertainty is between $\Delta_s - k_0 \cdot \sigma$ and $\Delta_s + k_0 \cdot \sigma$. Since the bias Δ_s can take any value from

$-\Delta$ to $+\Delta$, the smallest possible value of the overall measurement uncertainty is $-\Delta - k_0 \cdot \sigma$, and the largest possible value of the overall measurement uncertainty is $\Delta + k_0 \cdot \sigma$.

Thus, for a measuring instrument with

- a standard deviation σ of the un-biased component of measurement uncertainty and
- an upper bound Δ on the bias,

the overall measurement uncertainty is bounded by the value $\Delta + k_0 \cdot \sigma$, where the value k_0 is determined by the desired confidence level.

Resulting formulas for the accuracy of a repeated measurement. When we repeat the same measurement n times and take the average of n measurement results, the bias remains the same, while the standard deviation of the un-biased component of the measurement uncertainty decreases $\sqrt{n}$ times. Thus, after n measurements, the overall measurement uncertainty is bounded by the value $\Delta + k_0 \cdot \dfrac{\sigma}{\sqrt{n}}$.

So, we arrive at the following formulation of the trade-off problem.

Trade-off problem for engineering. In the situation when we know the overall accuracy Δ_0, and we want to minimize the cost of the resulting measurement, the trade-off problem takes the following form:

$$\text{Minimize } n \cdot F(\Delta, \sigma) \text{ under the constraint } \Delta + k_0 \cdot \frac{\sigma}{\sqrt{n}} \le \Delta_0, \tag{1}$$

where $F(\Delta, \sigma)$ is the cost of a single measurement performed by a measuring instrument whose bias is bounded by Δ and for which the un-biased uncertainty component has a standard deviation σ.

Trade-off problem for science. In the situation when we are given the limit F_0 on the cost, and the problem is to achieve the highest possible accuracy within this cost, we arrive at the following problem

$$\text{Minimize } \Delta + k_0 \cdot \frac{\sigma}{\sqrt{n}} \text{ under the constraint } n \cdot F(\Delta, \sigma) \le F_0. \tag{2}$$

4 Solving the trade-off problem in the general case

Mathematical comment. The number of measurement n is a discrete variable. In general, optimization with respect to discrete variables requires much more computations than continuous optimization (see, e.g., [9]). Since our formulation is approximate anyway, we will treat n as a real-valued variable – with the idea that in a practical implementation, we should take, as the actual sample size, the closest integer to the corresponding real number solution $n_{\rm opt}$.

Towards resulting formulas. For both constraint optimization problems, the Lagrange multiplier method leads to the following unconstraint optimization problem:

$$n \cdot F(\Delta, \sigma) + \lambda \cdot \left(\Delta + k_0 \cdot \frac{\sigma}{\sqrt{n}} - \Delta_0\right) \to \min_{\Delta, \sigma, n}, \tag{3}$$

where λ can be determined by one of the formulas

$$\Delta + k_0 \cdot \frac{\sigma}{\sqrt{n}} = \Delta_0, \quad n \cdot F(\Delta, \sigma) = F_0. \tag{4}$$

Equating the derivatives of the objective function (with respect to the unknowns Δ, σ, and n) to 0, we conclude that

$$n \cdot \frac{\partial F}{\partial \Delta} + \lambda = 0; \quad n \cdot \frac{\partial F}{\partial \sigma} + \lambda \cdot \frac{k_0}{\sqrt{n}} = 0; \quad F - \frac{1}{2} \cdot \lambda \cdot k_0 \cdot \frac{\sigma}{n^{3/2}} = 0. \tag{5}$$

Substituting the expression for λ from the first equation into the second one, we conclude that

$$n = k_0^2 \cdot \frac{(\partial F/\partial \Delta)^2}{(\partial F/\partial \sigma)^2}. \tag{6}$$

Substituting these expression into the other equations from (5) and into the equations (4), we get the following non-linear equations with two unknowns Δ and σ:

$$F + \frac{1}{2} \cdot \sigma \cdot \frac{\partial F}{\partial \sigma} = 0; \tag{7}$$

$$\Delta + \frac{\sigma \cdot (\partial F/\partial \sigma)}{\partial F/\partial \Delta} = \Delta_0; \quad k_0^2 \cdot \frac{(\partial F/\partial \Delta)^2}{(\partial F/\partial \sigma)^2} \cdot F = F_0. \tag{8}$$

So, we arrive at the following algorithm:

General formulas: results. For each of the optimization problems (1) and (2), to find the optimal accuracy values Δ and σ and the optimal sample size n, we do the following:

- First, we determine the optimal accuracy, i.e., the optimal values of Δ and σ, by solving a system of two non-linear equations with two unknowns Δ and σ: the equation (7) and one of the equations (8) (depending on what problem we are solving).
- After that, we determine the optimal sample size n by using the formula (6).

For practical engineering problems, we need more explicit and easy-to-use recommendations. The above formulas provide a general theoretical solution to the trade-off problem, but to use them in practice, we need more easy-to-use recommendations. In practice, however, we do not have the explicit formula $F(\Delta, \sigma)$ that determines how the cost of the measurement depends on its accuracy. Therefore, to make our recommendations more practically useful, we must also provide some guidance on how to determine this dependence – and then use the recommended dependence to simplify the above recommendations.

5 How Does the Cost of a Measurement Depend on Its Accuracy?

Two characteristics of uncertainty: Δ and σ. In our description, we use two parameters to characterize the measurement's accuracy: the upper bound Δ on the bias and the standard deviation σ of the un-biased component of measurement uncertainty.

It is difficult to describe how the cost of a measurement depends on σ. The standard deviation σ is determined by the noise level, so decreasing σ requires a serious re-design of the measuring instrument. For example, to get a standard measuring instrument, one thing designers usually do is place the instrument in liquid helium so as to eliminate the thermal noise as much as possible; another idea is to place the measuring instrument into a metal cage, to eliminate the effect of the outside electromagnetic fields on the measuring instrument's electronics.

Once we have eliminated the obvious sources of noise, eliminating a new source of noise is a creative problem, requiring a lot of ingenuity, and it is difficult to estimate how the cost of such decrease depends on σ.

The inability to easily describe the dependence of cost on σ may not be that crucial. The inability to easily handle the characteristic σ of the un-biased component of measurement uncertainty may not be so bad because, as we have mentioned, the un-biased component is the one that can be drastically decreased by increasing the sample size – in full accordance with the traditionally used simplifying engineering assumptions about uncertainty.

As we have mentioned, in terms of decreasing the overall accuracy, it is much more important to decrease the bias, i.e., to decrease the value Δ. Let us therefore analyze how the cost of a measurement depends on Δ.

How we can reduce Δ: reminder. As we have mentioned, we can decrease the characteristic Δ of the bias by calibrating our measuring instrument against the standard one.

After N repeated measurements, we get a bias Δ_s whose standard deviation is $\approx \sigma/\sqrt{N}$ (and whose distribution, due to the Central Limit Theorem, is close to Gaussian). Thus, with the same confidence level as we use to bound the overall measurement uncertainty, we can conclude that $|\Delta_s| \le k_0 \cdot \sigma/\sqrt{N}$.

Calibration is not a one-time procedure. To properly take calibration into account, it is important to recall that calibration is not a one-time procedure. Indeed, most devices deteriorate with time. In particular, measuring instruments, if not periodically maintained, become less and less accurate. Because of this, in measurement practices, calibration is not a one-time procedure, it needs to be done periodically.

How frequently do we need to calibrate a device? The change of Δ_s with time t is slow and smooth. A smooth dependence can be represented by a Taylor series $\Delta_s(t) = \Delta_s(0) + k \cdot t + c \cdot t^2 + \ldots$ In the first approximation, we can restrict ourselves to the main – linear – term (linear trend) in this expansion, and thus, in effect, assume that the change of Δ_s with time t is linear.

Thus, if by calibrating the instrument, we guaranteed that $|\Delta_s| \le \Delta$, then after time t, we can only guarantee that $|\Delta_s| \le \Delta + k \cdot t$. Once the upper bound on Δ_s reaches the level that we do not want to exceed, this means that a new calibration is in order. Usually (see, e.g., [17]), to guarantee the bound Δ throughout the entire calibration cycle, we, e.g., initially calibrate it to be

below $\Delta/2$, and then re-calibrate at a time t_0 when $\Delta/2 + k \cdot t_0 = \Delta$. In such a situation, the time t_0 between calibrations is equal to $t_0 = \Delta/(2 \cdot k)$.

How the calibration-based reduction procedure translates into the cost of a measurement: the main case. As we have just mentioned, the way to decrease Δ is to calibrate the measuring instrument. Thus, the resulting *additional* cost of a measurement comes from the cost of this calibration (spread over all the measurements performed between calibrations).

Comment. Of course, the overall cost of the measurement also includes other costs: the cost of buying (or designing) the measuring instrument, the cost of actually performing the individual measurement themselves, etc. However, these costs are usually order of magnitude smaller than the costs of calibration – since the calibration requires the use of a drastically more expensive standard measuring instrument. Thus, in our first-approximation analysis, we will ignore the costs of actual measurements and assume that the overall costs are approximately equal to the calibration costs.

These is an additional reason why the costs of actual measurement can be safely ignored when we compare measurement corresponding to different values Δ: costs are (approximately) the same for all values Δ; what increases when we want a smaller Δ is only the cost of the corresponding calibration.

Each calibration procedure consists of two stages:

- first, we transport the measuring instrument to the location of a standard – e.g., to the National Institute of Standard and Technology (NIST) or one of the regional standardization centers – and set up the comparison measurements by the tested and the standard instruments;
- second, we perform the measurements themselves.

Correspondingly, the cost of calibration can be estimated as the sum of the costs of these two stages.

Operating the standard measuring instrument is usually a very expensive procedure. So, setting it up for comparison with different measuring instruments requires a lot of time and a lot of adjustment. Once the set-up is done, the second stage is fast and automatic – and therefore not that expensive.

As a result, usually, the cost of the first stage is the dominating factor. So, we can reasonably assume that the cost of the calibration is just the cost of the set-up – i.e., the cost of the first stage of the calibration procedure.

By definition, the set-up does not depend on how many times N we perform the comparison measurements. Thus, in the first approximation, we can simply assume that each calibration requires a flat rate f_0.

The interval between time calibrations is $t_0 = \Delta/(2 \cdot k)$, then during a fixed period of time T_0 (e.g., 10 years), we need

$$\frac{T_0}{t_0} = \frac{T_0}{\Delta/(2 \cdot k)} = \frac{2 \cdot k \cdot T_0}{\Delta}$$

calibrations. Multiplying this number by the cost f_0 of each calibration, we get the overall cost of all the calibrations performed during the fixed time T_0 as $\dfrac{2 \cdot k \cdot T_0 \cdot f_0}{\Delta}$. Finally, dividing this cost by the estimated number N_0 of measurements performed during the period of time T_0, we estimate the cost $F(\Delta)$ of an individual measurement as

$$F(\Delta) = \frac{c}{\Delta}, \tag{9}$$

where we denoted

$$c \stackrel{\text{def}}{=} \frac{2 \cdot k \cdot T_0 \cdot f_0}{N_0}. \tag{10}$$

Comment. The above formula was first described, in a somewhat simplified form, in [7].

This formula is in good accordance with chemistry-related measurements. It is worth mentioning that the dependence $c \sim 1/\Delta$ also occurs in measurements related to chemical analysis. Indeed, in these measurements, the accuracy of the measurement result is largely determined by the quality of the reagents, i.e., mainly, by the concentration level δ of the unwanted chemicals (pollutants) in a reagent mix. Specifically, the maximum possible measurement uncertainty Δ is proportional to this concentration δ, i.e., $\Delta \approx c_0 \cdot \delta$.

According to [20], the cost of reducing pollutants to a level δ is proportional to $1/\delta$. Since the accuracy Δ is proportional to δ, the dependence of the cost of the accuracy is also inverse proportional to Δ, i.e., $F(\Delta) = c/\Delta$ for some constant c.

This formula is in good accordance with actual prices of different measurements. This dependence is in good agreement by the experimental data on the cost of measurements of chemical-related measurements. For

example, in a typical pollution measurement, a measurement with the 25% accuracy costs $\approx$ \$200, while if we want to get 7% accuracy, then we have to use a better reagent grade in our measurements which costs between \$500 and \$1,000. Here, the 3–4 times increase in accuracy (i.e., 3–4 times decrease in measurement uncertainty) leads to approximately the same (4–5) times increase in cost – which is indeed in good accordance with the dependence $F(\Delta) \approx c/\Delta$.

How the calibration-based reduction procedure translates into the cost of a measurement: cases of more accurate measurements. In deriving the formula $F(\Delta) \approx c/\Delta$, we assumed that the cost of actually performing the measurements with the standard instrument is much smaller than the cost of setting up the calibration experiment. This is a reasonable assumption if the overall number of calibration-related measurement N is not too large.

How many measurement do we need? After N measurements, we get the accuracy $\Delta = k_0 \cdot \sigma/\sqrt{N}$. Thus, for a measuring instrument with standard deviation σ, if we want to achieve the bias level Δ, we must use

$$N = k_0 \cdot \frac{\sigma^2}{\Delta^2} \tag{11}$$

measurements.

When the desired accuracy Δ is very small, the number of calibration-related measurements N is therefore very large. For large N, the duration of the calibration-related measurements exceeds the duration of the set-up. Since the most expensive part of the calibration procedure is the use of the standard measuring instrument, the cost of this procedure is proportional to the overall time during which we use this instrument. When N is large, this time is roughly proportional to N.

In this case, instead of a flat fee f_0, the cost of each calibration becomes proportional to N, i.e., equal to $f_1 \cdot N$, where f_1 is the cost per time of using the standard measuring instrument multiplied by the time of each calibration measurement. Due to the formula (11), the resulting cost of each calibration is equal to $f_1 \cdot k_0 \cdot \dfrac{\sigma^2}{\Delta^2}$. To get the cost of a single measurement, we must multiply this cost by the number of calibrations $\dfrac{2 \cdot k \cdot T_0}{\Delta}$ required during the time period T_0, and then divide by the typical number of measurements performed during this period of time. As a result, the cost of a single measurement becomes $\dfrac{\text{const}}{\Delta^3}$.

The cost of measurements beyond calibration: general discussion. In many scientific cutting-edge experiments, we want to achieve higher accuracy than was possible before. In such situations, we cannot simply use the existing standard measuring instrument to calibrate the new one, because we want to achieve the accuracy that no standard measuring instrument has achieved earlier.

In this case, how can we increase the accuracy depends on the specific quantity that we want to measure.

The cost of measurements beyond calibration: example. For example, in radioastrometry – the art of determining the locations of celestial objects from radioastronomical observation – the accuracy of a measurement by a single radio telescope is $\Delta \approx \lambda/D$, where λ is the wavelength of the radiowaves on which we are observing the source, and D is the diameter of the telescope; see, e.g., [21]. For a telescope of a linear size D, just the amount of material is proportional to its volume, i.e., to D^3; the cost F of designing a telescope is even higher – it is proportional to D^4. Since $D \approx \text{const}/\Delta$, in this case, we have $F(\Delta) \approx \text{const}/\Delta^4$.

The cost of measurements beyond calibration: power laws. The above dependence is a particular case of the *power law* $F(\Delta) \approx \text{const}/\Delta^\alpha$. Power laws are, actually, rather typical descriptions of the dependence of the cost of an individual measurement on its accuracy.

In [13], we explain why in the general case, power laws are indeed reasonable approximation: crudely speaking, in the absence of a preferred value of the measured quantity, it is reasonable to assume that the dependence does not change if we change the measuring unit (i.e., that it is scale invariant), and power laws are the only scale-invariant dependencies.

Comment. The same arguments about scale invariance apply when we try to find out how the cost of a measurement depends on the standard deviation. So, it is reasonable to assume that this dependence is also described by a power law $F(\sigma) \approx \text{const}/\sigma^\beta$ for some constant β.

6 Trade-off between accuracy and sample size in different cost models

Let us plug in different cost models into the above general solution for the tradeoff problem and find out what is the optimal trade-off between accuracy and sample size in the above cost models.

Since the above cost models only describe the dependence of the cost of Δ and n, we will assume that the characteristic σ (of the un-biased component of measurement uncertainty) is fixed, so we can only select the accuracy characteristic Δ and the sample size n.

Basic cost model: engineering situation. Let us start with the basic cost model, according to which $F(\Delta) = c/\Delta$. Within this model, we can explicitly solve the above system of equations. As a result, for the engineering situation, we conclude that

$$n_{\text{opt}} = \frac{9 \cdot k_0^2 \cdot \sigma^2}{4 \cdot \Delta_0^2}; \quad \Delta_{\text{opt}} = \frac{1}{3} \cdot \Delta_0. \tag{12}$$

Observation. In this case, the overall Δ_0 on the measurement uncertainty is the sum of the bounds coming from two uncertainty components:

- the bound Δ_0 that comes from the bias component, and
- the bound $k_0 \cdot \dfrac{\sigma}{\sqrt{n}}$ that comes from un-biased component of the measurement uncertainty.

In the optimal trade-off, the first component is equal to 1/3 of the bound on overall measurement uncertainty, and therefore, the second component is equal to 2/3 of the bound on overall measurement uncertainty. As a result, we conclude that when the measurement uncertainty comes from several components, in the optimal trade-off, these uncertainty components are of approximately the same size.

Heuristic consequence of this observation. As a result of this qualitative idea, it is reasonable to use the following heuristic rule when looking for a good (not necessarily optimal) trade-off: split the overall measurement uncertainty into equal parts.

In the above example, this would mean taking $\Delta = (1/2) \cdot \Delta_0$ (and, correspondingly, $k_0 \cdot \dfrac{\sigma}{\sqrt{n}} = (1/2) \cdot \Delta_0$) instead of the optimal value $\Delta = (1/3) \cdot \Delta_0$.

How non-optimal is this heuristic solution?

For the optimal solution $\Delta = (1/3) \cdot \Delta_0$, the resulting value of the objective function (1) (representing the overall measurement cost) is $\frac{27}{4} \cdot \frac{k_0^2 \cdot \sigma^2 \cdot c}{\Delta_0^2}$, while for $\Delta = (1/2) \cdot \Delta_0$, the cost is $8 \cdot \frac{k_0^2 \cdot \sigma^2 \cdot c}{\Delta_0^2}$ – only $\approx 20\%$ larger.

If we take into account that all our models are approximate, this means that the heuristic trade-off solution is practically as good as the optimal one.

Basic cost model: science situation. In the science situation (2), we get

$$n_{\text{opt}} = \left(\frac{F_0 \cdot k_0 \cdot \sigma}{2 \cdot c}\right)^{2/3}; \quad \Delta_{\text{opt}} = \frac{n_{\text{opt}} \cdot c}{F_0}. \tag{13}$$

Cases of more accurate and cutting-edge measurements. When $F(\Delta) = c/\Delta^\alpha$, for the engineering case, we get

$$n_{\text{opt}} = \frac{(\alpha+2)^2 \cdot k_0^2 \cdot \sigma^2}{4 \cdot \Delta_0^2}; \quad \Delta_0 = \frac{\alpha}{2+\alpha} \cdot \Delta_0.$$

For the science case,

$$n_{\text{opt}} = \left(\frac{F_0}{c}\right)^{2/(2+\alpha)} \cdot \left(\frac{k_0 \cdot \alpha}{2}\right)^{(2\alpha)/(2+\alpha)}; \quad \Delta_{\text{opt}} = \frac{\alpha}{2} \cdot k_0 \cdot \frac{\sigma}{\sqrt{n_{\text{opt}}}}.$$

In both cases, the uncertainty bound coming from the bias is approximately equal to the bound coming from the un-biased component of measurement uncertainty.

7 Case of dynamic measurements

Up to now, we have considered the case of static measurements, when the measured quantity is static (does not change over time). Let us now analyze the general case of *dynamic* measurements, when the measured quantity changes over time.

For such dynamic quantities, we may have two different objectives:

- We may be interested in knowing the *average* value of the measured quantity, e.g., the average concentration of a pollutant in a lake or the average day

temperature. In addition to to knowing the average, we may also want to know the standard deviation and/or other statistical characteristics.
- We may also want to know not only the average, but also the actual dependence of the measured quantity on space location and/or time.

For example:

- If we are interested in general weather patterns, e.g., as a part of the climatological analysis, then it is probably sufficient to measure the average temperature (or the average wind velocity) in a given area.
- On the other hand, if our intent is to provide the meteorological data to the planes flying in this area, then we would rather know how exactly the wind velocity depends on the location, so that the plane will be able to avoid locations where the winds are too strong.

In this paper, we analyze the trade-off between accuracy and sample size for both objectives.

8 First objective: measuring the average value of a varying quantity

Case of ideal measuring instruments: analysis. Let us start to analyze this situation with the case of an ideal measuring instrument, i.e., a measuring instrument for which the measurement uncertainty is negligible.

By using this ideal instrument, we can measure the value of the quantity of interest at different points and at different moments of time. After we perform n measurements and get n measurement results $x_1, \ldots, x_n$, a natural way to estimate the desired mean value $x_0 = E[x]$ of x is to use the arithmetic average $E \stackrel{\text{def}}{=} \dfrac{x_1 + \ldots + x_n}{n}$ of these measured values. It is reasonable to assume that the differences $x_i - x_0$ are independent random variables, with a known standard deviation σ_0.

In this case, due to the Central Limit Theorem, for large n, the difference $\Delta x_0 \stackrel{\text{def}}{=} E - x_0$ between the estimate E and the desired value x_0 is approximately normally distributed with 0 average and standard deviation $\sigma_0/\sqrt{n}$.

So, even for measurements with the ideal measuring instrument, the result E of measuring x_0 is not exact; we can only guarantee (with the corresponding level of confidence) that the measurement uncertainty Δx_0 is bounded by the value $k_0 \cdot \sigma_0/\sqrt{n}$.

Comment. If we do not know this standard deviation, we can estimate it based on the measurement results $x_1, \ldots, x_n$, by using the standard statistical formulas, such as

$$\sigma_0 \approx \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - E)^2}.$$

Case of ideal measuring instruments: recommendations. In the case of ideal measuring instruments, if we want to achieve the desired overall accuracy Δ_0 with a given confidence, then the sample size n must be determined by the condition that $k_0 \cdot \sigma_0 / \sqrt{n} \le \Delta_0$, where k_0 corresponds to this confidence:

- 95% confidence corresponds to $k_0 = 2$,
- 99.9% corresponds to $k_0 = 3$, and
- confidence $1 - 10^{-6}\%$ corresponds to $k_0 = 6$.

The above condition is equivalent to $\sqrt{n} \ge \dfrac{k_0 \cdot \sigma_0}{\Delta_0}$, i.e., to $n \ge \dfrac{k_0^2 \cdot \sigma_0^2}{\Delta_0^2}$. To minimize the measurement costs, we must select the smallest sample size for which this inequality holds, i.e., select $n \approx \dfrac{k_0^2 \cdot \sigma_0^2}{\Delta_0^2}$.

Case of realistic measuring instruments: description. In practice, measuring instruments are not perfect, they have measurement uncertainty. Usually, we assume that we know the standard deviation σ of the corresponding measurement uncertainty, and we know the upper bound Δ on the possible values of the bias Δ_s: $|\Delta_s| \le \Delta$; see, e.g., [17].

Case of realistic measuring instruments: analysis. For realistic measuring instruments, for each measurement, the difference $\Delta x_i = \widetilde{x}_i - x_i$ between the measured and actual values of the quantity of interest is no longer negligible.

In this case, based on n measurement results $\widetilde{x}_1, \ldots, \widetilde{x}_n$, we do not get the arithmetic average E of the *actual* values, we only get the average

$$\widetilde{E} = \frac{\widetilde{x}_1 + \ldots + \widetilde{x}_n}{n}$$

of the *measured* values. We are using this average $\widetilde{E}$ as an estimate for the desired average x_0. There are two reasons why $\widetilde{E}$ is different from x_0:

- first, due to measurement uncertainty, $\widetilde{x}_i \ne x_i$, hence $\widetilde{E} \ne E$;
- second, due to the finite sample size, $E \ne x_0$.

As a result, the uncertainty Δx_0 with which this procedure measures x_0, i.e., the difference $\Delta x_0 \stackrel{\text{def}}{=} \widetilde{E} - x_0$, can be represented as the sum of two uncertainty components:

$$\widetilde{E} - x_0 = (\widetilde{E} - E) + (E - x_0). \tag{14}$$

If we use a measuring instrument for which the mean (bias) of the measurement uncertainty is Δ_s and standard deviation is σ, then for the difference of arithmetic averages, the mean is the same value Δ_s (bias) and the standard deviation is $\sqrt{n}$ times smaller: it is equal to $\sigma/\sqrt{n}$. We have just described the difference $E - x_0$: it is a random variable with 0 mean and standard deviation $\sigma_0/\sqrt{n}$.

Since the mean value of $E - x_0$ is 0 (by definition of x_0 as the mean of x_i), the mean value of the sum (14) is equal to the mean value of the first uncertainty component, i.e., to Δ_s.

It is reasonable to assume that the measurement uncertainty $\widetilde{x}_i - x_i$ (caused by the imperfections of the measurement procedure) and the deviations $x_i - x_0$ (caused by variability of the quantity of interest) are independent random variables. In this case, the variance of the sum (14) is equal to the sum of the corresponding variances, i.e., to

$$\frac{\sigma^2}{n} + \frac{\sigma_0^2}{n} = \frac{\sigma_t^2}{n},$$

where we denoted $\sigma_t \stackrel{\text{def}}{=} \sqrt{\sigma^2 + \sigma_0^2}$. Hence, the standard deviation of the total measurement uncertainty is equal to $\sigma_t/\sqrt{n}$.

So, the measurement uncertainty $\widetilde{E} - x_0$ is approximately normally distributed, with the mean Δ_s (about which we know that $|\Delta_s| \le \Delta$) and the standard deviation $\sigma_t/\sqrt{n}$. Thus, we can conclude that with a selected degree of confidence, the overall measurement uncertainty cannot exceed $\Delta + k_0 \cdot \dfrac{\sigma_t}{\sqrt{n}}$.

Case of realistic measuring instruments: recommendations. From the purely mathematical viewpoint, when the standard deviation σ of a measuring instrument is fixed, then, to determine Δ and n, we get exactly the same formulas as in the case of static measurements, with the only difference that:

- instead of the standard deviation σ of the measuring instrument,
- we now have the combined standard deviation $\sigma_t = \sqrt{\sigma^2 + \sigma_0^2}$ of the measuring instrument and of the measured quantity.

So, all the recommendations that we have developed for static measurements are also applicable here.

Example. If we want to achieve a given accuracy Δ_0 with the smallest possible cost, then, according to the above text, we should use the measuring instrument with accuracy $\Delta \approx (1/3) \cdot \Delta_0$. The sample size n is then determined by the formula $k_0 \cdot \dfrac{\sigma_t}{\sqrt{n}} = (2/3) \cdot \Delta_0$.

For measuring average, the optimal accuracy Δ if the same as for static measurements, but the optimal sample size is now determined by a new formula $n_{\text{opt}} = \dfrac{9 \cdot k_0^2 \cdot \sigma_t^2}{4 \cdot \Delta_0^2}$, with σ_t instead of σ. Since $\sigma_t > \sigma$, we will need a larger sample size n.

9 Second objective: measuring the actual dependence of the measured quantity on space location and/or on time

Formulation of the problem. In many real-life situations, we are interested not only in the average value of the measured quantity x, we are also interested in the actual dependence of this quantity on space and/or time.

Within this general scheme, there are several possible situations:

- We may have a quantity that does not depend on a spatial location but does depend on time – e.g., we may be interested in the temperature at a given location. In this case, we are interested to learn how this quantity x depends on the time t, i.e., we are interested to know the dependence $x(t)$.
- We may be interested in a quantity that does not change with time but does change from one spatial location to the other. For example:
  - in a geographic analysis, we may be interested in how the elevation x depends on the 2-D spatial location $t = (t_1, t_2)$;
  - in a geophysical analysis, we may be interested how in the density depends on a 3-D location $t = (t_1, t_2, t_3)$ inside the Earth.
- Finally, we may be interested in a quantity that changes both with time and from one spatial location to the other. For example:
  - we may be interested in learning how the surface temperature depends on time t_1 and on the 2-D spatial location (t_2, t_3);
  - we may be also interested in learning how the general temperature in the atmosphere depends on time t_1 and on the 3-D spatial location (t_2, t_3, t_4).

In all these cases, we are interested to know the dependence $x(t)$ of a measured quantity on the point $t = (t_1, \ldots, t_d)$ in d-dimensional space, where the

dimension d ranges from 1 (for the case when we have a quantity depending on time) to 4 (for the case when we are interested in the dependence both on time and on the 3-D spatial location).

Measurement inaccuracy caused by the finiteness of the sample. In practice, we can only measure the values of x at finitely many different locations, and we must use extrapolation to find the values at other locations. So, even if we use a perfect measuring instrument, for which the measurement uncertainty can be ignored, we still have an uncertainty caused by extrapolation.

For example, suppose that we have measured the values $x(t^{(i)})$ of the quantity x at moments of time $t^{(1)} < t^{(2)} < \ldots, < t^{(n)}$, and we want to describe the value $x(t)$ of this quantity at a different moment of time $t \neq t^{(i)}$, a moment of time at which no measurement has been made.

In practice, for most systems, we know the limit g on how fast the value of the quantity x can change with time (or from one spatial location to the other). So, when, e.g., $t^{(1)} < t < t^{(2)}$, we can conclude that $|x(t)-x(t^{(1)})| \leq g\cdot|t-t^{(1)}|$, i.e., that $x(t) \in [x(t^{(1)}) - g \cdot |t - t^{(1)}|, x(t^{(1)}) + g \cdot |t - t^{(1)}|]$. Thus, even when we have an ideal measuring instrument, the fact that we only have a finite sample $t^{(1)}, \ldots, t^{(n)}$ leads to uncertainty in our knowledge of the values $x(t)$ for $t \neq x^{(i)}$.

Estimate of the measurement uncertainty for a given measurement accuracy and given sample size. Let us consider a general situation when we perform measurements with a guaranteed accuracy Δ, and when we measure the quantity x at n different points $t^{(1)}, \ldots, t^{(n)}$ in the d-dimensional space. As a result of this measurement, we get n values $\widetilde{x}_i$ that are Δ-close to the actual values of the quantity x at the corresponding point $t^{(i)}$: $|\widetilde{x}_i-x(t^{(i)})| \leq \Delta$.

If we are interested in the value $x(t)$ of the quantity x at a point $t \neq t^{(i)}$, then we have to use one of the measured values $\widetilde{x}_i$.

We assume that we know the rate g with which $x(t)$ changes with t. Thus, if we use the the result $\widetilde{x}_i$ of measuring $x(t^{(i)})$ to estimate $x(t)$, we can guarantee that $|x(t^{(i)}) - x(t)| \leq g \cdot \rho(t, t^{(i)})$, where $\rho(a, b)$ denotes the distance between the two points in the d-dimensional space. Since $|\widetilde{x}_i - x(t^{(i)})| \leq \Delta$, we can thus conclude that $|\widetilde{x}_i - x(t)| \leq |\widetilde{x}_i - x(t^{(i)})| + |x(t^{(i)}) - x(t)| \leq \Delta + g \cdot \rho(t, t^{(i)})$, i.e.,

$$|\widetilde{x}_i - x(t)| \leq \Delta + g \cdot \rho(t, t^{(i)}). \tag{15}$$

Thus, the smaller the distance between t and $t^{(i)}$, the smaller the resulting

measurement uncertainty. So, to get the most accurate estimate for $x(t)$, we must select, for this estimate, the point $t^{(i)}$ which is the closest to t.

In general, once we fix the accuracy Δ, the sample size n, and the points $t^{(1)}, \ldots, t^{(n)}$ at which the measurement are performed, we can guarantee that for every t, the value $x(t)$ can be reconstructed with the accuracy $\Delta + g \cdot \rho_0$, where ρ_0 is the largest possible distance between a point t and the sample set $\{t^{(1)}, \ldots, t^{(n)}\}$.

Thus, once we fixed Δ and n, we should select the points $t^{(i)}$ in such a way that this "largest distance" ρ_0 attains the smallest possible value.

In the 1-D case, the corresponding allocation is easy to describe. Indeed, suppose that we want to allocate such points $t^{(i)}$ on the interval $[0, T]$. We want to minimize the distance ρ_0 corresponding to a given sample size n – or, equivalently, to minimize the sample size given a distance ρ_0. Every point t is ρ_0-close to one of the sample points $t^{(i)}$, so it belongs to the corresponding interval

$$[t^{(i)} - \rho_0, t^{(i)} + \rho_0].$$

Thus, the interval $[0, T]$ of width T is covered by the union of n intervals $[t^{(i)} - \rho_0, t^{(i)} + \rho_0]$ of widths $2\rho_0$. The width T of the covered interval cannot exceed the sum of the widths of the covering intervals, so we have $T \le n \cdot (2\rho_0)$, hence always $\rho_0 \ge T/(2n)$. Actually, we can have $\rho_0 = T/2n$ if we select the points $t^{(i)} = (i - 0.5) \cdot (T/n)$. Then:

- for the values $t \in [0, T/n]$, we take, as the estimate for $x(t)$, the result $\widetilde{x}_1$ of measuring $x(t^{(1)}) = x(T/(2n))$;
- for the values $t \in [T/n, 2T/n]$, we take, as the estimate for $x(t)$, the result $\widetilde{x}_2$ of measuring $x(t^{(2)}) = x((3/2) \cdot (T/n))$;
- ...
- for the values $t \in [(i-1) \cdot T/n, i \cdot T/n]$, we take, as the estimate for $x(t)$, the result $\widetilde{x}_i$ of measuring $x(t^{(i)}) = x((i - 1/2) \cdot (T/n))$;
- ...

So, the optimal location of points is when they are on a grid $t^{(1)} = 0.5 \cdot T/n$, $t^{(2)} = 1.5 \cdot T/n$, $t^{(3)} = 2.5 \cdot T/n$, ..., and each point $t^{(i)}$ "serves" the values t from the corresponding interval $[(i-1) \cdot T/n, i \cdot T/n]$ (the interval that contains this point $t^{(i)}$ as its center), serves in the sense that for each point t from this interval, as the measured value of $x(t)$, we take the value $x^{(i)}$. These intervals corresponding to individual points $t^{(i)}$ cover the entire interval $[0, T]$ without intersection,

In this optimal location, when we perform n measurements, we get $\rho_0 = T/(2n)$.

Similarly, in the general d-dimensional case, we can place n points on a d-dimensional grid. In this case, each point $t^{(i)}$ "serves" the corresponding cube; these cubes cover the whole domain without intersection. If we denote, by V, the d-dimensional volume of the spatial (or spatio-temporal) domain that we want to cover, then we can conclude that each point $x^{(i)}$ serves the cube of volume V/n. Since the volume of a d-dimensional cube of linear size Δt is equal to $(\Delta t)^d$, we can thus conclude that the linear size of each of the cubes serves by a measurement point is $(V/n)^{1/d}$.

Within this cube, each point $t^{(i)}$ is located at the center of the corresponding cube. Thus, for each point t within this cube and for each coordinate j, the absolute value $|t_j - t_j^{(i)}|$ between the j-th coordinate of this point t and the j-th coordinate of the cube's center $t^{(i)}$ does not exceed one half of the cube's linear size: $|t_j - t_j^{(i)}| \le (1/2) \cdot (V/n)^{1/d}$. Therefore, for

$$\rho(t, t^{(i)}) = \sqrt{\left(t_1 - t_1^{(i)}\right)^2 + \ldots + \left(t_d - t_d^{(i)}\right)^2},$$

we get

$$\rho(t, t^{(i)}) \le \rho \stackrel{\text{def}}{=} \sqrt{d \cdot \left(\frac{1}{2} \cdot \left(\frac{V}{n}\right)^{1/d}\right)^2} = \sqrt{d} \cdot \frac{1}{2} \cdot \frac{V^{1/d}}{n^{1/d}}.$$

We have already mentioned that for every point t, the accuracy with which we can reconstruct $x(t)$ is bounded by the value $\Delta + g \cdot \rho_0$. Thus, this accuracy is bounded by $\Delta + g \cdot \sqrt{d} \cdot \frac{1}{2} \cdot \frac{V^{1/d}}{n^{1/d}}$.

We are now ready to formally describe the corresponding trade-off problems.

Trade-off problems for engineering and science: formulation. In engineering applications, we know the overall accuracy Δ_0, and we want to minimize the cost of the resulting measurement. In this case, the trade-off problem takes the following form:

$$\text{Minimize } n \cdot F(\Delta) \to \min_{\Delta, n} \text{ under the constraint } \Delta + \frac{g_0}{n^{1/d}} = \Delta_0, \qquad (16)$$

where $F(\Delta)$ is a cost of a single measurement made by a measuring instrument with accuracy Δ, and we denoted

$$g_0 \stackrel{\text{def}}{=} g \cdot \sqrt{d} \cdot \frac{1}{2} \cdot V^{1/d}. \qquad (17)$$

In scientific applications, when we are given the cost F_0, and the problem is to achieve the highest possible accuracy within this cost. In this case, we arrive

at the following problem

$$\text{Minimize } \Delta + \frac{g_0}{n^{1/d}} \to \min_{\Delta,n} \text{ under the constraint } n \cdot F(\Delta) = F_0. \tag{18}$$

Engineering situation: solution. For the basic cost model $F(\Delta) = c/\Delta$, the engineering problem (16) has the following solution:

$$\Delta_{\text{opt}} = \frac{1}{d+1} \cdot \Delta_0; \quad n_{\text{opt}} = \left(\frac{g_0}{\Delta_0} \cdot \frac{d+1}{d} \right)^d. \tag{19}$$

Similarly to the static case, the optimal trade-off between accuracy and the sample size is attained when both uncertainty components are of approximately the same size.

Science situation: solution. For the basic cost model $F(\Delta) = c/\Delta$, the science problem (16) has the following solution:

$$n_{\text{opt}} = \left(\frac{F_0}{c} \cdot \frac{g_0}{d} \right)^{d/(d+1)}; \quad \Delta_{\text{opt}} = \frac{n_{\text{opt}} \cdot c}{F_0}. \tag{20}$$

In this case too, in the optimal trade-off, the uncertainty bound coming from the accuracy of individual measurements is approximately equal to the uncertainty bound coming from the finiteness of the sample.

Case of non-smooth processes: how to describe them. In the above text, we considered the case the dependence of the quantity x on time and/or space t is smooth. In this case, for small changes Δt, this dependence can be approximately described by a linear function $x(t + \Delta t) = x(t) + g_1 \cdot \Delta t_1 + \ldots + g_d \cdot \Delta t_d$. So, if we know the upper bound g on the length $\|(g_1, \ldots, g_d)\|$ of the gradient of $x(t)$, we can bound the difference $x(t + \Delta t) - x(t)$ between the values of the quantity x at close points $t + \Delta t$ and t by the product $g \cdot \|\Delta t\| = g \cdot \rho(t, t + \Delta t)$.

In practice, we often encounter non-smooth processes. For example, meteorological data exhibit random change (similar to the Brownian motion); as the result of this, the dependence of the corresponding quantities x on time and spatial coordinates is not smooth.

For the particular case of a Brownian motion, the difference between the values of the quantity x at nearly points grows as the square root of the distance between these points: $|x(t+\Delta t)-x(t)| \le C \cdot \|\Delta t\|^{1/2}$ for some real number C. In

many physical processes, this dependence can be described by a more general power law, i.e., $|x(t+\Delta t)-x(t)| \le C \cdot \|\Delta t\|^{\beta}$ for some real numbers C and $\beta \in (0,1)$. Such processes are a particular case of *fractals*; see, e.g., [12] (This notion is closely related with the notion of a fractal dimension: namely, the graph of the corresponding dependence $x(t)$ has a fractal dimension $d+(1-\beta)$.)

In [13], it is explained why scale invariance naturally leads to the power law – and thus, to the fractal dependence.

Measurement uncertainty in the case of non-smooth processes. Let us use these formulas to estimate measurement uncertainty for the case of non-smooth processes. We have already mentioned that if we perform (appropriately located) n measurements in a d-dimensional space, then the distance from each point t of the domain of interest to one of the points $t^{(i)}$ in which the measurement was made does not exceed $\rho_0 = \sqrt{d} \cdot \dfrac{1}{2} \cdot \dfrac{V^{1/d}}{n^{1/d}}$.

In the fractal case, we can conclude that the uncertainty of approximating the desired value $x(t)$ with the measured value $x(t^{(i)})$ does not exceed $C \cdot \rho^{\beta}$. Thus, if we perform n measurements with a measuring device of accuracy Δ, the resulting accuracy in reconstructing all the values of $x(t)$ is bounded by the value

$$\Delta + C \cdot \rho_0^{\beta} = \Delta + C \cdot d^{\beta/2} \cdot \frac{1}{2^{\beta}} \cdot \frac{V^{\beta/d}}{n^{\beta/d}} = \Delta + \frac{g_{\beta}}{n^{\beta/d}},$$

where we denoted

$$g_{\beta} \stackrel{\text{def}}{=} C \cdot d^{\beta/2} \cdot \frac{1}{2^{\beta/d}} \cdot V^{\beta/d}.$$

Trade-off problems for engineering and science: formulation and solution. In the situation when we know the overall accuracy Δ_0, and we want to minimize the cost of the resulting measurement, the trade-off problem takes the following form:

$$\text{Minimize } n \cdot F(\Delta) \text{ under the constraint } \Delta + \frac{g_{\beta}}{n^{\beta/d}} = \Delta_0. \tag{21}$$

In the situation when we are given the limit F_0 on the cost, and the problem is to achieve the highest possible accuracy within this cost, we arrive at the following problem

$$\text{Minimize } \Delta + \frac{g_{\beta}}{n^{\beta/d}} \text{ under the constraint } n \cdot F(\Delta) = F_0. \tag{22}$$

From the mathematical viewpoint, these formulas are similar to the formulas corresponding to the smooth case, with the only difference that instead of raising n to the power $1/d$, we now raise n to the power $1/d'$, where $d' \stackrel{\text{def}}{=} d/\beta$.

Thus, for the basic cost model $F(\Delta) = c/\Delta$, the engineering problem has the following solution:

$$\Delta_{\text{opt}} = \frac{\beta}{d+\beta} \cdot \Delta_0; \quad n_{\text{opt}} = \left(\frac{g_\beta}{\Delta_0} \cdot \frac{d+\beta}{d} \right)^d. \tag{23}$$

For the basic cost model $F(\Delta) = c/\Delta$, the science problem has the following solution:

$$n_{\text{opt}} = \left(\frac{F_0}{c} \cdot \frac{g_\beta}{d} \right)^{d/(d+\beta)}; \quad \Delta_{\text{opt}} = \frac{n_{\text{opt}} \cdot c}{F_0}. \tag{24}$$

in this case too, in the optimal trade-off, both uncertainty components are of approximately the same value.

Case of more accurate measuring instruments. In the above text, we have shown that for more accurate measuring instrument, the cost $F(\Delta)$ of a measurement depends on its accuracy as $F(\Delta) = c/\Delta^\alpha$. Once we go beyond the basic cost model $\alpha = 1$, we get $\alpha = 3$, and then, as we increase accuracy, we switch to a different value α.

For such a power law, in the engineering case, the optimal accuracy is $\Delta_{\text{opt}} = \dfrac{\alpha}{\alpha+2} \cdot \Delta_0$. In particular, for $\alpha = 3$, we have $\Delta_{\text{opt}} = \dfrac{3}{5} \cdot \Delta_0$.

10 Case study: in brief

A real-life example in which we used similar arguments to made a selection between the accuracy and the sample size is the design of radioastronomical telescope system [1–4,8,10,11]. As we have mentioned, for the radiotelescope of diameter D, the measurement accuracy is proportional to λ/D, and the cost is proportional to D^4.

The design of a large system of radiotelescopes has several objectives:

- first, we would like to solve *radioastrometry* problems, i.e., determine the location of the radiosources with as much accuracy as possible;

- second, we would like to solve the *radioimaging* problems, i.e., for each of the radiosources, we would like to know not only its location, but also its *image* – i.e., how the intensity (and polarization) of the source changes from one point of this source to the other.

In the first problem, we are interested in measuring a well-defined unchanging quantity. In the second problem, we are interested in finding the actual dependence of the measured quantity on the spatial location.

In the second problem, similar to what we discussed in the general case, the more samples we take (i.e., the more telescopes we build), the more points we will get on the image. On the other hand, within a given overall cost, if we build more telescopes, then the amount of money allocated to each telescope will be smaller, so each telescope will be small ($D' \ll D$), and the resulting accuracy $\Delta \sim 1/D$ of each of the many measurements will be not so good.

In our analysis, we have found an optimal trade-off between accuracy and sample size. This analysis was used in the design of the successful Russian network of radiotelescopes.

11 Conclusions

In many practical situations, we are not satisfied with the accuracy of the existing measurements. There are two possible ways to improve the measurement accuracy. First, instead of a *single* measurement, we can make *repeated* measurements; the additional information coming from these additional measurements can improve the accuracy of the result of this series of measurements. Second, we can replace the *current* measuring instrument with a *more accurate* one; correspondingly, we can use a more accurate (and more expensive) measurement procedure provided by a measuring lab – e.g., a procedure that includes the use of a higher quality reagent. In general, we can combine these two ways, and make *repeated* measurements with a *more accurate* measuring instrument.

What is the appropriate trade-off between sample size and accuracy? Traditional engineering approach to this problem assumes that we know the exact probability distribution of all the measurement uncertainties. In many practical situations, however, we do not know the exact distributions. For example, we often only know the upper bound on the corresponding measurement (or estimation) uncertainty; in this case, after the measurements, we only know the interval of possible values of the quantity of interest. We first show that in such situations, traditional engineering approach can sometimes be misleading, so for interval uncertainty, new techniques are needed. Then, we describe

proper techniques for achieving optimal trade-off between sample size and accuracy under interval uncertainty.

In general, if the measurement uncertainty consists of several components, then the optimal trade-off between the accuracy Δ and the same size n occurs when these components are approximately of the same size.

In particular, if we want to achieve the overall accuracy Δ_0, as a first approximation, it is reasonable to take $\Delta = \Delta_0/2$ – and select the sample size for which the resulting overall measurement uncertainty is Δ_0.

A more accurate description of optimal selections in different situations is as follows:

- for the case when we measure a single well-defined quantity (or the average value of varying quantity), we should take $\Delta = \dfrac{1}{3} \cdot \Delta_0$;
- for the case when we are interested in reconstructing all the values $x(t)$ of a smooth quantity x depending on d parameters $t = (t_1, \ldots, t_d)$, we should take $\Delta = \dfrac{1}{d+1} \cdot \Delta_0$;
- for the case when are interested in reconstructing all the values $x(t)$ of a non-smooth quantity x depending on d parameters $t = (t_1, \ldots, t_d)$, we should take $\Delta = \dfrac{\beta}{d+\beta} \cdot \Delta_0$, where β is the exponent of the power law that describes how the difference $x(t + \Delta t) - x(t)$ changes with $\|\Delta t\|$;
- for the case of more accurate measuring instruments, when the cost $F(\Delta)$ of a single measurement starts growing as c/Δ^3, we should take $\Delta = \dfrac{3}{5} \cdot \Delta_0$. In general, if $F(\Delta) = c/\Delta^\alpha$, we should take $\Delta = \dfrac{\alpha}{\alpha+2} \cdot \Delta_0$.

References

[1] A. Dravskikh, A.M. Finkelstein, V. Kreinovich, "Astrometric and geodetic applications of VLBI 'arc method'", *Modern Astrometry, Proceedings of the IAU Colloquium No. 48*, Vienna, 1978, 143–153.

[2] A. F. Dravskikh, O. M. Kosheleva, A. M. Finkelstein, P. A. Fridman, "Possibility of using a reference-object method to form a phase-stable multielement long-baseline interferometric system", *Bulletin of the Special Astrophysical Observatory – North Caucasus*, Allerton Press, N. Y., Vol. 16, 1984, pp. 72–80.

[3] A.F. Dravskikh, O.M. Kosheleva, V. Kreinovich, A.M. Finkelstein, "The method of arcs and differential astrometry," *Soviet Astronomy Letters*, 1979, Vol. 5, No. 3, pp. 160–162.

[4] A. F. Dravskikh, O. M. Kosheleva, V. Ya. Kreinovich, A. M. Finkelstein, "Optimization of the procedure for measuring arcs by radiointerferometry", *Soviet Astronomy Letters*, 1979, Vol. 5, No. 4, pp. 227–228.

[5] J. Eng, "Sample Size Estimation: How Many Individuals Should Be Studied?", *Radiology*, 2003, Vol. 227, pp. 309–313.

[6] J. J. Hox, *Multilevel Analysis: Techniques and Applications*, Lawrence Erlbaum Associates, 2002.

[7] V. Kreinovich, *How to compute the price of a measuring instrument?*, Leningrad Center for New Information Technology "Informatika", Technical Report, Leningrad, 1989 (in Russian).

[8] V. Kreinovich, A. Bernat, O. Kosheleva, A. Finkelstein, "Interval estimates for closure phase and closure amplitude imaging in radio astronomy", *Interval Computations*, 1992, No. 2(4), pp. 51–71.

[9] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1997.

[10] V. Kreinovich, S. A. Starks, D. Iourinski, O. Kosheleva, and A. Finkelstein, "Open-ended configurations of radio telescopes: a geometrical analysis", *Geombinatorics*, 2003, Vol. 13, No. 2, pp. 79–85.

[11] V. Kreinovich, S. A. Starks, O. Kosheleva, and A. Finkelstein, "Open-ended configurations of radio telescopes: towards optimal design", *Proceedings of the 2002 World Automation Congress WAC'2002*, Orlando, Florida, June 9–13, 2002, pp. 101–106.

[12] B. B. Mandelbrot, *The fractal geometry of Nature*, Freeman, San Francisco, 1982.

[13] H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.

[14] P. V. Novitskii and I. A. Zograph, *Estimating the Measurement Errors*, Energoatomizdat, Leningrad, 1991 (in Russian).

[15] N. A. Obuchowski and D. K. Mcclish, "Sample Size Determination For Diagnostic Accuracy Studies Involving Binormal ROC Curve Indices", *Statistics in Medicine*, 1997, Vol. 16, No. 13, pp. 1529–1542.

[16] A. I. Orlov, "How often are the observations normal?", *Industrial Laboratory*, 1991, Vol. 57. No. 7, pp. 770–772.

[17] S. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, Springer-Verlag, New York, 2005.

[18] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[19] A. K. Skidmore, "Accuracy assessment of spatial information", In: *Spatial Statistics for Remote Sensing*, Springer Netherlands, 2002, pp. 197–209.

[20] D. Stevens, "Analysis of biological systems", *Proceedings of the NIH MARC Winter Institute on Undergraduate Education in Biology*, Santa Cruz, California, January 7–11, 2005.

[21] G. L. Verschuur and K. I. Kellerman, "Galactic and Extragalactic Radio Astronomy", Springer-Verlag, N.Y., 1988.