Comparisons of Measurement Results as Constraints on Accuracies of Measuring Instruments: When Can We Determine the Accuracies from These Constraints?

Christian Servin¹ and Vladik Kreinovich² ¹Information Technology Department El Paso Community College 919 Hunter, El Paso TX 79915, USA cservin@gmail.com ²Department of Computer Science University of Texas at El Paso El Paso, TX 79968, USA vladik@utep.edu

Abstract

For a measuring instrument, a usual way to find the probability distribution of its measurement errors is to compare its results with the results of measuring the same quantity with a much more accurate instrument. But what if we are interested in estimating the measurement accuracy of a state-of-the-art measuring instrument, for which no more accurate instrument is possible? In this paper, we show that while in general, such estimation is not possible; however, can uniquely determine the corresponding probability distributions if we have several state-of-the-art measuring instruments, and for one of them, the corresponding probability distribution is symmetric.

1 Formulation of the Problem

Need to determine accuracies of measurement instruments. Most information comes from measurements. Measurement results are never absolutely accurate: the measurement result \tilde{x} is, in general, different from the actual (unknown) value x of the corresponding quantity; see, e.g., [7]. To properly process data, it is therefore important to know how accurate are our measurements.

Ideally, we would like to know what are the possible values of measurement errors $\Delta x \stackrel{\text{def}}{=} \tilde{x} - x$, and how frequent are different possible values of Δx . In

other words, we would like to know the probability distribution on the set of all possible values of the measurement error Δx .

How accuracies are usually determined: by using a second, much more accurate measuring instrument. One usual way to find the desired probability distribution is to have a second measuring instrument which is much more accurate than the one that we want to estimate. In this case, the measurement error $\Delta x_2 = \tilde{x}_2 - x$ of this second instrument is much smaller than $\Delta x = \tilde{x} - x$ and thus, the difference $\tilde{x} - \tilde{x}_2 = (\tilde{x} - x) - (\tilde{x}_2 - x)$ between the two measurement results can serve as a good approximation to the measurement error. From the sample of such differences, we can therefore find the desired probability distribution for Δx .

What if we do not have a more accurate measuring instrument? But what if the measuring instrument whose accuracy we want to estimate is among the best? In this case, we do not have a much more accurate measuring instrument. What can we do in this case?

In such situations, we can use the fact that there usually, there are *several* measuring instrument of the type that we want to analyze. Due to measurement errors, for the same quantity, these instruments, in general, produce slightly different measurement results. It is therefore desirable to try to extract the information about measurement accuracies from the differences between these measurement results.

Two possible situations. In some cases, we have a stable manufacturing process that produces several practical identical measuring instruments, for which the probability distributions of measurement error are the same. In such cases, all we need to find is this common probability distribution.

In other cases, we cannot ignore the differences between different instruments. In this case, for each individual measuring instrument, we need to find its own probability distribution.

What is known: case of normal distribution. In many practical situations, the measurement error is caused by the joint effect of numerous independent small factors. In such situations, the Central Limit Theorem (see, e.g., [9]) implies that this distribution is close to Gaussian.

A Gaussian distribution is uniquely determined by its mean (bias) and standard deviation σ . When we only know the differences, we cannot determine the bias: it could be that all the measuring instruments have the same bias, and we will never determine that since we only see the differences. Thus, it makes sense to limit ourselves only to the *random component* of the measurement error, i.e., to the measurement error minus its mean value.

For this "re-normalized" measurement error Δx , the mean is 0. So, all we need to determine is the standard deviation σ . These standard deviations can indeed be determined; see, e.g., [4, 8].

Specifically, hen we have two identical independent measuring instruments, with normally distributed measurement errors Δx_1 and Δx_2 , then the difference $\tilde{x}_2 - \tilde{x}_1$ is also normally distributed, with variance $V = \sigma^2 + \sigma^2 = 2\sigma^2$. Thus,

once we experimentally determine the variance V of this observable difference, we can compute the desired variance σ^2 as $\sigma^2 = \frac{V}{2}$.

When we have several different measuring instruments, with unknown standard deviations $\sigma_1, \sigma_2, \sigma_3, \ldots$, then for each observable difference $\tilde{x}_i - \tilde{x}_j$ the variance is equal to $V_{ij} = \sigma_i^2 + \sigma_j^2$. Thus, once we experimentally determine the three variances V_{12}, V_{23} , and V_{13} , we can find the desired standard deviations by solving the corresponding system of three equations with three unknowns: $V_{12} = \sigma_1^2 + \sigma_2^2, V_{23} = \sigma_2^2 + \sigma_3^2$, and $V_{13} = \sigma_1^2 + \sigma_3^2$, whose solution is:

$$\sigma_1^2 = \frac{V_{12} + V_{13} - V_{23}}{2}, \quad \sigma_2^2 = \frac{V_{12} + V_{23} - V_{13}}{2},$$
$$\sigma_3^2 = \frac{V_{13} + V_{23} - V_{12}}{2}.$$

Problem: what if distributions are not Gaussian? Empirical analysis of measuring instruments shows that only slightly more than a half of them have Gaussian measurement errors [3, 6]. What happens in the non-Gaussian case? In such cases, sometimes, we simply cannot uniquely reconstruct the corresponding distributions; see, e.g., [8]. In this paper, we explain when such a reconstruction is possible and when it is not possible.

2 Idea: Let Us Use Moments

Motivation for using moments. As we have mentioned, a Gaussian distribution with zero mean is uniquely determined by its second moment $M_2 = \sigma^2$. This means that all higher moments $M_k \stackrel{\text{def}}{=} E[(\Delta x)^k]$ are uniquely determined by the value M_2 .

In general, we may have values of M_k which are different from the corresponding Gaussian values. Thus, to describe a general distribution, in addition to the second moment, we also need to describe its higher moments as well.

Moments are sufficient to uniquely describe a distribution: reminder. But even if we know all the moments, will it be sufficient to uniquely determine the corresponding probability distribution? The answer is yes, it is possible, and let us provide a simple reminder of why it is possible – and how can we reconstruct the corresponding distribution.

The usual way to represent a probability distribution of a random variable Δx is by describing its probability density function (pdf) $\rho(\Delta x)$. In many situations, it is convenient to use its *characteristic function*

$$\chi(\omega) \stackrel{\text{def}}{=} E[\exp(\mathbf{i} \cdot \omega \cdot \Delta x)],$$

where $i \stackrel{\text{def}}{=} \sqrt{-1}$, i.e.,

$$\chi(\omega) = \int \rho(\Delta x) \cdot \exp(\mathbf{i} \cdot \omega \cdot \Delta x) \, d\Delta x.$$

From the mathematical viewpoint, the characteristic function is the Fourier transform of the pdf, and it is known that we can uniquely reconstruct a function from its Fourier transform (this reconstruction is known as the *inverse Fourier transform*); see, e.g., [1, 2, 5, 10].

On the other hand, if we use Taylor expansion of the exponential function

$$\exp(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \ldots + \frac{z^k}{k!} + \ldots,$$

then the characteristic functions takes the form

$$\chi(\omega) = E\left[1 + \mathbf{i} \cdot \omega \cdot \Delta x - \frac{1}{2!} \cdot \omega^2 \cdot (\Delta x)^2 + \ldots + \frac{\mathbf{i}^k}{k!} \cdot \omega^k \cdot (\Delta x)^k + \ldots\right],$$

i.e.,

$$\chi(\omega) = 1 - \frac{1}{2} \cdot \omega^2 \cdot M_2 + \ldots + \frac{\mathbf{i}^k}{k!} \cdot \omega^k \cdot M_k + \ldots$$

Thus, if we know all the moments M_k , we can uniquely reconstruct the characteristic function and thus, uniquely reconstruct the desired pdf.

Important fact: for a symmetric distribution, odd moments are zeros. In the following analysis, it is important to use the fact that for a symmetric distribution, i.e., a distribution for which $\rho(-\Delta x) = \rho(\Delta x)$, add odd moments M_{2s+1} are equal to 0:

$$M_{2s+1} = \int \rho(\Delta x) \cdot (\Delta x)^{2s+1} \, d\Delta x.$$

Indeed, if we replace Δx to $\Delta x' \stackrel{\text{def}}{=} -\Delta x$, then $d\Delta x = -d\Delta x'$, $(\Delta x)^{2s+1} = -(\Delta x')^{2s+1}$ and thus, the above integral takes the form

$$M_{2s+1} = -\int \rho(-\Delta x') \cdot (\Delta x')^{2s+1} d\Delta x' = -\int \rho(\Delta x') \cdot (\Delta x')^{2s+1} d\Delta x',$$

so $M_{2s+1} = -M_{2s+1}$ and hence, $M_{2s+1} = 0$.

3 Case When Have Several Identical Measuring Instruments

Description of the case: reminder. In this cases, we have several measuring instruments, with the same probability distribution and thus, with the same moments M_2 , M_3 , etc. The only available information consists of the differences $\Delta x_1 - \Delta x_2 = \tilde{x}_1 - \tilde{x}_2$. Based on the observations, we can determine the probability distribution for each such difference, and thus, we can determine the moments M'_k of this difference.

We would like to use these observable moments $M'_k = E[(\Delta x_1 - \Delta x_2)^k]$ to find the desired differences $M_k = E[(\Delta x)^k]$.

What is known: case of second moments. For k = 2, we have $M'_2 = 2M_2$ and thus, we can uniquely reconstruct the desired second moment M_2 from the observed second moment M'_2 .

Natural next case: third moments. Can we similarly reconstruct the desired third moment $M_3 = E[(\Delta x)^3]$ based on the observed third moment $M'_{3} = E[(\Delta x_{1} - \Delta x_{2})^{3}]?$ Here,

$$(\Delta x_1 - \Delta x_2)^3 = (\Delta x_1)^3 - 3 \cdot (\Delta x_1)^2 \cdot \Delta x_2 + 3 \cdot \Delta x_1 \cdot (\Delta x_2)^2 - (\Delta x_2)^3$$

so, due to linearity of the mean and to the fact that the measurement errors Δx_1 and Δx_2 corresponding to two measuring instruments are assumed to be independent, we conclude that

$$M'_{3} = E[(\Delta x_{1} - \Delta x_{2})^{3}] = E[(\Delta x_{1})^{3}] - 3 \cdot E[(\Delta x_{1})^{2}] \cdot E[\Delta x_{2}] + 3 \cdot E[\Delta x_{1}] \cdot E[(\Delta x_{2})^{2}] - E[(\Delta x_{2})^{3}].$$

In this case, $E[\Delta x_{i}] = 0$ and $E(\Delta x_{1})^{3}] = E[(\Delta x_{2})^{3}] = M_{3}$, so

$$M_3' = M_3 - M_3 = 0.$$

In other words, the observed third moment M'_3 is always equal to 0, and thus, carries no information about M_3 .

So, the only case when we can reconstruct M_3 is when we know it already. One such case is when we know that the distribution is symmetric. In turns out that in this case, we can reconstruct all the moments and thus, we can uniquely reconstruct the original probability distribution.

When the probability distribution of the measurement error is symmetric, this distribution can be uniquely determined from the observed differences. For a symmetric distribution, all odd moments are equal to 0. Thus, to uniquely determine a symmetric distribution, it is sufficient to determine all its even moments M_{2s} . Let us prove, by induction, that we can reconstruct all these even moments.

We already know that we can reconstruct M_2 . Let us assume that we already know how to reconstruct the moments M_2, \ldots, M_{2s} . Let us show how to reconstruct the next moment $M_{2s+2} = E[(\Delta x)^{2s+2}]$. For this, we will use the observed moment $M'_{2s+2} = E[(\Delta x_1 - \Delta x_2)^{2s+2}]$. Here,

$$(\Delta x_1 - \Delta x_2)^{2s+2} = (\Delta x_1)^{2s+2} - (2s+2) \cdot (\Delta x_1)^{2s+1} \cdot \Delta x_2 + \frac{(2s+2) \cdot (2s+1)}{1 \cdot 2} \cdot (\Delta x_1)^{2s} \cdot (\Delta x_2)^2 - \dots + \frac{(2s+2) \cdot (2s+1)}{1 \cdot 2} \cdot (\Delta x_1)^2 \cdot (\Delta x_2)^{2s} - (2s+2) \cdot \Delta x_1 \cdot (\Delta x_2)^{2s+1} + (\Delta x_2)^{2s+2}$$
Thus

Thus,

$$M'_{2s+2} = E[(\Delta x_1)^{2s+2}] - (2s+2) \cdot E[(\Delta x_1)^{2s+1}] \cdot E[\Delta x_2] +$$

$$\frac{(2s+2)\cdot(2s+1)}{1\cdot 2}\cdot E[(\Delta x_1)^{2s}]\cdot E[(\Delta x_2)^2] - \ldots + \frac{(2s+2)\cdot(2s+1)}{1\cdot 2}\cdot E[(\Delta x_1)^2]\cdot E[(\Delta x_2)^{2s}] - (2s+2)\cdot E[\Delta x_1]\cdot E[(\Delta x_2)^{2s+1}] + E[(\Delta x_2)^{2s+2}],$$

i.e.,

$$M'_{2s+2} = M_{2s+2} + \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_{2s} \cdot M_2 + \dots + \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_2 \cdot M_{2s} + M_{2s+2}.$$

Thus,

$$2M_{2s+2} = M'_{2s+2} - \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_{2s} \cdot M_2 - \dots - \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_2 \cdot M_{2s}.$$

We know the value M'_{2s+2} , and we assumed that we have already shown that we can uniquely determine the moments M_2, \ldots, M_{2s} . Thus, we can indeed uniquely determine the moment M_{2s+2} .

Induction proves that we can indeed determine all the even moments.

4 Case When Have Several Different Measuring Instruments

Description of the case: reminder. In this case, we have several measuring instruments with, in general, different probability distributions. For each of the measuring instruments i, we want to find the corresponding moments

$$M_{k,i} = E[(\Delta x_i)^k].$$

To find these moments, we can use the observe moments

$$M'_{k,i,j} = E[(\Delta x_i - \Delta x_j)^k].$$

What is known: case of second moments. For k = 2, we have $M'_{2,i,j} = M_{2,i} + M_{2,j}$, so we can uniquely reconstruct the desired second moments $M_{2,i,j}$ from the observed moments $M'_{2,i,j}$ by using the following formulas:

$$M_{2,1} = \frac{M'_{2,1,2} + M'_{2,1,3} - M'_{2,2,3}}{2}, \quad M_{2,2} = \frac{M'_{2,1,2} + M'_{2,2,3} - M'_{2,1,3}}{2},$$
$$M_{2,3} = \frac{M'_{2,1,3} + M'_{2,2,3} - M'_{2,1,2}}{2}.$$

Natural next case: third moments. Can we similarly reconstruct the desired third moments $M_{3,i} = E[(\Delta x_i)^3]$ based on the observed third moments $M'_{3,i,j} = E[(\Delta x_i - \Delta x_i)^3]$? Here,

$$(\Delta x_i - \Delta x_j)^3 = (\Delta x_i)^3 - 3 \cdot (\Delta x_i)^2 \cdot \Delta x_j + 3 \cdot \Delta x_i \cdot (\Delta x_j)^2 - (\Delta x_j)^3,$$

so, due to linearity of the mean and to the fact that the measurement errors Δx_i and Δx_j corresponding to two measuring instruments are assumed to be independent, we conclude that

$$M'_{3,i,j} = E[(\Delta x_i - \Delta x_j)^3] = E[(\Delta x_i)^3] - 3 \cdot E[(\Delta x_i)^2] \cdot E[\Delta x_j] + 3 \cdot E[\Delta x_i] \cdot E[(\Delta x_j)^2] - E[(\Delta x_j)^3].$$

is case, $E[\Delta x_i] = E[\Delta x_j] = 0$ and $E(\Delta x_i)^3] = M_{3,i}$, so

In th

$$M'_{3,i,j} = M_{3,i} - M_{3,j}$$

Since we only know the differences between the their moments, we cannot uniquely reconstruct these moments $M_{3,i}$: for example, if we add a constant to all the values $M_{3,i}$, all the observed differences will not change.

So, the only case when we can reconstruct the third moments $M_{3,i}$ is when we have some information about them already. One such case is when we know that for one of the measuring instruments, the probability distribution of measurement errors is symmetric. In turns out that in this case, we can reconstruct all the moments and thus, we can uniquely reconstruct all the original probability distributions.

When the probability distribution of one of the measurement errors is symmetric, all distributions can be uniquely determined from the observed differences. Without losing generality, let us assume that the probability distribution of the measurement error is symmetric for the 1st measuring instrument. For a symmetric distribution, all odd moments are equal to 0; thus, we have $M_{2s+1,1} = 0$ for all s. Let us prove, by induction, that we can reconstruct all the moments of all the distributions.

We already know that we can reconstruct the second moments $M_{2,i}$. Let us assume that we already know how to reconstruct the moments $M_{2,i}, \ldots, M_{n,i}$. Let us show how to reconstruct the next moments $M_{n+1,i} = E[(\Delta x_i)^{n+1}]$. For this, we will use the observed moments $M'_{n+1,i,j} = E[(\Delta x_i - \Delta x_j)^{n+1}]$. We will consider two cases:

- when n is odd, i.e., n = 2s + 1 and n + 2 = 2s + 2, and
- when n is even, i.e., n = 2s and n + 1 = 2s + 1.

First case. Let us first consider the first case. Here,

$$(\Delta x_i - \Delta x_j)^{2s+2} = (\Delta x_i)^{2s+2} - (2s+2) \cdot (\Delta x_i)^{2s+1} \cdot \Delta x_j + \frac{(2s+2) \cdot (2s+1)}{1 \cdot 2} \cdot (\Delta x_i)^{2s} \cdot (\Delta x_j)^2 - \dots + \frac{(2s+2) \cdot (2s+1)}{1 \cdot 2} \cdot (\Delta x_i)^2 \cdot (\Delta x_j)^{2s} - (2s+2) \cdot \Delta x_i \cdot (\Delta x_j)^{2s+1} + (\Delta x_j)^{2s+2}.$$

Thus,

$$M'_{2s+2,i,j} = E[(\Delta x_i)^{2s+2}] - (2s+2) \cdot E[(\Delta x_i)^{2s+1}] \cdot E[\Delta x_j] + \frac{(2s+2) \cdot (2s+1)}{1 \cdot 2} \cdot E[(\Delta x_i)^{2s}] \cdot E[(\Delta x_j)^2] - \dots + \frac{(2s+2) \cdot (2s+1)}{1 \cdot 2} \cdot E[(\Delta x_i)^2] \cdot E[(\Delta x_j)^{2s}] - (2s+2) \cdot E[\Delta x_i] \cdot E[(\Delta x_j)^{2s+1}] + E[(\Delta x_j)^{2s+2}],$$

i.e.,

$$M'_{2s+2,i,j} = M_{2s+2,i} + \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_{2s,i} \cdot M_{2,j} + \dots + \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_{2,i} \cdot M_{2s,j} + M_{2s+2,j}.$$

Thus,

$$M_{2s+2,i} + M_{2s+2,j} = s_{ij} \stackrel{\text{def}}{=} M'_{2s+2,i,j} - \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_{2s,i} \cdot M_{2,j} - \dots - \frac{(2s+2)\cdot(2s+1)}{1\cdot 2} \cdot M_{2i} \cdot M_{2s,j}.$$

We know the value $M'_{2s+2,i,j}$, and we assumed that we have already shown that we can uniquely determine the moments $M_{2,i}, \ldots, M_{2s+1,i}$. Thus, we can indeed uniquely determine the values $s_{ij} = M_{2s+2,i} + M_{2s+2,j}$.

Based on these values, we can uniquely reconstruct the moments $M_{n+1,i} = M_{2s+2,i}$ as follows:

$$M_{2s+2,1} = \frac{s_{12} + s_{13} - s_{23}}{2}, \quad M_{2s+2,2} = \frac{s_{12} + s_{23} - s_{13}}{2},$$
$$M_{2s+2,3} = \frac{s_{13} + s_{23} - s_{12}}{2}.$$

Second case. Let us now consider the second case, when n = 2s and n + 1 = 2s + 1. Since we assumed that for the first measuring instrument, the probability distribution is symmetric, we get $M_{2s+1,1} = E[(\Delta x_1)^{2s+1}] = 0$.

For every $i \neq 1$, we have

$$(\Delta x_i - \Delta x_1)^{2s+1} = (\Delta x_i)^{2s+1} - (2s+2) \cdot (\Delta x_i)^{2s} \cdot \Delta x_1 + \frac{(2s+1) \cdot 2s}{1 \cdot 2} \cdot (\Delta x_i)^{2s-1} \cdot (\Delta x_1)^2 - \dots + \frac{(2s+1) \cdot 2s}{1 \cdot 2} \cdot (\Delta x_i)^2 \cdot (\Delta x_1)^{2s-1} - (2s+1) \cdot \Delta x_i \cdot (\Delta x_1)^{2s} + (\Delta x_1)^{2s+1}.$$

Thus,

$$M'_{2s+1,i,1} = E[(\Delta x_i)^{2s+1}] - (2s+2) \cdot E[(\Delta x_i)^{2s+1}] \cdot E[\Delta x_1] + \frac{(2s+1) \cdot 2s}{1 \cdot 2} \cdot E[(\Delta x_i)^{2s-1}] \cdot E[(\Delta x_1)^2] - \ldots + \frac{(2s+1) \cdot 2s}{1 \cdot 2} \cdot E[(\Delta x_i)^2] \cdot E[(\Delta x_1)^{2s-1}] - (2s+1) \cdot E[\Delta x_i] \cdot E[(\Delta x_1)^{2s}] + E[(\Delta x_1)^{2s+1}],$$

i.e.,

$$M'_{2s+1,i,1} = M_{2s+1,i} + \frac{(2s+1)\cdot 2s}{1\cdot 2} \cdot M_{2s-1,i} \cdot M_{2,1} + \dots +$$

Thus,

$$M_{2s+1,i} = M'_{2s+1,i,1} - \frac{(2s+1)\cdot 2s}{1\cdot 2} \cdot M_{2s,i} \cdot M_{2,1} - \dots$$

We know the value $M'_{2s+1,i,1}$, and we assumed that we have already shown that we can uniquely determine the moments $M_{2,i}, \ldots, M_{2s,i}$. Thus, we can indeed uniquely determine the moments $M_{n+1,i} = M_{2s+1,i}$.

Conclusion. In both cases, the induction step is proven, so induction proves that we can indeed determine all the moments of all the distributions.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

- R. N. Bracewell, Fourier Transform and Its Applications, McGraw Hill, New York, 1978.
- [2] Th. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts, 2009.
- [3] P. V. Novitskii and I. A. Zograph, *Estimating the Measurement Errors*, Energoatomizdat, Leningrad, 1991 (in Russian).

- [4] O. Ochoa, A. Velasco, and C. Servin, "Towards Model Fusion in Geophysics: How to Estimate Accuracy of Different Models", *Journal of Uncertain Systems*, 2013, Vol. 7, No. 3, pp. 190–197.
- [5] S. Orfanidis, *Introduction to Signal Processing*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
- [6] A. I. Orlov, "How often are the observations normal?", *Industrial Laboratory*, 1991, Vol. 57. No. 7, pp. 770–772
- [7] S. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, American Institute of Physics, New York, 2005.
- [8] C. Servin and V. Kreinovich, Propagation of Interval and Probabilistic Uncertainty in Cyberinfrastructure-Related Data Processing and Data Fusion, Springer Verlag, Berlin, Heidelberg, 2015.
- [9] D. J. Sheskin, Handbook of Parametric and Nonparametric Statistical Procedures, Chapman and Hall/CRC Press, Boca Raton, Florida, 2011.
- [10] I. N. Sneddon, Fourier Transforms, Dover Publ., New York, 2010.