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

Christian Servin[1] and Vladik Kreinovich[2]

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

1. Need to Determine Accuracies of Measurement Instruments

- Most information comes from measurements.
- Measurement results are never absolutely accurate.
- The measurement result $\widetilde{x}$ is, in general, different from the actual (unknown) value x of the corr. quantity.
- 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 $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$, 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 Δx.

2. How Accuracies Are Usually Determined: by Using a Second, Much More Accurate Measuring Instrument

- A 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 $\Delta x_2 = \widetilde{x}_2 - x$ of this second instrument is much smaller than $\Delta x = \widetilde{x} - x$.
- Thus, the difference $\widetilde{x} - \widetilde{x}_2 = (\widetilde{x} - x) - (\widetilde{x}_2 - x) \approx \Delta x$ can serve as a good approximation to Δx.
- From the sample of such differences, we can therefore find the desired probability distribution for Δx.

3. 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?
- Usually, there are *several* measuring instrument of the type that we want to analyze.
- Due to measurement errors, for same quantity, these instruments produce different measurement results.
- Let's try to extract the measurement accuracy from the differences between these measurement results.

4. Two Possible Situations

- In some cases, we have a stable manufacturing process.
- This process produces several practical identical measuring instruments.
- For these instruments, 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 such cases, for each individual measuring instrument, we need to find its own probability distribution.

5. What is Known: Case of Normal Distribution

- Often, the measurement error is caused by the joint effect of numerous independent small factors.
- In such situations, the Central Limit Theorem 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, we should limit ourselves to the *random component* $\Delta x - E[\Delta x]$ of the measurement error.

6. Case of Normal Distribution (cont-d)

- For this "re-normalized" measurement error Δx, the mean is 0, so, all we need to determine is σ.
- For two identical independent instruments, $\widetilde{x}_2 - \widetilde{x}_1$ is also normal, with variance $V = \sigma^2 + \sigma^2 = 2\sigma^2$.
- Thus, once we experimentally determine the variance V of $\widetilde{x}_2 - \widetilde{x}_1$, we *can* compute $\sigma^2 = \dfrac{V}{2}$.
- When instruments are different, the variance of $\widetilde{x}_i - \widetilde{x}_j$ is equal to $V_{ij} = \sigma_i^2 + \sigma_j^2$.
- Thus, for empirical variances V_{12}, V_{23}, and V_{13}, we have:

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

- So, we *can* compute $\sigma_1^2 = \dfrac{V_{12} + V_{13} - V_{23}}{2}$,

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

7. What If Distributions Are Not Gaussian?

- Empirical analysis shows that only slightly more than a half of instruments have Gaussian measurement errors.
- What happens in the non-Gaussian case?
- It is known that, sometimes, we simply cannot uniquely reconstruct the corresponding distributions.
- In this talk, we explain when such a reconstruction is possible and when it is not possible.

8. Idea: Let Us Use Moments

- 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 M_2, we also need to describe M_3, M_4, ...
- If we know all the moments, then we can uniquely determine the corresponding probability distribution.
- Indeed, the usual way to represent a random variable Δx is by its pdf $\rho(\Delta x)$.

9. Let Us Use Moments (cont-d)

- In many situations, it is convenient to use its *characteristic function*

$$\chi(\omega) \stackrel{\text{def}}{=} E[\exp(\mathrm{i}\cdot\omega\cdot\Delta x)] = \int \rho(\Delta x)\cdot\exp(\mathrm{i}\cdot\omega\cdot\Delta x)\, d(\Delta x).$$

- From the mathematical viewpoint, the characteristic function is the Fourier transform of the pdf.
- It is known that we can uniquely reconstruct a function from its Fourier transform: by inverse transform.
- On the other hand,

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

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

- If we know M_k, then we know $\chi(\omega)$, hence $\rho(x)$.

10. Important Fact: For a Symmetric Distribution, Odd Moments Are Zeros

- For a symmetric distribution, for which $\rho(-\Delta x) = \rho(\Delta x)$, all 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, $M_{2s+1} = 0$.

11. Case When Have Several Identical Measuring Instruments

- In this case, all measuring instruments have the same moments $M_k = E[(\Delta x)^k]$.
- The only available information consists of the differences $\Delta x_1 - \Delta x_2 = \widetilde{x}_1 - \widetilde{x}_2$.
- We can thus determine the moments $M'_k = E[(\Delta x_1 - \Delta x_2)^k]$.
- We would like to find M_k.
- For $k = 2$, we have $M'_2 = 2M_2$, so $M_2 = 0.5 M'_2$.
- For $k = 3$, we get $M'_3 = 0$.
- 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.

12. When the Probability Distribution of the Measurement Error Is Symmetric

- 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}.
- We show, by induction, that we can reconstruct all even moments.
- We already know that we can reconstruct M_2.
- Let us assume that we already know how to reconstruct the moments $M_2, \ldots, M_{2s}$.
- To reconstruct $M_{2s+2} = E[(\Delta x)^{2s+2}]$, let us use

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

13. Case of Symmetric Distribution (cont-d)

- Here,

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

- So,

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

- Thus,

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

- We know the value M'_{2s+2}, and we assumed that we already know $M_2, \ldots, M_{2s}$.
- Thus, we can uniquely determine M_{2s+2}.
- Induction proves that we can indeed determine all the even moments.

14. Case of Different Measuring Instruments

- We observe the moments $M'_{k,i,j} = E[(\Delta x_i - \Delta x_j)^k]$.
- We want to find the moments $M_{k,i} = E[(\Delta x_i)^k]$.
- We already know that in general, the reconstruction is not possible.
- Let us show, by induction, that reconstruction is possible if one of the distributions $i = 1$ is symmetric.
- For $k = 2$, we have $M'_{2,i,j} = M_{2,i} + M_{2,j}$, so we can find $M_{2,i}$; e.g., $M_{2,1} = \dfrac{M'_{2,1,2} + M'_{2,1,3} - M'_{2,2,3}}{2}$.
- Similar formulas help reconstruct even moments.
- $M'_{2s+1,i,1} = M_{2s+1,i} + \frac{(2s+1)\cdot 2s}{1\cdot 2} \cdot M_{2s-1,i} \cdot M_{2,1} + \ldots,$ so

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

15. Acknowledgments

This work was supported in part by the National Science Foundation grants:

- HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence), and
- DUE-0926721.