Accuracy of Data Fusion: Interval (and Fuzzy) Case

Christian Servin[1], Olga Kosheleva[2], and Vladik Kreinovich[2]

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

1. Need for Estimation

- In the grand scheme of things, the main objectives of science and engineering are:
  - to get a good understanding of the current state of the world and how this state will change, and
  - to come up with recipes of how to make sure that this change will go in favorable directions.
- To describe the state of the world, we need to describe the numerical values of all the physical quantities.
- To describe the corresponding recommendations, we need to describe:
  - the numerical values of all the parameters of the corresponding designs and/or
  - the numerical values of the appropriate controls.
- These values usually depend on the parameters describing the current state of the world.

2. Need for Estimation (cont-d)

- We can estimate the corresponding physical quantities directly:
  - either by measuring them
  - or by relying on experts who provide the related estimates.
- If this is not possible, we can estimate the desired quantities indirectly:
  - by measuring and/or estimating related quantities and
  - by using the known relation between all these quantities to estimate of the desired quantities.

3. Need to Improve Accuracy and Need for Data Fusion

- Sometimes, even the measurements by state-of-the-art measuring instruments are not accurate enough.
- For example, we want to estimate density etc. several kilometers below the surface.
- Often, these estimates are not sufficiently accurate:
  - to predict the location of mineral deposits or
  - to predict an earthquake.
- A natural way to improve accuracy is:
  - to perform more estimations of the same quantity, and then
  - to combine ("fuse") the resulting estimates into a single more accurate one.

4. How Accurate is the Result of Data Fusion?

- A natural question is: how accurate is the result of the data fusion?
- The answer to this question is known for probabilistic uncertainty.
- However, for important cases of interval and fuzzy uncertainty, no such general formulas have been known.
- We show how to derive the corresponding formulas.

5. Towards Formulating the Problem in Precise Terms

- Let us assume that we have n estimates $x_1, \ldots, x_n$ of the same quantity x.
- Often, we know the probability distributions of the corresponding approximation errors

$$\Delta x_i \stackrel{\text{def}}{=} x_i - x.$$

- We are interested in values x_i obtained by state-of-the-art measurements.
- Otherwise, there is no big need for fusion, we could simply use more accurate measuring instruments.
- State-of-the-art means, in particular, that all the usual ways to improve accuracy have already been applied.

6. Formulating the Problem (cont-d)

- For example:
  - if a measuring instrument has a bias,
  - i.e., if the mean value of the approximation error is different from 0,
  - then we can detect this bias if we *calibrate* the instrument.
- For that, we repeatedly compare:
  - this instrument's results and
  - the results of measuring the same quantity by a "standard" (more accurate) measuring instrument.
- Once the bias is known, we can simply subtract this bias from all the measurement results.
- Thus, when we deal with state-of-the-art measuring instruments, we can safely assume that the bias is 0.

7. Formulating the Problem (cont-d)

- So, the mean value of the estimation error is 0.
- Similarly, we can safely assume that all known major sources of measurement errors have been taken care of.
- E.g., measurements are affected by the 50–60 Hz electromagnetic signals emitted by electric outlets.
- So we can assume that these signals have been screened away.
- In general, we can assume that all reasonably major sources of measurement errors have been eliminated.
- Thus, the remaining estimate comes from the joint effect of a large number of small error components.
- It is known that in such situations, the resulting error distribution is close to Gaussian.

8. Formulating the Problem (cont-d)

- This follows from the Central Limit Theorem:
  - when the number N of such small components tends to infinity,
  - the distribution of the sum of these components tends to Gaussian.
- So, we can safely assume that each estimation error Δx_i is normally distributed.
- A normal distribution is uniquely determined by its mean and its standard deviation.
- We know that the mean of Δx_i is 0.
- So, knowing the distribution means that we know the standard deviation σ_i of the corresponding estimate.
- It is also reasonable to assume that errors corresponding to different measurements are independent.

9. Resulting Formulation of the Problem

- We have n estimates $x_1, \ldots, x_n$ of the same quantity y.
- For each i, the measurement error $x_i - x$ is normally distributed with 0 mean and known st. dev. σ_i.
- We also know that the measurement errors corresponding to different distributions are independent.
- We would like to find a combined estimate $\widetilde{x}$ and estimate how accurate is this estimate.

10. Fusing Measurement Results: Derivation of the Formula

- Each measurement error $x_i - x$ is normally distributed.
- So, the corresponding probability density function (pdf) $\rho_i(x)$ has the form

$$\rho_i(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma_i} \cdot \exp\left(-\frac{(x_i - x)^2}{2\sigma_i^2}\right).$$

- Measurement errors corresponding to different measurements are independent.
- So, the overall pdf is equal to the product of the corresponding probability densities:

$$\rho(x) = \prod_{i=1}^{n} \rho_i(x) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi} \cdot \sigma_i} \cdot \exp\left(-\frac{(x_i - x)^2}{2\sigma_i^2}\right)\right).$$

11. Fusing Measurement Results (cont-d)

- As the desired fused estimate $\widetilde{x}$, it is reasonable to select:
  - the most probable value x, i.e.,
  - the value for which the above expression attains its largest possible value.
- This idea is known as *Maximum Likelihood Method*.
- This maximization problem can be simplified if we take into account that $f(z) = -\ln(z)$ is decreasing.
- Thus, maximizing the above expression is equivalent to minimizing its negative logarithm

$$-\ln(\rho(x)) = \text{const} + \sum_{i=1}^{n} \frac{(x_i - x)^2}{2\sigma_i^2}.$$

- Here const denotes terms that do not depend on x.

12. Fusing Measurement Results (cont-d)

- Differentiating $-\ln(\rho(x))$ with respect to the unknown x and equating the derivative to 0, we get

$$\sum_{i=1}^{n} \frac{x - x_i}{\sigma_i^2} = 0, \text{ i.e., } x \cdot \sum_{i=1}^{n} \sigma^{-2} = \sum_{i=1}^{n} x_i \cdot \sigma_i^{-2}.$$

- Thus, we arrive at the following formula.

13. Probabilistic Fusion: Resulting Formulas

- *Problem:* how to fuse measurement results $x_1, \ldots, x_n$ measured with accuracies $\sigma_1, \ldots, \sigma_n$.
- *Solution:* $\widetilde{x} = \dfrac{\sum\limits_{i=1}^{n} x_i \cdot \sigma_i^{-2}}{\sum\limits_{i=1}^{n} \sigma_i^{-2}}$.
- How accurate is the fused estimate?
- The probability distribution for different values x is given by the above formula, i.e.:

$$\rho(x) = \left(\frac{1}{\sqrt{2\pi} \cdot \prod\limits_{i=1}^{n} \sigma_i} \cdot \exp\left(-\sum_{i=1}^{n} \frac{(x_i - x)^2}{2\sigma_i^2} \right) \right).$$

- One can easily see that the expression under the exponent is a quadratic function of x.

14. Probabilistic Fusion (cont-d)

- Thus, the distribution for x is also Gaussian, i.e., has the form $\rho(x) = \dfrac{1}{\sqrt{2\pi}\cdot\sigma}\cdot\exp\left(-\dfrac{(\widetilde{x}-x)^2}{2\sigma^2}\right)$.
- Comparing the coefficients for x^2 under the exponential function in two expressions, we conclude that

$$\sum_{i=1}^{n}\frac{1}{\sigma_i^2}=\frac{1}{\sigma^2},\ \text{so } \sigma^2=\frac{1}{\sum\limits_{i=1}^{n}\frac{1}{\sigma_i^2}}.$$

- For $n=2$, $\sigma^2 = \dfrac{1}{\dfrac{1}{\sigma_1^2}+\dfrac{1}{\sigma_2^2}} = \dfrac{\sigma_1^2\cdot\sigma_2^2}{\sigma_1^2+\sigma_2^2}$.
- Another important case is when all the measurements have the same accuracy, i.e., when $\sigma_1 = \ldots = \sigma_n$.
- In this case, we have $\dfrac{1}{\sigma^2}=\dfrac{n}{\sigma_1^2}$, so $\sigma^2=\dfrac{\sigma_1^2}{n}$ and $\sigma=\dfrac{\sigma_1}{\sqrt{n}}$.

15. Need for Interval Uncertainty

- We assumed that we know the probability distribution of measurement errors $\Delta x_i \stackrel{\text{def}}{=} x_i - x$.
- Usually, this distribution is obtained by *calibrating* the measuring instrument.
- We compare the values x_i measured by this instrument and by the standard instrument (SI).
- SI is much more accurate.
- So, we can safely ignore the difference between SI's results and the actual values x.
- However, there are important cases when calibration is not done.

16. Need for Interval Uncertainty (cont-d)

- In state-of-the-art measurements, we use the most accurate measuring instruments.
- In this case, there are simply no more-accurate instruments which can be used for calibration.
- So calibration is not possible.
- At best, we can provide an upper bound Δ_i on the corresponding measurement error $\Delta x_i = x_i - x$:

$$|\Delta x_i| \leq \Delta_i.$$

- In this case:
  - once we know the measurement result x_i,
  - the only information that can conclude about the actual value x is that $x \in [x_i - \Delta_i, x_i + \Delta_i]$.
- Another case is measurements on the shop floor, during the manufacturing process.

17. Need for Interval Uncertainty (cont-d)

- In this case, theoretically, we could calibrate every single sensor, every single measuring instrument.
- However, calibration is expensive – since it involves using complex standard measuring instrument.
- So, in manufacturing, such calibration is often not done; then:
  - the only information that we have about a measuring instrument is
  - the upper bound on its measurement error.
- If we do not even know any such upper bound, then this is not a measuring instrument at all.
- Indeed, then, the actual value can be anything, no matter what the instrument shows.

18. How to Fuse Measurement Results in Case of Interval Uncertainty

- In each measurement, we get the value x_i with accuracy Δ_i.
- So, we conclude that the actual value x belongs to the interval $[x_i - \Delta_i, x_i + \Delta_i]$.
- Thus, x belongs to the intersection of all these n intervals:

$$[\underline{x}, \overline{x}] = \bigcap_{i=1}^{n} [x_i - \Delta_i, x_i + \Delta_i] = \left[\max_i (x_i - \Delta_i), \min_i (x_i + \Delta_i)\right].$$

19. We Need to Estimate Average Accuracy

- Each value x_i is the result of measuring the desired quantity x with accuracy Δ_i.
- Thus, each value x_i can take any value from the interval

$$[x - \Delta_i, x + \Delta_i].$$

- For the same measurement errors of two measurements, we can get different accuracies of the fusion result.
- Let's assume that we fuse the results of two measurements performed with the same accuracy Δ.
- We can have the exact same measurement result in both cases $x_1 = x_2$.
- In this case, the corresponding intervals are the same, and their intersection is the exact same interval

$$[x_1 - \Delta, x_1 + \Delta].$$

20. Need to Estimate Average Accuracy (cont-d)

- Thus, in this case, fusion does not improve the accuracy at all.
- On the other hand, we may have $x_1 = x + \Delta$ and $x_2 = x - \Delta$; then,

$$[x_1-\Delta, x_1+\Delta]\cap[x_2-\Delta, x_2+\Delta] = [x, x+2\Delta]\cap[x-2\Delta, x] = \{x\}.$$

- So, in this case, by fusing two measurement result, we get the exact value of the measured quantity.
- The accuracy of the fused result depends on the actual measurement results.
- So, the only thing that we can estimate is the *average* value of the corresponding estimation error.

21. What Probability Distribution Should We Use?

- For each measurement, all we know about the measurement error is that it is between $-\Delta_i$ and Δ_i.
- There is no reason to believe that some values from this interval $[-\Delta_i, \Delta_i]$ are more probable than others.
- Thus, it is reasonable to conclude that all these values should have the exact same probability.
- So, the measurement error should be uniformly distributed on the corresponding interval.
- This natural conclusion is known as *Laplace Indeterminacy Principle*.

22. What Probability Distribution (cont-d)

- Similarly, if we have n measurements, then:
  - all we know about n corresponding measurement errors $\Delta x_1, \ldots, \Delta x_n$ is that
  - the corresponding vector $(\Delta x_1, \ldots, \Delta x_n)$ is located in the box

$$[-\Delta_1, \Delta_1] \times \ldots \times [-\Delta_n, \Delta_n].$$

- Thus, it is reasonable to conclude that we have a uniform distribution on this box.
- So, all measurement errors are independent random variables.

23. What Is Known And What We Will Do in This Talk

- It is known that:
  - if we fuse several interval estimates with the same accuracy $\Delta_1 = \ldots = \Delta_n$,
  - then the average accuracy of the fused result is, for large n, asymptotically equal to $\dfrac{\Delta_1}{n}$.
- We can see that, in this case, the average measurement error decreases more than in the probabilistic case.
- Indeed, there, the average measurement error decreases as $1/\sqrt{n}$ (much slower).
- In this talk, we consider the general case of possibly different accuracies Δ_i.

24. Case of Two Fused Measurements $n = 2$: Analysis of the Problem

- We fuse two measurement results x_1 and x_2, measured with accuracies Δ_1 and Δ_2.
- Without losing generality, we can take $\Delta_1 \le \Delta_2$.
- Based on the two measurement results, we get the following upper bound u on the actual value x:

$$u = \min(x_1 + \Delta_1, x_2 + \Delta_2).$$

- We are interested in the average (expected) value Δ of the difference $u - x$.
- There is a symmetry with respect to the change of $x \to -x$ that swaps lower ℓ and upper u bounds.
- Thus, we will the exact same average value for the difference $x - \ell$.

25. Case of $n = 2$ (cont-d)

- To compute this average value, let us find the probability distribution of the difference $\Delta x \stackrel{\text{def}}{=} u - x$.
- Here, for both i, we have $x_i = x + \Delta x_i$, thus,

$$u = \min(x + \Delta x_1 + \Delta_1, x + \Delta x_2 + \Delta_2) =$$

$$x + \min(\Delta x_1 + \Delta_1, \Delta x_2 + \Delta_2).$$

- Thus, $\Delta x = u - x = \min(\Delta x_1 + \Delta_1, \Delta x_2 + \Delta_2)$.
- Let us compute, for each real number z, the probability

$$\text{Prob}(z \le \Delta x) = \text{Prob}(z \le \min(\Delta x_1 + \Delta_1, \Delta x_2 + \Delta_2)).$$

- z is smaller than the minimum of the two numbers if and only it is smaller than both of them, so

$$\text{Prob}(z \le \Delta x) = \text{Prob}(z \le \Delta x_1 + \Delta_1 \,\&\, z \le \Delta x_2 + \Delta_2).$$

26. Case of $n = 2$ (cont-d)

- Let us rewrite each of the resulting inequalities so that it will have:
  - the value Δx_i on one side on the corresponding inequality and
  - all other terms on the other side:

$$\text{Prob}(z \le \Delta x) = \text{Prob}(\Delta x_1 \ge z - \Delta_1 \,\&\, \Delta x_2 \ge z - \Delta_2).$$

- We assumed that the measurement errors Δx_1 and Δx_2 are independent.
- So, the probability that both inequalities hold is equal to the product:

$$\text{Prob}(z \le \Delta x) = \text{Prob}(\Delta x_1 \ge z - \Delta_1) \cdot \text{Prob}(\Delta x_2 \ge z - \Delta_2).$$

- Each Δx_i is uniformly distributed on the corresponding interval $[-\Delta_i, \Delta_i]$ of width $2\Delta_i$.

27. Case of $n = 2$ (cont-d)

- So, the probability to be on each subinterval is proportional to the width of this subinterval.
- To be precise, it is equal to the ratio of the width of the subinterval to the width of the original interval.
- For each threshold z_i, the inequality $\Delta x_i \ge z_i$ is satisfied on the subinterval $[z_i, \Delta_i]$ of width $\Delta_i - z_i$.
- Thus, the probability that this inequality is satisfied is equal to the ratio $\dfrac{\Delta_i - z_i}{2\Delta_i}$.
- In particular, for $z_i = z - \Delta_i$, we get

$$\text{Prob}(\Delta x_i \ge z - \Delta_i) = \frac{\Delta_i - (z - \Delta_i)}{2\Delta_i} = \frac{2\Delta_i - z}{2\Delta_i}.$$

- Thus, we have

$$\text{Prob}(z \le \Delta x) = \frac{2\Delta_1 - z}{2\Delta_1} \cdot \frac{2\Delta_2 - z}{2\Delta_2} = \frac{(2\Delta_1 - z) \cdot (2\Delta_2 - z)}{4\Delta_1 \cdot \Delta_2}.$$

28. Case of $n = 2$ (cont-d)

- So, the cumulative distribution function (cdf) $F(z) = \text{Prob}(\Delta x \le z)$ is equal to

$$F(z) = 1 - \text{Prob}(z \le \Delta x) = 1 - \frac{(2\Delta_1 - z) \cdot (2\Delta_2 - z)}{4\Delta_1 \cdot \Delta_2}.$$

- The corresponding probability density function $\rho(z)$ can be obtained if we differentiate the cdf:

$$\rho(z) = \frac{dF(z)}{dz} = \frac{(2\Delta_1 - z) + (2\Delta_1 - z)}{4\Delta_1 \cdot \Delta_2} =$$

$$\frac{2\Delta_1 + 2\Delta_2 - 2z}{4\Delta_1 \cdot \Delta_2} = \frac{\Delta_1 + \Delta_2 - z}{2\Delta_1 \cdot \Delta_2}.$$

- The difference $\Delta x = u - x$ is always greater than or equal to 0 – since u is the upper bound for x.

- The largest possible value of this difference is $2\Delta_1$, when $x_1 = x - \Delta_1$, $x_2 = x$, and $u = x_2 + \Delta_1$.

29. Case of $n = 2$ (cont-d)

- The average (expected) value Δ of this difference can thus be computed as

$$\Delta = \int_0^{2\Delta_1} z \cdot \rho(z)\,dz =$$

$$\frac{1}{2\Delta_1 \cdot \Delta_2} \cdot \int_0^{2\Delta_1} [z \cdot (\Delta_1 + \Delta_2) - z^2]\,dz =$$

$$\frac{1}{2\Delta_1 \cdot \Delta_2} \cdot \left[\frac{z^2}{2} \cdot (\Delta_1 + \Delta_2) - \frac{z^3}{3}\right]_0^{2\Delta_1} =$$

$$\frac{1}{2\Delta_1 \cdot \Delta_2} \cdot \left[\frac{4\Delta_1^2}{2} \cdot (\Delta_1 + \Delta_2) - \frac{8\Delta_1^3}{3}\right] =$$

$$\frac{1}{2\Delta_1 \cdot \Delta_2} \cdot \left(2\Delta_1^3 + 2\Delta_1^2 \cdot \Delta_2 - \frac{8}{3} \cdot \Delta_1^3\right) =$$

$$\frac{2\Delta_1^2 \cdot \Delta_2 - \dfrac{2}{3} \cdot \Delta_1^3}{2\Delta_1 \cdot \Delta_2} = \Delta_1 - \frac{1}{3} \cdot \frac{\Delta_1^2}{\Delta_2}.$$

30. Case of $n = 2$: Conclusions

- We fuse two measurements with interval uncertainties

$$\Delta_1 \le \Delta_2.$$

- Then, the average accuracy Δ of the fused result is

$$\Delta = \Delta_1 - \frac{1}{3} \cdot \frac{\Delta_1^2}{\Delta_2}.$$

- In particular, when $\Delta_2 \to \infty$, we get $\Delta \to \Delta_1$.
- This makes perfect sense:
  - very inaccurate measurements do not add any information,
  - so accuracy is not improved.

31. Case of $n = 2$: Conclusions (cont-d)

- When $\Delta_1 = \Delta_2$, we get $\Delta = \dfrac{2}{3} \cdot \Delta_1$.
- In other words, the average inaccuracy decreases by a factor of 1.5.
- In the probabilistic case, it only decreases by a smaller factor of $\sqrt{2} \approx 1.41$.
- Let us show that for $n = 2$, interval fusion always leads to a larger decrease of measurement error.

32. For $n = 2$, Interval Fusion Is, on Average, More Accurate

- Let us start with the same accuracy values $\Delta_1 = \sigma_1$ and $\Delta_2 = \sigma_2$.
- We want to prove that the result of interval fusion is always narrower, i.e., that

$$\Delta_1 - \frac{1}{3} \cdot \frac{\Delta_1^2}{\Delta_2} < \sqrt{\frac{\Delta_1^2 \cdot \Delta_2^2}{\Delta_1^2 + \Delta_2^2}} = \frac{\Delta_1 \cdot \Delta_2}{\sqrt{\Delta_1^2 + \Delta_2^2}}.$$

- Let us divide both sides of this inequality by Δ_1 and express both sides in terms of the ratio $r \stackrel{\text{def}}{=} \frac{\Delta_2}{\Delta_1} \geq 1$.
- Then the desired inequality gets the following equivalent form

$$1 - \frac{1}{3r} < \frac{r}{\sqrt{1 + r^2}}.$$

33. Interval Fusion Is More Accurate (cont-d)

- Multiplying both sides by r and by $\sqrt{1+r^2}$, we get yet another equivalent inequality
$$\left(r - \frac{1}{3}\right) \cdot \sqrt{1+r^2} < r^2.$$
- Squaring both sides, we get the following equivalent inequality
$$(r^2+1) \cdot \left(k^2 - \frac{2}{3}r + \frac{1}{9}\right) < r^4.$$
- Opening parentheses, we get
$$r^4 - \frac{2}{3}r^3 + \frac{1}{9}r^2 + r^2 - \frac{2}{3}r + \frac{1}{9} < r^4.$$
- Let us move all the terms to the right-hand side and combining terms proportional to the same power of r:
$$\frac{2}{3}r^3 - \frac{10}{9}r^2 + \frac{2}{3}r - \frac{1}{9} > 0.$$

34. Interval Fusion Is More Accurate (cont-d)

- Let us prove that this inequality holds for all r.
- Indeed, this inequality clearly holds for $r = 1$: then, the left-hand side is equal to

$$\frac{2}{3} - \frac{10}{9} + \frac{2}{3} - \frac{1}{9} = \frac{6 - 10 + 6 - 1}{9} = \frac{1}{9} > 0.$$

- Let us prove that the left-hand side is increasing and thus, it is positive for all $r > 1$ as well.
- Indeed, the derivative of this left-hand side is equal to

$$2r^2 - \frac{20}{9}r + \frac{2}{3}.$$

- The discriminant of this quadratic equation is equal to

$$\left(\frac{20}{9}\right)^2 - 4 \cdot 2 \cdot \frac{2}{3} = \frac{400}{81} - \frac{16}{3} = \frac{400 - 16 \cdot 27}{81} = \frac{400 - 432}{81} = -\frac{32}{81} < 0.$$

35. Interval Fusion Is More Accurate (cont-d)

- So, the quadratic expression for the derivative is always non-negative.
- Thus, the inequality holds for all $r \geq 1$.
- Since it is equivalent to the desired inequality, the desired inequality also holds always.
- The statement is proven.

36. Case When All the Measurements Are Equally Accurate $\Delta_1 = \ldots = \Delta_n$

- In this case, similar to the case $n = 2$, we conclude that
$$F(z) = 1 - \frac{2\Delta_1 - z}{2\Delta_1} \cdot \ldots \cdot \frac{2\Delta_n - z}{2\Delta_n} = 1 - \left(\frac{2\Delta_1 - z}{2\Delta_1}\right)^n.$$
- Here, z can take any value from 0 to $2\Delta_1$, so the ratio $y \stackrel{\text{def}}{=} \dfrac{z}{2\Delta_1}$ takes values from the interval $[0, 1]$.
- In terms of y, we have $F(y) = 1 - (1 - y)^n$, so
$$\rho(y) = \frac{dF(y)}{dy} = n \cdot (1 - y)^{n-1}.$$
- Thus, the average value $E[y]$ of y is equal to
$$E[y] = \int_0^1 n \cdot (1 - y)^{n-1} \cdot y\, dy.$$

37. Case When $\Delta_1 = \ldots = \Delta_n$ (cont-d)

- Introducing an auxiliary variable $u = 1 - y$ for which $y = 1 - u$ and for which $u \in [0, 1]$, we get

$$E[y] = n \cdot \int_0^1 (1-u) \cdot u^{n-1}\, du = n \cdot \int_0^1 (u^{n-1} - u^n)\, du =$$

$$n \cdot \left[\frac{u^n}{n} - \frac{u^{n+1}}{n+1}\right]_0^1 = n \cdot \left(\frac{1}{n} - \frac{1}{n+1}\right) =$$

$$n \cdot \frac{(n+1) - n}{n \cdot (n+1)} = n \cdot \frac{1}{n \cdot (n+1)} = \frac{1}{n+1}.$$

- Thus, for the accuracy $\Delta = r \cdot (2\Delta_1)$ of the fusion result, we get $\Delta = \dfrac{2}{n+1} \cdot \Delta_1$.

- For the case of $n = 2$, this is exactly what we got based on our general formula.

38. General Case

- In the general case, one can actually also get an explicit formula for each n, with $\Delta_1 \le \ldots \le \Delta_n$.
- Indeed, here, we get the following cumulative distribution function:

$$F(z) = 1 - \frac{2\Delta_1 - z}{2\Delta_1} \cdot \ldots \cdot \frac{2\Delta_n - z}{2\Delta_n}.$$

- The right-hand side of this formula is a polynomial.
- Thus, by differentiation, we can get an explicit polynomial formula for the derivative $\rho(z) = \dfrac{dF(z)}{dz}$.
- Thus, we get an explicit polynomial formula for the resulting value

$$\Delta = \int_0^{2\Delta_1} \rho(z) \cdot z\, dz.$$

39. Case of Fuzzy Estimates

- It is known that in fuzzy logic:
  - the usual way of processing fuzzy estimates – by using Zadeh's extension principle
  - is equivalent to processing α-cut intervals for all $\alpha \in [0, 1]$.
- Thus, the above formulas can be applied to the fuzzy case as well.

40. Acknowledgments

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

- 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science)
- and HRD-1242122 (Cyber-ShARE Center of Excellence).