Relationship Between Measurement Results and Expert Estimates of Cumulative Quantities, on the Example of Pavement Roughness

Edgar Daniel Rodriguez Velasquez[1,2],
Carlos M. Chang Albitres[2], and Vladik Kreinovich[3]
[1]Universidad de Piura in Peru (UDEP), edgar.rodriguez@udep.pe
Departments of [2]Civil Engineering and [3]Computer Science
University of Texas at El Paso, El Paso, Texas 79968, USA,
edrodriguezvelasquez@miners.utep.edu, cchangalbitres2@utep.edu,
vladik@utep.edu

1. Cumulative Quantities

- Many physical quantities can be measured directly: e.g., we can directly measure mass, acceleration, force.
- However, we are often interested in *cumulative* quantities that combine values corresponding to:
  - different moments of time and/or
  - different locations.
- For example:
  - when we are studying public health or pollution or economic characteristics,
  - we are often interested in characteristics describing the whole city, the whole region, the whole country.

2. Formulation of the Problem

- Cumulative characteristics are not easy to measure.
- To measure each such characteristic, we need:
  - to perform a large number of measurements, and then
  - to use an appropriate algorithm to combine these results into a single value.
- Such measurements are complicated.
- So, we often have to supplement the measurement results with expert estimates.
- To process such data, it is desirable to describe both estimates in the same scale:
  - to estimate the actual value of the corresponding quantity based on the expert estimate, and
  - vice versa.

3. Case Study: Estimating Pavement Roughness

- Estimating road roughness is an important problem.
- Indeed, road pavements need to be maintained and repaired.
- Both maintenance and repair are expensive.
- So, it is desirable to estimate the pavement roughness as accurately as possible.
- If we overestimate the road roughness, we will waste money on "repairing" an already good road.
- If we underestimate the road roughness, the road segment will be left unrepaired and deteriorate further.
- As a result, the cost of future repair will skyrocket.
- The standard way to measure the pavement roughness is to use the International Roughness Index (IRI).

4. Estimating Pavement Roughness (cont-d)

- Crudely speaking, IRI describes the effect of the pavement roughness on a standardized model of a vehicle.
- Measuring IRI is not easy, because the real vehicles differ from this standardized model.
- As a result, we measure roughness by some instruments and use these measurements to estimate IRI.
- For example, we can:
  - perform measurements by driving an available vehicle along this road segment,
  - extract the local roughness characteristics from the effect of the pavement on this vehicle, and then
  - estimate the effect of the same pavement on the standardized vehicle.

5. Estimating Pavement Roughness (cont-d)

- In view of this difficulty, in many cases, practitioners rely on expert estimates of the pavement roughness.
- The corr. measure – estimated on a scale from 0 to 5 – is known as the Present Serviceability Rating (PSR).

6. Empirical Relation Between Measurement Results and Expert Estimates

- The empirical relation between PSR and IRI is described by the 1994 Al-Omari-Darter formula:

$$\mathrm{PSR} = 5 \cdot \exp(-0.0041 \cdot \mathrm{IRI}).$$

- This formula remains actively used in pavement engineering.
- It works much better than many previously proposed alternative formulas, such as

$$\mathrm{PSR} = a + b \cdot \sqrt{\mathrm{IRI}}.$$

- However, it is not clear why namely this formula works so well.

7. What We Do in This Talk

- We propose a possible explanation for the above empirical formula.
- This explanation will be general: it will apply to all possible cases of cumulative quantities.
- We will come up with a general formula $y = f(x)$ that describes how:
  - a subjective estimate y of a cumulative quantity
  - depends on the result x of its measurement.
- As a case study, we will use gauging road roughness.

8. Main Idea

- In general, the numerical value of a *subjective estimate* depends on the scale.
- In road roughness estimates, we usually use a 0-to-5 scale.
- In other applications, it may be more customary to use 0-to-10 or 0-to-1 scales.
- A usual way to transform between the two scales is to multiply all the values by a corresponding factor.
- For example, to transform from 0-to-10 to 0-to-1 scale, we multiply all the values by $\lambda = 0.1$.
- In other transitions, we can use transformations $y \to \lambda \cdot y$ with different re-scaling factors λ.
- There is no major advantage in selecting a specific scale.

9. Main Idea (cont-d)

- So, subjective estimates are defined modulo such a re-scaling transformation $y \to \lambda \cdot y$.
- At first glance, the result of *measuring* a cumulative quantity may look uniquely determined.
- However, a detailed analysis shows that there is some non-uniqueness here as well.
- Indeed, the result of a cumulative measurement comes from combining values measured:
  - at different moments of time and/or
  - values corresponding to different spatial locations.
- For each individual measurement, the probability of a sensor's malfunction may be low.
- However, often, we perform a large number of measurements.

10. Main Idea (cont-d)

- So, some of them bound to be caused by such malfunctions and are, thus, outliers.
- It is well known that even a single outlier can drastically change the average.
- So, to avoid such influence, the usual algorithms first filter out possible outliers.
- This filtering is not an exact science; we can set up:
  - slightly different thresholds for detecting an outlier,
  - slightly different threshold for allowed number of remaining outliers, etc.
- We may get a computation result that only takes actual signals into account.
- With a different setting, we may get a different result, affected by a few outliers.

11. Main Idea (cont-d)

- Let's denote the average value of an outlier is L and the average number of such outliers is n.
- Then, the second scheme, in effect, adds a constant $n \cdot L$ to the cumulative value computed by the first scheme.
- So, the measured value of a cumulative quantity is defined modulo an addition of some value:

$$x \to x + a \text{ for some constant } a.$$

12. Motivation for Invariance

- We do not know exactly what is the ideal threshold, so we have no reason to select a specific shift as ideal.
- It is therefore reasonable to require:
  - that the desired formula $y = f(x)$ not depend on the choice of such a shift, i.e.,
  - that the corresponding dependence not change if we simply replace x with $x' = x + a$.
- Of course, we cannot just require that $f(x) = f(x + a)$ for all x and all a.
- Indeed, in this case, the function $f(x)$ will simply be a constant, but y increases with x.
- But this is clearly not how invariance is usually defined.
- For example, for many physical interactions, there is no fixed unit of time.

13. Motivation for Invariance (cont-d)

- So, formulas should not change if we simply change a unit for measuring time: $t' = \lambda \cdot t$.
- The formula $d = v \cdot t$ relating the distance d, the velocity v, and the time t should not change.
- We want to make this formula true when time is measured in the new units.
- So, we may need to also appropriately change the units of other related quantities.
- In the above example, we need to appropriately change the unit for measuring velocity, so that:
  - not only time units are changed, e.g., from hours to second, but
  - velocities are also changed from km/hour to km/sec.

14. Motivation for Invariance (cont-d)

- So, if we re-scale x, the formula $y = f(x)$ should remain valid if we appropriately re-scale y.
- As we have mentioned earlier, possible re-scalings of the subjective estimate y have the form $y \to y' = \lambda \cdot y$.
- Thus, for each a, there exists $\lambda(a)$ (depending on a) for which $y = f(x)$ implies that $y' = f(x')$, where

$$x' \stackrel{\text{def}}{=} x + a \text{ and } y' \stackrel{\text{def}}{=} \lambda \cdot y.$$

15. Definitions and the Main Result

- A monotonic function $f(x)$ is called *unit-invariant* if:
  - for every real number a, there exists a positive real number $\lambda(a)$ for which, for each x and y,
  - if $y = f(x)$, then $y' = f(x')$, where $x' \stackrel{\text{def}}{=} x + a$ and $y' \stackrel{def}{=} \lambda(a) \cdot y$.
- **Proposition.** *A function $f(x)$ is unit-invariant if and only if it has the form*

$$f(x) = C \cdot \exp(-b \cdot x) \textit{ for some } C \textit{ and } b.$$

- For road roughness, this result explains the empirical formula.

16. Proof

- It is easy to check that every function $y = f(x) = C \cdot \exp(-b \cdot x)$ is indeed unit-invariant.
- Indeed, for each a, we have

$$f(x') = f(x + a) = C \cdot \exp(-b \cdot (x + a)) =$$

$$C \cdot \exp(-b \cdot x - b \cdot a) = \lambda(a) \cdot C \cdot \exp(-b \cdot x).$$

- Here we denoted $\lambda(a) \stackrel{\text{def}}{=} \exp(-b \cdot a)$.
- Thus here, indeed, $y = f(x)$ implies that $y' = f(x')$.

17. Proof (cont-d)

- Vice versa, let us assume that the function $f(x)$ is unit-invariant.
- Then, for each a, the condition $y = f(x)$ implies that $y' = f(x')$, i.e., that $\lambda(a) \cdot y = f(x+a)$.
- Substituting $y = f(x)$ into this equality, we conclude that $f(x+a) = \lambda(a) \cdot f(x)$.
- It is known that every monotonic solution of this functional equation has the form

$$f(x) = C \cdot \exp(-b \cdot x) \text{ for some } C \text{ and } b.$$

- The proposition is proven.

18. Conclusions

- In pavement engineering, it is important to accurately gauge the quality of road segments.
- Such estimates help us decide how to best distribute the available resources between different road segments.
- So, proper and timely maintenance is performed on road segments whose quality has deteriorated.
- Thus, to avoid future costly repairs of untreated road segments.
- The standard way to gauge the quality of a road segment is International Roughness Index (IRI).
- It requires a large amount of costly measurements.
- As a result, it is not practically possible to regularly measure IRI of all road segments.

19. Conclusions (cont-d)

- So, IRI measurements are usually restricted to major roads.
- For local roads, we need to an indirect way to estimate their quality.
- To estimate the quality of a road segment, we:
  - combine user estimates of different segment properties
  - into a single index known as Present Serviceability Rating (PSR).
- There is an empirical formula relating IRI and PSR.
- However, one of the limitations of this formula is that it purely heuristic.
- This formula lacks a theoretical explanation and thus, the practitioners may be not fully trusting its results.

20. Conclusions (cont-d)

- In this paper, we provide such a theoretical explanation.
- We hope that the resulting increased trust in this formula will help enhance its use.
- Thus, it will help with roads management.

21. 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).