How to Deal with High-Impact Low-Probability Events: Theoretical Explanation of the Empirically Successful Fuzzy-Like Technique

Juan Ulloa, Aaron Velasco,
Olga Kosheleva, Vladik Kreinovich,
University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA
julloa@utep.edu, aavelasco@utep.edu
olgak@utep.edu, vladik@utep.edu

1. Need to deal with high-impact low-probability events

- In many decision-making situations, we need to take into account high-impact low-probability events.
- In civil engineering, we need to take into account the possibility of rare strong earthquakes that could destroy the designed buildings.
- In information security, we need to take into account the low-probability scenario in which:
  - the adversary can break through all our security barriers and
  - thus, inflict a high-impact damage.
- How can we take such events into account?

2. Cannot we use the usual risk-based approach?

- At first glance, the solution is straightforward.
- According to decision theory, in decision making, we should select the alternative for which the expected utility is the largest.
- This means that the expected loss is the smallest.
- The expected loss is equal to the product $p \cdot \ell$ of:
  - the event's probability p and
  - the corresponding loss ℓ.
- So, this product should be the numerical measure that describes how we should take such events into account.
- This measure can be described in purely probabilistic terms.

3. Cannot we use the usual risk-based approach (cont-d)

- For example, the earthquake's damage ℓ to a city can be described by multiplying:
  - the probability P that in this event, a randomly selected building will be damaged, and
  - the average amount of damage D to an affected building.
- For $\ell = P \cdot D$, the expected loss $p \cdot \ell$ takes the form $p \cdot P \cdot D$.
- It is, thus, proportional to the product $p \cdot P$ of the two probabilities:
  - the probability p that such an event will occur, and
  - the probability P that this event will damage a randomly selected building.

4. The usual risk-based approach underestimates the risk

- We want to estimate the probability that the event occurred *and* that it damaged the randomly selected building.
- The above product formula $p \cdot P$ is valid if these two events are independent.
- However, for high-impact low-probability events, there is often a correlation between these two events – which makes the product formula not valid.
- Let us explain this on the example of earthquakes.
- In California or Japan, where reasonable-size earthquakes are frequent, everything is designed with this in mind.
- So such earthquakes do not cause any major damage.
- In contrast, in place like El Paso – where we live – earthquakes are very rare.

5. The usual risk-based approach underestimates the risk (cont-d)

- As a result, many buildings are not designed with such earthquakes in mind. So:
  - if a similar-strength earthquake happens in El Paso – and it will happen sometimes in the next few hundred years,
  - it will cause a huge damage.
- Statistical estimates for P mainly take into account most frequent events – i.e., mostly events from high-frequency zones like California.
- So, if we use these largely-California-based estimates to estimate El Paso risks, we will be strongly underestimating the risk.

6. So what can we do: empirically successful way to take such events into account

- It is desirable to perform statistical analysis of high-impact low-probability events.
- Such an analysis was performed by researchers from the US National Institute of Standards and Technology (NIST).
- The results of their analysis are summarized in the 2012 NIST document.
- Here is the main table from this document.
- In this table, VL means very low, L means low, M means moderate, H mean high, and VH means very high.
- For now, ignore the underlining – it is not from the original table, it was done by us, and it will be explained later.

7. Empirically successful way to take such events into account (cont-d)

p \ P	VL	L	M	H	VH
VL	L	L	<u>L</u>	<u>L</u>	<u>L</u>
L	VL	L	L	L	<u>M</u>
M	VL	L	M	M	<u>H</u>
H	VL	L	M	H	<u>VH</u>
VH	VL	L	M	H	VH

- Almost all the entries fit the fuzzy-like formula $\min(p, P)$.
- There are exception of several entries from the first row and from the last column – entries that we underlined.

8. But why?

- But why this empirical table has this particular form?
- A naive answer is that in this case, naive fuzzy – with minimum – works better than naive probability – with the product.
- In other words, paraphrasing Orwell's "Animal Farm": fuzzy good, probability bad.
- But why is probability bad for low-frequency events – while it works perfectly well for the cases when the frequency is not low?

9. What we do in this talk

- In this talk , we provide an explanation for the above empirical table.
- First, we explain, in detail, the non-underlines part of the table.
- Then, we provide qualitative arguments explaining why underlined entries in this table are different from minimum.

10. What do we know about the desired probability of both events happening?

- The expected loss is equal to the damage D multiplied by the probability t that both events occur:
  - that the low-probability event happens, and
  - that this event causes a randomly selected building to be destroyed.
- All we know is the probabilities p and P of these two events.
- We know that there is a correlation between them, but we do not know the values of this correlation.
- In this case, all we know about the probability t of both events happening is that this value must satisfy the following inequalities:

$$\max(p + P - 1, 0) \leq t \leq \min(p, P).$$

- These inequalities were first derived by Frechét.

11. What do we know about the desired probability of both events happening (cont-d)

- We consider low-probability events, i.e., events for which $p \ll 1$.
- So, unless $P \approx 1$ – which is the case of the last column of or table – we have $p + P \le 1$ and thus, $\max(p + P - 1, 0) = 0$.
- In this case, the above double inequality takes the following form:

$$0 \le t \le \min(p, P).$$

12. What do we know about the expected loss and the expected utility?

- Because of the bounds on the probability t, the expected loss $t \cdot D$ satisfies the following inequality:

$$0 \le t \cdot D \le \min(p, P) \cdot D.$$

- So, for the expected utility u – which is equal to minus the expected loss – we have the following inequality:

$$-\min(p, P) \le u \le 0.$$

- In other words, all we know about the expected utility u is that it is locates somewhere on an interval $[\underline{u}, \overline{u}]$, where

$$\underline{u} \stackrel{\text{def}}{=} -\min(p, P) \cdot D \text{ and } \overline{u} \stackrel{\text{def}}{=} 0.$$

13. How should be make a decision under this interval uncertainty?

- What does decision theory recommend in situations when we only know bounds on the expected utility?
- It recommends selecting an alternative with the largest possible value of the following combination

$$\alpha_H \cdot \overline{u} + (1 - \alpha_H) \cdot \underline{u}, \text{ for some } \alpha_H \in [0, 1].$$

- This expression was first derived by the economist Leo Hurwicz – who later got Nobel prize for his research.
- The coefficient α_H is known as the *optimism-pessimism parameter*. The name comes from the following.
- For $\alpha_H = 1$, the above expression turns into $\overline{u}$.
- This means that the decision maker only takes into account the best-case scenario and ignores all other possibilities.
- This is the case of extreme optimism.

14. How should be make a decision under this interval uncertainty (cont-d)

- For $\alpha_H = 0$, the above expression turns into $\underline{u}$.
- This means that the decision maker only takes into account the worse-case scenario and ignores all other possibilities.
- This is the case of extreme pessimism.
- Intermediate values α_H mean that the decision maker takes different possible scenarios into account.
- In particular, in our case, the Hurwicz's combination takes the following form:

$$\alpha_H \cdot 0 + (1 - \alpha_H) \cdot (-\min(p, P) \cdot D) = -\min(p, P) \cdot (1 - \alpha_H) \cdot D.$$

- So, the risk is proportional to the minimum $\min(p, P)$.
- This is exactly what most entries in the above table say.

15. Remaining questions: why some entries differ from min?

- To complete our explanations, it is necessary to explain why:
  - in two cases: in the first row and in the last column.
  - some entries differ from $\min(p, P)$.
- Let us explain these two cases one by one.

16. Why some entries in the first row are different from min?

- Humans have a tendency to ignore low-probability events when making decisions.
- For example, in many papers, events with probability less than 5% were considered to be impossible.
- This led to so many irreproducible results that the American Statistical Association (ASA) had to issue a special statement about it.
- In spite of this highly publicized statement, many practitioners continue to ignore low-probability events when making decisions.
- Because of this phenomenon, there is a risk that a low-probability event will be ignored.

17. Why some entries in the first row are different from min (cont-d)

- To make sure that the event is *not* ignored, NIST researchers recommend to increase the probability t when p is very low.
- This affects the first row of the above table – the row corresponding to events with very low (VL) probability.

18. Why some entries in the last column are different from min?

- As mentioned in Kahneman's books, humans have a tendency to underestimate high probabilities when making decisions.
- We need to counteract this subjective underestimation.
- So, NIST researchers proposed to increase recommended values t when the probability of damage is very high (VH).
- This corresponds to the last column.

19. Acknowledgments

This work was supported in part:

- by the US National Science Foundation grants:
  - 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science),
  - HRD-1834620 and HRD-2034030 (CAHSI Includes),
  - EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES),
- by the AT&T Fellowship in Information Technology, and
- by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Focus Program SPP 100+ 2388, Grant Nr. 501624329,