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

Juan Ulloa[1], Aaron Velasco[2], Olga Kosheleva[3], and Vladik Kreinovich[1]
[1]Department of Computer Science
[2]Department of Earth, Environment, and Resource Sciences
[3]Department of Teacher Education
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
julloa7@utep.edu, aavelasco@utep.edu, olgak@utep.edu, vladik@utep.edu

Abstract

When making decisions, it is important to take into account high-impact low-probability events. For such events, traditional probability-based approach – which considers the product of the probability p that this event happens and the probability P that a randomly selected building will be destroyed – often underestimates risks. Available data has lead to an empirical table that provides a more adequate risk estimate. Most of the entries in this table correspond to the fuzzy-like formula $\min(p, P)$. This paper explains this empirical result. Specifically, it explains both the effectiveness of the min formula – and also explains deviations from this formula.

1 Formulation of the Problem

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. For example:

- 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, etc.

How can we take such events into account?

Cannot we use the usual risk-based approach? At first glance, the solution is straightforward. According to decision theory (see, e.g., [2, 3, 7, 10, 12, 13, 17]), in decision making, we should select the alternative for which the expected utility is the largest – i.e., in this case, 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. 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 $p \cdot P \cdot D$, and 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.

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. 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 line California. So, if we use these largely-California-based estimates o estimate El Paso risks, we will be strongly underestimating the risk.

So what can we do: empirically successful way to take such events into account. Statistical analysis of numerous high-impact low-probability events was performed by researchers from the US National Institute of Standards and

Technology (NIST). The results of their analysis are summarized in the NIST document [15]. 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:

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

With the exception of several entries from the first row and from last column – entries that we underlines – all the entries fit the fuzzy-like formula $\min(p, P)$ (see, e.g., [1, 6, 11, 14, 16, 20]) – as opposed to the above-mentioned probability-like product formula.

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?

What we do in this paper. In this paper, 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.

2 Our Explanation

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 – first derived by Frechet (see, e.g., [18]):

$$\max(p + P - 1, 0) \le t \le \min(p, P). \tag{1}$$

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 double inequality (1) takes the following form:

$$0 \le t \le \min(p, P). \tag{2}$$

What do we know about the expected loss and the expected utility? Because of the bounds (2) on the probability t, the expected loss $t \cdot D$ satisfies the following inequality:

$$0 \le t \cdot D \le \min(p, P) \cdot D. \tag{3}$$

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

$$-\min(p, P) \le u \le 0. \tag{4}$$

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

$$\underline{u} \stackrel{\text{def}}{=} -\min(p, P) \cdot D \text{ and } \overline{u} \stackrel{\text{def}}{=} 0. \tag{5}$$

How should be make a decision under this interval uncertainty? In situations when we only know the expected utility, decision theory recommends selected an alternative with the largest possible value of the following combination

$$\alpha_H \cdot \overline{u} + (1 - \alpha_H) \cdot \underline{u}, \tag{6}$$

for some coefficient $\alpha_H \in [0, 1]$; see, e.g., [4, 7, 10]. 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 expression (6) 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.
- For $\alpha_H = 0$, the expression (6) 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, for the case when the interval is described by the formula (5), the Hurwicz's combination (6) takes the following form:

$$\alpha_H \cdot 0 + (1 - \alpha_H) \cdot (-\min(p, P) \cdot D) = -\min(p, P) \cdot (1 - \alpha_H) \cdot D. \quad (7)$$

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

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.

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 – which led to so many irreproducible results that the American Statistical Association (ASA) had to issue a special statement about it [19]. 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. To make sure that the event is *not* ignored, NIST researchers recommend to increase the probability t for such cases, when p is very low. This affects the first row of the above table – the row corresponding to events with very low (VL) probability.

Why some entries in the last column are different from min? As mentioned in [5, 8, 9], humans have a tendency to underestimate high probabilities when making decisions.

To counteract this subjective undersatimation, NIST researchers proposed to increase recommended values t for the case when the probability of damage is very high (VH) – which corresponds to the last column.

Acknowledgments

This work was supported in part by the 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), and by the AT&T Fellowship in Information Technology.

It was also supported by a grant from the Hungarian National Research, Development and Innovation Office (NRDI), and by the Institute for Risk and Reliability, Leibniz Universitaet Hannover, Germany.

The authors are greatly thankful to Oscar Perez for his help and encouragement.

References

[1] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[2] P. C. Fishburn, *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.

[3] P. C. Fishburn, *Nonlinear Preference and Utility Theory*, The John Hopkins Press, Baltimore, Maryland, 1988.

[4] L. Hurwicz, *Optimality Criteria for Decision Making Under Ignorance*, Cowles Commission Discussion Paper, Statistics, No. 370, 1951.

[5] D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus, and Giroux, New York, 2011.

[6] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[7] V. Kreinovich, "Decision making under interval uncertainty (and beyond)", In: P. Guo and W. Pedrycz (eds.), *Human-Centric Decision-Making Models for Social Sciences*, Springer Verlag, 2014, pp. 163–193.

[8] J. Lorkowski and V. Kreinovich, "Fuzzy logic ideas can help in explaining Kahneman and Tversky's empirical decision weights", In: L. Zadeh et al. (Eds.), *Recent Developments and New Direction in Soft-Computing Foundations and Applications*, Springer Verlag, 2016, pp. 89–98.

[9] J. Lorkowski and V. Kreinovich, *Bounded Rationality in Decision Making Under Uncertainty: Towards Optimal Granularity*, Springer Verlag, Cham, Switzerland, 2018.

[10] R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.

[11] J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.

[12] H. T. Nguyen, O. Kosheleva, and V. Kreinovich, "Decision making beyond Arrow's 'impossibility theorem', with the analysis of effects of collusion and mutual attraction", *International Journal of Intelligent Systems*, 2009, Vol. 24, No. 1, pp. 27–47.

[13] H. T. Nguyen, V. Kreinovich, B. Wu, and G. Xiang, *Computing Statistics under Interval and Fuzzy Uncertainty*, Springer Verlag, Berlin, Heidelberg, 2012.

[14] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[15] National Institute of Standards and Technology, *Information Security: Guide for Conducting Risak Assessments*, NIST Special Publication 800-30, Revision 1, September 2012.

[16] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[17] H. Raiffa, *Decision Analysis*, McGraw-Hill, Columbus, Ohio, 1997.

[18] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

[19] R. L. Wasserstein and N. A. Lazar, "The ASA's statement on p-values: context, process, and purpose", *The American Statistician*, 2016, Vol. 70, No. 2, pp. 129–133, https://doi.org/10.1080/00031305.2016.1154108

[20] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.