Distortion Models for the Probability of Rare Events: an Application on Human Reliability*

Pablo-Ramsés Alonso-Martín UO257584@UNIOVI.ES
Ignacio Montes IMONTES@UNIOVI.ES
Enrique Miranda MIRANDAENRIQUE@UNIOVI.ES
Department of Statistics and O.R., University of Oviedo, Spain

The estimation of the probability of rare events is a complex problem, where the use of samples may sometimes lead to over or underestimating the probabilities and where the elicitation by experts may be subject to a number of biases. In order to decrease the dependency on the data and increase the robustness of the inferences, one possibility is to consider a *distortion model* [1, 2] around the one determined from the data. These include several imprecise probability models, such as linear-vacuous mixtures, transformations of a probability measure by means of an increasing function or balls centered at a probability measure and with a radius representing the distance at which we want to be robust.

If the available data lead to a probability measure P_0 on the different rare events that may arise, the last approach consists in considering a credal set $B_d^{\delta}(P_0)$ whose associated lower and upper probabilities $\underline{P}$ and $\overline{P}$ satisfy $\overline{P}(A) > 0$ and $\underline{P}(A) < 1$ for every non-trivial event A. In this manner, when for a given event A its estimated probability satisfies $P_0(A) = 0$, the upper probability $\overline{P}(A)$ will be strictly positive, avoiding underestimation; while if $P_0(A) > 0$, the lower probability will satisfy $\underline{P}(A) < P_0(A)$, avoiding overestimations.

The proposed approach is illustrated in the context of *human reliability*, where the goal is to estimate the probabilities of human error. This problem has been tackled using imprecise probabilities in [3, 4]. Specifically, using a dataset with information about 238 major accidents built in [5], in [4] it is built a Bayesian network summarising the connections between the different factors. The problem arises when an estimated conditional probability cannot be computed because the probability of the conditioning event is zero. In that case, [4] proposes to consider a vacuous model, i.e., all the set of the probability measures; this transforms the Bayesian network into a *credal network*, that is then used to give estimations of the probabilities of human error.

In this work, we propose an alternative consisting in distorting the probability measure derived from the data set by means of a linear-vacuous mixture. Our choice of this distortion model is due to several reasons: (i) it produces strictly positive upper probabilities, that allow us to apply regular extension; (ii) it is closed under conditioning; (iii) the associated credal sets have a small number of extreme points, which reduces the complexity with respect to other alternatives; and (iv) it satisfies the property of 2-monotonicity, which simplifies the computation of the conditional model. We consider the credal network associated with this distortion model and compute the estimations of the (lower and upper) probabilities of error, analysing the imprecision for different values of the distortion factor. Comparing our results with those in [4], both approaches turn out to be consistent. This leads us to conclude that our approach for estimating probabilities of rare events is a robust alternative that, according to our computations, performs satisfactory.

References

[1] I. Montes, E. Miranda, and S. Destercke. Unifying neighbourhood and distortion models: Part I- New results on old models. *International Journal of General Systems*, 49(6):602–635, 2020.

[2] I. Montes, E. Miranda, and S. Destercke. Unifying neighbourhood and distortion models: Part II- New models and synthesis. *International Journal of General Systems*, 49(6):636–674, 2020.

[3] C. Morais, R. Moura, M. Beer, and E. Patelli. Attempt to predict human error probability in different industry sectors using data from major accidents and Bayesian networks. In *Proceedings of PSAM*, 2018.

[4] C. Morais, S. Tolo, R. Moura, M. Beer, and E. Patelli. Tackling the lack of data for human error probability with credal network. In *Proceedings of ESREL*, 2019.

[5] R. Moura, M. Beer, E. Patelli, J. Lewis, and F. Knoll. Learning from major accidents to improve system design. *Safety Science*, 84:37–45, 2016.

* *We acknowledge the financial support of project PGC2018-098623-B-I00 from the Ministry of Science, Innovation and Universities of Spain.*