UNIVERSITY OF HULL | SCHOOL OF MATHEMATICS & PHYSICAL SCIENCES

Empirical Interpretation of Imprecise Probabilities

Marco E. G. V. Cattaneo

introduction

- imprecise probabilities can have a clear empirical/frequentist meaning only if they can be estimated from data
- consider for example a (potentially infinite) sequence of bags containing only white and black marbles: we draw one marble at random from each bag, where the proportion of black marbles in the i-th bag is
$$p_i \in [\underline{p}, \overline{p}] \subseteq [0, 1]$$
- if $\underline{p} = \overline{p}$, then $[\underline{p}, \overline{p}]$ represents a precise probability (P), which can be estimated from data without problems (Bernoulli, 1713)
- if $\underline{p} < \overline{p}$, then $[\underline{p}, \overline{p}]$ represents an imprecise probability (IP): can it still be estimated from data?

interpretations of $[\underline{p}, \overline{p}]$

- which sequences of proportions p_i are compatible with the IP $[\underline{p}, \overline{p}]$?
- epistemological indeterminacy interpretation (Walley and Fine, 1982), used e.g. in the theory of Markov chains with IPs (Kozine and Utkin, 2002):
$$p_i = p \in [\underline{p}, \overline{p}]$$
- ontological indeterminacy interpretation (Walley and Fine, 1982), used e.g. in the theories of Markov chains with IPs (Hartfiel, 1998) and probabilistic graphical models with IPs (Cozman, 2005):
$$p_i \in [\underline{p}, \overline{p}]$$
- id-ontological (identifiable ontological indeterminacy interpretation), making $[\underline{p}, \overline{p}]$ identifiable:
$$p_i \in [\underline{p}, \overline{p}] = \left[\liminf_{i\to\infty} p_i,\ \limsup_{i\to\infty} p_i\right]$$
or more generally
$$p_i \in [\underline{p}, \overline{p}] = \left[\underline{\alpha}(p_1, p_2, \ldots),\ \overline{\alpha}(p_1, p_2, \ldots)\right],$$
where $\underline{\alpha}, \overline{\alpha} : [0,1]^{\mathbb{N}} \to [0,1]$ do not depend on any finite number of their arguments

levels of estimability of $[\underline{p}, \overline{p}]$

- let $X_i \in \{0, 1\}$ describe the color of the marble drawn from the i-th bag (e.g. 0 for white and 1 for black), and let $\mathcal{I}$ be the set of all intervals $[\underline{p}, \overline{p}]$ that are considered possible (with $0 \leq \underline{p} \leq \overline{p} \leq 1$)
- we are interested in arbitrarily good estimators $\underline{\pi}_n, \overline{\pi}_n : \{0,1\}^n \to [0,1]$ of $\underline{p}, \overline{p}$: for all $\varepsilon > 0$, all $\delta > 0$, all $[\underline{p}, \overline{p}] \in \mathcal{I}$, and all (precise) probability measures P corresponding to the sequences p_i compatible with $[\underline{p}, \overline{p}]$, there is an N such that for all $n \geq N$,
$$P\left(\left|\underline{\pi}_n(X_1, \ldots, X_n) - \underline{p}\right| > \varepsilon\right) \leq \delta$$
$$P\left(\left|\overline{\pi}_n(X_1, \ldots, X_n) - \overline{p}\right| > \varepsilon\right) \leq \delta$$
- ideal: uniformly consistent estimability (i.e. N cannot depend on $[\underline{p}, \overline{p}]$ or P), meaning that we can construct arbitrarily short confidence intervals for $\underline{p}$ and $\overline{p}$ with arbitrarily high confidence levels (when n is sufficiently large)
- minimal: IP-consistent estimability (i.e. consistent in terms of IPs: N can depend on $[\underline{p}, \overline{p}]$, but not on P), called strong estimability by Walley and Fine (1982), and almost equivalent to the testability of $[\underline{p}, \overline{p}]$ with arbitrarily low significance level and arbitrarily high power for a fixed alternative (when n is sufficiently large)
- inadequate: P-consistent estimability (i.e. consistent in terms of Ps: N can depend on both $[\underline{p}, \overline{p}]$ and P), meaning that $\underline{p}$ and $\overline{p}$ can be estimated arbitrarily well under each compatible sequence p_i (when n is sufficiently large), but the level of precision of the estimator can depend on the particular sequence p_i

conclusion

- IPs $[\underline{p}, \overline{p}]$ can be empirically distinguished only if they are disjoint
- finite-sample IPs $[\min\{p_1, \ldots, p_n\},\ \max\{p_1, \ldots, p_n\}]$ cannot be estimated from data
- the paper summarizes several results that are not surprising, but important to clarify the limited empirical/frequentist meaning of IPs

estimability of $[\underline{p}, \overline{p}]$

interpretation of $[\underline{p}, \overline{p}]$:

estimability of $\underline{p}, \overline{p}$: necessary and sufficient conditions on $[\underline{p}, \overline{p}] \in \mathcal{I}$:	epistemological: $p_i = p \in [\underline{p}, \overline{p}]$	ontological: $p_i \in [\underline{p}, \overline{p}]$	id-ontological: $p_i \in [\underline{p}, \overline{p}]$ s.t. $\underline{p} = \liminf_{i\to\infty} p_i$, $\overline{p} = \limsup_{i\to\infty} p_i$
ideal: uniformly consistent	pairwise disjoint and IPs isolated	pairwise disjoint and IPs isolated	pairwise disjoint and IPs isolated
minimal: IP-consistent	pairwise disjoint	pairwise disjoint	pairwise disjoint
inadequate: P-consistent	pairwise disjoint	pairwise disjoint	?

example of estimators of $\underline{p}, \overline{p}$

$$\underline{\pi}_n(X_1, \ldots, X_n) = \inf\left\{\underline{p} : [\underline{p}, \overline{p}] \in \mathcal{I},\ \overline{p} + c_n > \frac{1}{n}\sum_{i=1}^{n} X_i\right\}$$
$$\overline{\pi}_n(X_1, \ldots, X_n) = \sup\left\{\overline{p} : [\underline{p}, \overline{p}] \in \mathcal{I},\ \underline{p} - c_n < \frac{1}{n}\sum_{i=1}^{n} X_i\right\}$$

satisfy all the above properties (when the corresponding necessary and sufficient conditions on $\mathcal{I}$ are fulfilled), where c_n is any sequence of real numbers such that $\lim_{n\to\infty} c_n = 0$ and $\lim_{n\to\infty} \sqrt{n}\, c_n = +\infty$, while $\inf \varnothing$ and $\sup \varnothing$ can be defined arbitrarily

estimability of $[\min\{p_1, \ldots, p_n\},\ \max\{p_1, \ldots, p_n\}]$

interpretation of $[\underline{p}, \overline{p}]$:

estimability of $\min / \max\{p_1, \ldots, p_n\}$: necessary and sufficient conditions on $[\underline{p}, \overline{p}] \in \mathcal{I}$:	epistemological: $p_i = p \in [\underline{p}, \overline{p}]$	ontological: $p_i \in [\underline{p}, \overline{p}]$	id-ontological: $p_i \in [\underline{p}, \overline{p}]$ s.t. $\underline{p} = \liminf_{i\to\infty} p_i$, $\overline{p} = \limsup_{i\to\infty} p_i$
ideal: uniformly consistent		no IPs	no IPs
minimal: IP-consistent		no IPs	no IPs
inadequate: P-consistent		no IPs	?

example of estimators of $\min\{p_1, \ldots, p_n\}$, $\max\{p_1, \ldots, p_n\}$

$$\underline{\pi}_n(X_1, \ldots, X_n) = \overline{\pi}_n(X_1, \ldots, X_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

satisfy all the above properties (when the corresponding necessary and sufficient conditions on $\mathcal{I}$ are fulfilled)

references

J. Bernoulli. *Ars Conjectandi*. Thurneysen Brothers, 1713.

F. G. Cozman. Graphical models for imprecise probabilities. *International Journal of Approximate Reasoning*, 39:167–184, 2005.

D. J. Hartfiel. *Markov Set-Chains*. Springer, 1998.

I. O. Kozine and L. V. Utkin. Interval-valued finite Markov chains. *Reliable Computing*, 8: 97–113, 2002.

P. Walley and T. L. Fine. Towards a frequentist theory of upper and lower probability. *The Annals of Statistics*, 10:741–761, 1982.