Empirical Interpretation of Imprecise Probabilities

Marco Cattaneo

School of Mathematics and Physical Sciences
University of Hull

ISIPTA '17 – ECSQARU 2017, Lugano, Switzerland
10 July 2017

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:

$$p_i \in [\underline{p}, \overline{p}] = \left[\liminf_{i\to\infty} p_i,\ \limsup_{i\to\infty} p_i\right]$$

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

- assuming that we have a sufficiently large number n of drawings
- ideal: uniformly consistent estimability, meaning that we can construct arbitrarily short confidence intervals for $\underline{p}$ and $\overline{p}$ with arbitrarily high confidence levels
- minimal: IP-consistent estimability (i.e. consistent in terms of IPs), 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
- inadequate: P-consistent estimability (i.e. consistent in terms of Ps), meaning that $\underline{p}$ and $\overline{p}$ can be estimated arbitrarily well under each compatible sequence p_i, but the level of precision of the estimator can depend on the particular sequence p_i

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

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

estimability of $\underline{p}, \overline{p}$:

necessary and sufficient conditions on possible $[\underline{p}, \overline{p}]$:	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	?

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

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

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

necessary and sufficient conditions on possible $[\underline{p}, \overline{p}]$:	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	?

conclusion

- IPs $[\underline{p}, \overline{p}]$ can be empirically distinguished only if they are disjoint
- finite-sample IPs $[\min\{p_1, \dots, p_n\}, \max\{p_1, \dots, p_n\}]$ 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
- examples of estimators with the required properties are given in the paper

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.