On the Robustness of Imprecise Probability Methods

Marco Cattaneo
Department of Statistics, LMU Munich

ISIPTA '13, Compiègne, France
2 July 2013

introduction

- IP methods are often claimed to be robust (or more robust than conventional methods): **are they really robust?**

introduction

- IP methods are often claimed to be robust (or more robust than conventional methods): **are they really robust?**
- "robustness signifies insensitivity to small deviations from the assumptions"
  (Huber, 1981: *Robust Stat.*, p. 1)

introduction

- IP methods are often claimed to be robust (or more robust than conventional methods): **are they really robust?**

- "robustness signifies insensitivity to small deviations from the assumptions"
  (Huber, 1981: *Robust Stat.*, p. 1)

- in the IP approach:

introduction

- IP methods are often claimed to be robust (or more robust than conventional methods): **are they really robust?**
- "robustness signifies insensitivity to small deviations from the assumptions"
  (Huber, 1981: *Robust Stat.*, p. 1)
- in the IP approach:
  - probability values $P(A)$ need not be precisely chosen, they are replaced by intervals $[\underline{P}(A), \overline{P}(A)]$,

introduction

- IP methods are often claimed to be robust (or more robust than conventional methods): **are they really robust?**
- "robustness signifies insensitivity to small deviations from the assumptions"
  (Huber, 1981: *Robust Stat.*, p. 1)
- in the IP approach:
  - probability values $P(A)$ need not be precisely chosen, they are replaced by intervals $[\underline{P}(A), \overline{P}(A)]$,
  - but this means choosing two values precisely: $\underline{P}(A)$ and $\overline{P}(A)$

introduction

- IP methods are often claimed to be robust (or more robust than conventional methods): **are they really robust?**
- "robustness signifies insensitivity to small deviations from the assumptions"
  (Huber, 1981: *Robust Stat.*, p. 1)
- in the IP approach:
  - probability values $P(A)$ need not be precisely chosen, they are replaced by intervals $[\underline{P}(A), \overline{P}(A)]$,
  - but this means choosing two values precisely: $\underline{P}(A)$ and $\overline{P}(A)$
- the robustness of the conventional methods refers to the arbitrariness in the choice of $P(A)$, while the robustness of the IP methods refers to the arbitrariness in the choices of $\underline{P}(A)$ and $\overline{P}(A)$

robust or not robust?

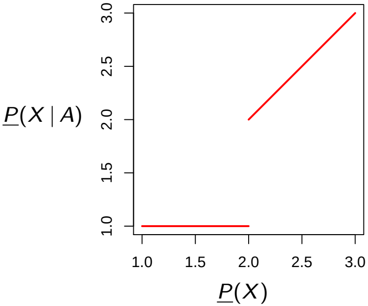
- natural extension of IP models: **robust** (Troffaes and Hable, *ISIPTA '11*)

robust or not robust?

- natural extension of IP models: **robust** (Troffaes and Hable, *ISIPTA '11*)
- updating of IP models (by natural/regular extension): **not robust** in general

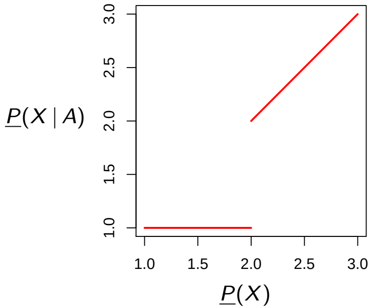
robust or not robust?

- natural extension of IP models: **robust** (Troffaes and Hable, *ISIPTA '11*)
- updating of IP models (by natural/regular extension): **not robust** in general
- e.g., $X \in \{1, 2, 3\}$, unique assessment: $\underline{P}(X)$, observation: $A = \{X \neq 2\}$



robust or not robust?

- natural extension of IP models: **robust** (Troffaes and Hable, *ISIPTA '11*)
- updating of IP models (by natural/regular extension): **not robust** in general
- e.g., $X \in \{1, 2, 3\}$, unique assessment: $\underline{P}(X)$, observation: $A = \{X \neq 2\}$



- by contrast, updating of precise probabilities is continuous

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)
- e.g., sequence of binary experiments, starting with "complete ignorance":

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"

  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"

  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$
    - $s = ?$

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"

  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$
    - $s = ?$
  - IP approach:

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$
    - $s = ?$
  - IP approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$
    - $s = ?$
  - IP approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate priors)

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"
  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$
    - $s = ?$
  - IP approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate priors)
    - $t \in (0, 1)$

doubtful assumptions

- "conclusions drawn from the imprecise model are automatically robust, because they do not rely on arbitrary or doubtful assumptions"

  (Walley, 1991: *Stat. Reasoning with IP*, p. 5)

- e.g., sequence of binary experiments, starting with "complete ignorance":
  - Bayesian approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate prior)
    - $t = 1/2$
    - $s = ?$
  - IP approach:
    - $X_1, X_2, \ldots \overset{i.i.d.}{\sim} Ber(\theta)$ conditional on θ (exchangeability)
    - $\theta \sim Beta(s, t)$ (conjugate priors)
    - $t \in (0, 1)$
    - $s = ?$

misleading comparisons

- imprecise methods based on IP models are often compared with precise methods based on precise probabilities

misleading comparisons

- imprecise methods based on IP models are often compared with precise methods based on precise probabilities

- e.g., imprecise classifiers based on IDM priors are compared with Bayesian classifiers based on uniform priors

misleading comparisons

- imprecise methods based on IP models are often compared with precise methods based on precise probabilities
- e.g., imprecise classifiers based on IDM priors are compared with Bayesian classifiers based on uniform priors
- in some situations, imprecise methods can be more robust, but they can be based on precise probabilities as well:

		probability model:	
		precise	imprecise
method:	precise	$\checkmark$	$\checkmark$
	imprecise	$\checkmark$	$\checkmark$

misleading comparisons

- imprecise methods based on IP models are often compared with precise methods based on precise probabilities
- e.g., imprecise classifiers based on IDM priors are compared with Bayesian classifiers based on uniform priors
- in some situations, imprecise methods can be more robust, but they can be based on precise probabilities as well:

		probability model:	
		precise	imprecise
method:	precise	✓	✓
	imprecise	✓	✓

- the gain in robustness is obtained by allowing the methods to be imprecise, and not necessarily by basing them on IP models

conclusion

- IP methods are robust if they are insensitive to small deviations from the assumed IP model

conclusion

- IP methods are robust if they are insensitive to small deviations from the assumed IP model
- there seems to be **no reason** to claim that IP methods are in general robust (or more robust than conventional methods)

conclusion

- IP methods are robust if they are insensitive to small deviations from the assumed IP model
- there seems to be **no reason** to claim that IP methods are in general robust (or more robust than conventional methods)
- this is particularly important in statistics, where the (higher) robustness of the IP approach could have been one of the few general advantages over the Bayesian approach

conclusion

- IP methods are robust if they are insensitive to small deviations from the assumed IP model
- there seems to be **no reason** to claim that IP methods are in general robust (or more robust than conventional methods)
- this is particularly important in statistics, where the (higher) robustness of the IP approach could have been one of the few general advantages over the Bayesian approach
- nevertheless, IP models (with the sensitivity analysis interpretation) can be useful as a tool for studying the robustness of Bayesian methods