Updating and avoiding sure loss

Marco Cattaneo
Department of Physics and Mathematics, University of Hull

WPMSIIP 2014, Ghent, Belgium
9 September 2014

introduction

updating rules	Bayesian	coherent	α-cut
continuity	✓	×	✓
vacuous priors	×	×	✓
iterative consistency	✓	✓	×
coherence	✓	✓	×

updating rules

- when an event $B \subset \Omega$ is observed, previsions must be updated

updating rules

- when an event $B \subset \Omega$ is observed, previsions must be updated
- Bayesian updating of linear previsions P with $P(B) > 0$:

$$P \mapsto P(\,\cdot \mid B) \quad \text{with} \quad P(X \mid B) = \frac{P(X\, I_B)}{P(B)} \quad \text{for all bounded} \quad X : \Omega \to \mathbb{R}$$

updating rules

- when an event $B \subset \Omega$ is observed, previsions must be updated

- Bayesian updating of linear previsions P with $P(B) > 0$:

$$P \mapsto P(\cdot \mid B) \quad \text{with} \quad P(X \mid B) = \frac{P(X\, I_B)}{P(B)} \quad \text{for all bounded} \quad X : \Omega \to \mathbb{R}$$

- generalizations of Bayesian updating to (coherent) lower previsions $\underline{P}$ with $\overline{P}(B) > 0$:

updating rules

- when an event $B \subset \Omega$ is observed, previsions must be updated
- Bayesian updating of linear previsions P with $P(B) > 0$:

$$P \mapsto P(\cdot \mid B) \quad \text{with} \quad P(X \mid B) = \frac{P(X\, I_B)}{P(B)} \quad \text{for all bounded} \quad X : \Omega \to \mathbb{R}$$

- generalizations of Bayesian updating to (coherent) lower previsions $\underline{P}$ with $\overline{P}(B) > 0$:
  - regular extension:

$$\underline{P} \mapsto \inf_{P \geq \underline{P},\, P(B) > 0} P(\cdot \mid B)$$

updating rules

- when an event $B \subset \Omega$ is observed, previsions must be updated

- Bayesian updating of linear previsions P with $P(B) > 0$:

$$P \mapsto P(\,\cdot \mid B) \quad \text{with} \quad P(X \mid B) = \frac{P(X\, I_B)}{P(B)} \quad \text{for all bounded} \quad X : \Omega \to \mathbb{R}$$

- generalizations of Bayesian updating to (coherent) lower previsions $\underline{P}$ with $\overline{P}(B) > 0$:

  - regular extension:

$$\underline{P} \mapsto \inf_{P \geq \underline{P},\, P(B) > 0} P(\,\cdot \mid B)$$

  - α-cut rule, where $\alpha \in (0, 1)$:

$$\underline{P} \mapsto \inf_{P \geq \underline{P},\, P(B) \geq \alpha\, \overline{P}(B)} P(\,\cdot \mid B)$$

continuity

- Bayesian updating is continuous with respect to the metric

$$d(P, P') = \sup_{X : \Omega \to [-1, 1]} |P(X) - P'(X)|$$

continuity

- Bayesian updating is continuous with respect to the metric

$$d(P, P') = \sup_{X\,:\,\Omega \to [-1,1]} \left|P(X) - P'(X)\right|$$

- α-cut updating is continuous with respect to the (Hausdorff) metric

$$d(\underline{P}, \underline{P}') = \sup_{X\,:\,\Omega \to [-1,1]} \left|\underline{P}(X) - \underline{P}'(X)\right|$$

  for all $\alpha \in (0, 1)$, while regular/natural/coherent updating has discontinuities at points $\underline{P}$ with $\overline{P}(B) > \underline{P}(B) = 0$ (Cattaneo, 2014)

vacuous priors

- contrary to linear previsions, lower previsions can describe prior ignorance about the unknowns of a statistical model

vacuous priors

- contrary to linear previsions, lower previsions can describe prior ignorance about the unknowns of a statistical model

- contrary to regular/natural/coherent updating, the α-cut rule can update vacuous priors to non-vacuous posteriors in statistical analyses

vacuous priors

- contrary to linear previsions, lower previsions can describe prior ignorance about the unknowns of a statistical model

- contrary to regular/natural/coherent updating, the α-cut rule can update vacuous priors to non-vacuous posteriors in statistical analyses

- for regular statistical models, the imprecise previsions obtained from vacuous priors by means of α-cut updating are confidence intervals with (asymptotic) level $F_{\chi^2}(-2 \ln \alpha)$ (Wilks, 1938)

iterative consistency

- Bayesian updating is iteratively consistent:

$$P((\,\cdot \mid B) \mid C) = P(\,\cdot \mid B \cap C) = P((\,\cdot \mid C) \mid B)$$

iterative consistency

- Bayesian updating is iteratively consistent:

$$P((\cdot \mid B) \mid C) = P(\cdot \mid B \cap C) = P((\cdot \mid C) \mid B)$$

- contrary to regular/natural updating, the α-cut rule is not iteratively consistent in general

iterative consistency

- Bayesian updating is iteratively consistent:

$$P((\cdot \mid B) \mid C) = P(\cdot \mid B \cap C) = P((\cdot \mid C) \mid B)$$

- contrary to regular/natural updating, the α-cut rule is not iteratively consistent in general

- this can be remedied by recording the whole (second-order) likelihood function $lik(P) \propto P(observations)$ as the second level of a hierarchical model (Cattaneo, 2008)

iterative consistency

- Bayesian updating is iteratively consistent:

$$P((\,\cdot \mid B) \mid C) = P(\,\cdot \mid B \cap C) = P((\,\cdot \mid C) \mid B)$$

- contrary to regular/natural updating, the α-cut rule is not iteratively consistent in general

- this can be remedied by recording the whole (second-order) likelihood function $lik(P) \propto P(observations)$ as the second level of a hierarchical model (Cattaneo, 2008)

- the α-cut is then a way of obtaining lower previsions from the hierarchical model

coherence

- contrary to Bayesian and regular/natural/coherent updating, the α-cut rule does not avoid sure loss in general

coherence

- contrary to Bayesian and regular/natural/coherent updating, the α-cut rule does not avoid sure loss in general

- example:

$$\begin{aligned} X &\sim \textit{Bernoulli}(\tfrac{1}{2}) \\ (Y \mid X = 0) &\sim \textit{uniform on } \{1, \ldots, n\} \\ (Y \mid X = 1) &\sim \textit{vacuous on } \{1, \ldots, n\} \end{aligned}$$

coherence

- contrary to Bayesian and regular/natural/coherent updating, the α-cut rule does not avoid sure loss in general

- example:

$$\begin{aligned} X &\sim \textit{Bernoulli}(\tfrac{1}{2}) \\ (Y \mid X = 0) &\sim \textit{uniform on } \{1, \dots, n\} \\ (Y \mid X = 1) &\sim \textit{vacuous on } \{1, \dots, n\} \end{aligned}$$

- regular/natural/coherent updating:

$$\underline{P}(X = 1 \mid Y) = 0 \quad \text{and} \quad \overline{P}(X = 1 \mid Y) = \frac{n}{n+1} \quad (\text{dilation})$$

coherence

- contrary to Bayesian and regular/natural/coherent updating, the α-cut rule does not avoid sure loss in general
- example:

$$\begin{aligned} X &\sim \textit{Bernoulli}(\tfrac{1}{2}) \\ (Y \mid X = 0) &\sim \textit{uniform on } \{1, \ldots, n\} \\ (Y \mid X = 1) &\sim \textit{vacuous on } \{1, \ldots, n\} \end{aligned}$$

- regular/natural/coherent updating:

$$\underline{P}(X = 1 \mid Y) = 0 \quad \text{and} \quad \overline{P}(X = 1 \mid Y) = \frac{n}{n+1} \quad \text{(dilation)}$$

- α-cut updating:

$$\underline{P}(X = 1 \mid Y) = \left(1 - \frac{1}{(n+1)\,\alpha}\right) \vee 0 \quad \xrightarrow{n \to \infty} \quad 1 \quad \text{(sure loss)}$$

conclusion

- the relative importance of properties like coherence, continuity, and ability of using vacuous priors depends on the application field and on the exact interpretation of imprecise probabilities

conclusion

- the relative importance of properties like coherence, continuity, and ability of using vacuous priors depends on the application field and on the exact interpretation of imprecise probabilities

- α-cut updating can be seen as a continuous approximation of coherent updating, but as a general approach it is only reasonable when the whole (second-order) likelihood function is recorded (iterative consistency)

conclusion

- the relative importance of properties like coherence, continuity, and ability of using vacuous priors depends on the application field and on the exact interpretation of imprecise probabilities
- α-cut updating can be seen as a continuous approximation of coherent updating, but as a general approach it is only reasonable when the whole (second-order) likelihood function is recorded (iterative consistency)
- no (reasonable) method using the second-order likelihood function to obtain lower previsions can avoid sure loss in general

references

Cattaneo, M. (2008). Fuzzy probabilities based on the likelihood function. In *Soft Methods for Handling Variability and Imprecision*. Springer, 43–50.

Cattaneo, M. (2014). A continuous updating rule for imprecise probabilities. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems, Part 3*. Springer, 426–435.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.* 9, 60–62.