On Prior-Data Conflict in Predictive Bernoulli Inferences

Gero Walter, Thomas Augustin
Department of Statistics
Ludwig-Maximilians-Universität (LMU) München, Germany

Frank P.A. Coolen
Department of Mathematics
Durham University, UK

Bernoulli Data and Prior-Data Conflict

Bernoulli Data

- Bernoulli observations are 0/1 observations (head/tails when tossing a coin)
- given: a set of observations (12 out of 16 tosses were heads)
- additional to observations, we have strong prior information (we are convinced that $P(\text{heads})$ should be around 0.75)
- interested in probability $\mathbf{P}$ that the next observation is a head. (predictive probability!)

The Beta-Bernoulli/Binomial Model (BBM)
(in Walley's parametrization)

$$\begin{array}{rl} \text{data}: & s \sim \text{Binom}(p, n) \\ \text{conjugate prior}: & p \sim \text{Beta}(n^{(0)}, y^{(0)}) \\ \hline \text{posterior}: & p \mid s \sim \text{Beta}(n^{(n)}, y^{(n)}) \end{array}$$

$$y^{(n)} = \frac{n^{(0)}}{n^{(0)}+n} \cdot y^{(0)} + \frac{n}{n^{(0)}+n} \cdot \frac{s}{n} = \mathbf{P}\,,$$

$$n^{(n)} = n^{(0)} + n\,.$$

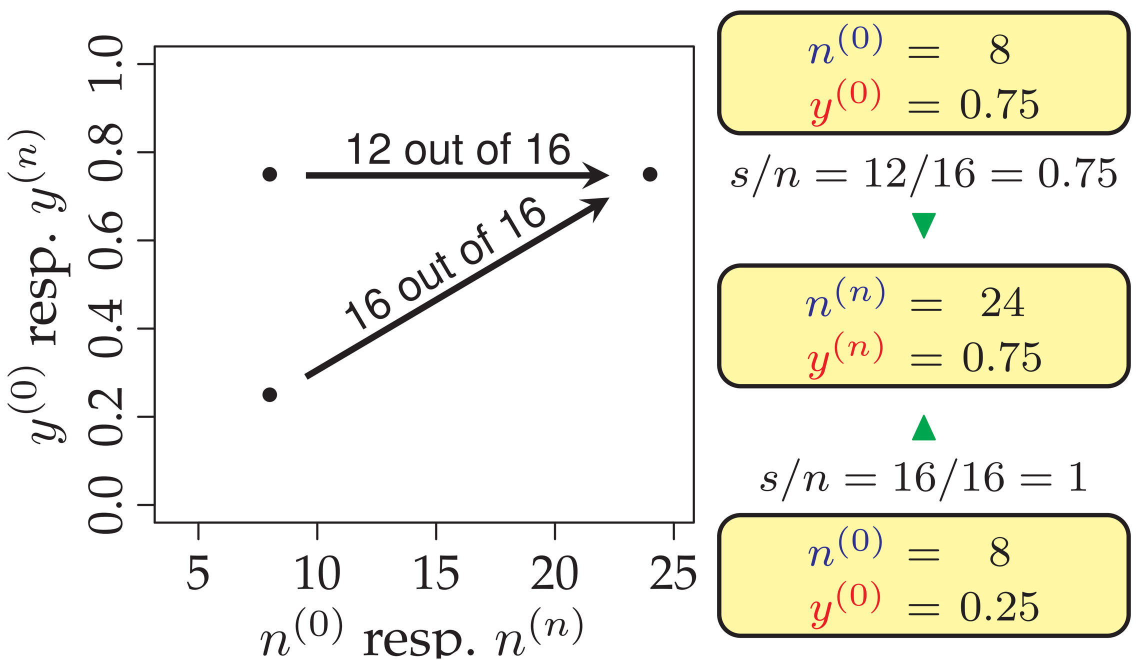
Prior-Data Conflict

If $P(\text{heads}) = p$ for the coin is actually very different from our prior guess $y^{(0)}$ (i.e., prior information and data are in conflict), this should show up in the predictive inferences (probability $\mathbf{P}$ and, e.g., confidence intervals). However, as

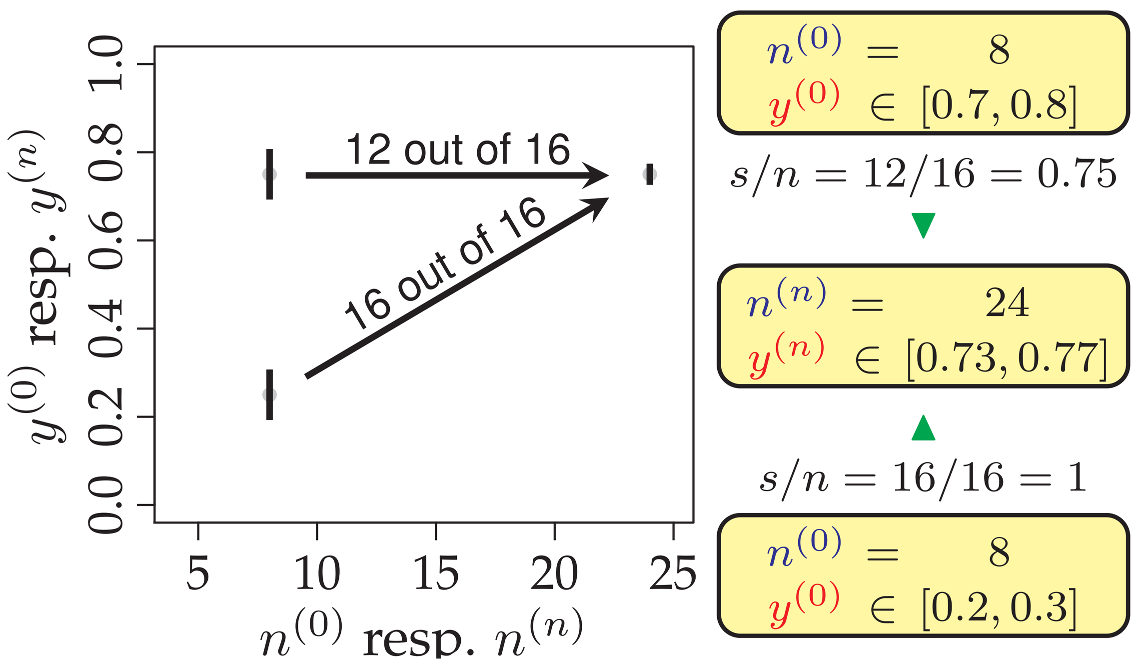
$$\text{Var}(p \mid s) = \frac{y^{(n)}(1-y^{(n)})}{n^{(n)}+1}\,,$$

a systematic reaction to prior-data conflict is not possible for the BBM.

Beta-Binomial Model (BBM)



Imprecise BBM $\hat{=}$ IDM (Walley 1996)



Weighted Inference

Combine predictive inferences of

1. an *uninformative* model ➡ $[\underline{\mathbf{P}}^u, \overline{\mathbf{P}}^u]$ from, e.g., a near-ignorance prior, here:

   $n^{(0)} = 1,\ y^{(0)} \in [0, 1]$

2. an *informative* model ➡ $[\underline{\mathbf{P}}^i, \overline{\mathbf{P}}^i]$ from, e.g., an informative prior, here:

   $n^{(0)} = n^i + 1,\ y^{(0)} \in \left[\frac{s^i}{n^i+1}, \frac{s^i+1}{n^i+1}\right]$

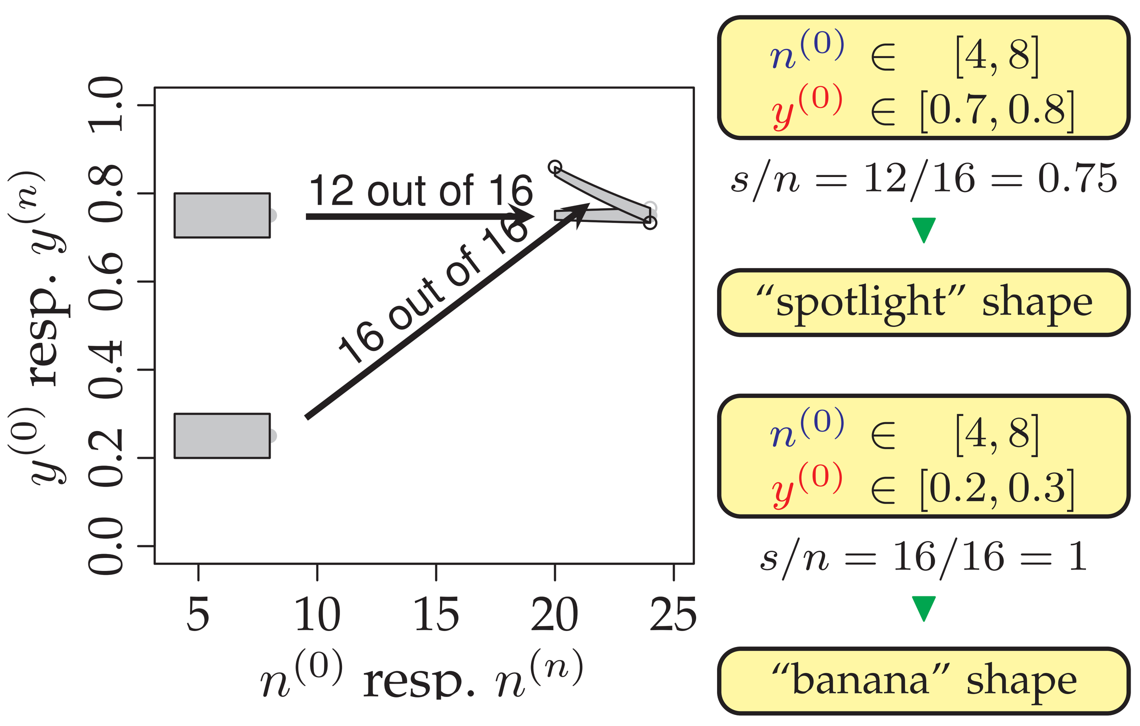
by weighing them imprecisely:

$$\underline{\mathbf{P}} = \min_{\alpha \in [\alpha_l, \alpha_r]} \underline{\mathbf{P}}_\alpha, \text{ where } \underline{\mathbf{P}}_\alpha = \alpha \underline{\mathbf{P}}^i + (1-\alpha)\underline{\mathbf{P}}^u$$

$$\overline{\mathbf{P}} = \max_{\alpha \in [\alpha_l, \alpha_r]} \overline{\mathbf{P}}_\alpha, \text{ where } \overline{\mathbf{P}}_\alpha = \alpha \overline{\mathbf{P}}^i + (1-\alpha)\overline{\mathbf{P}}^u$$

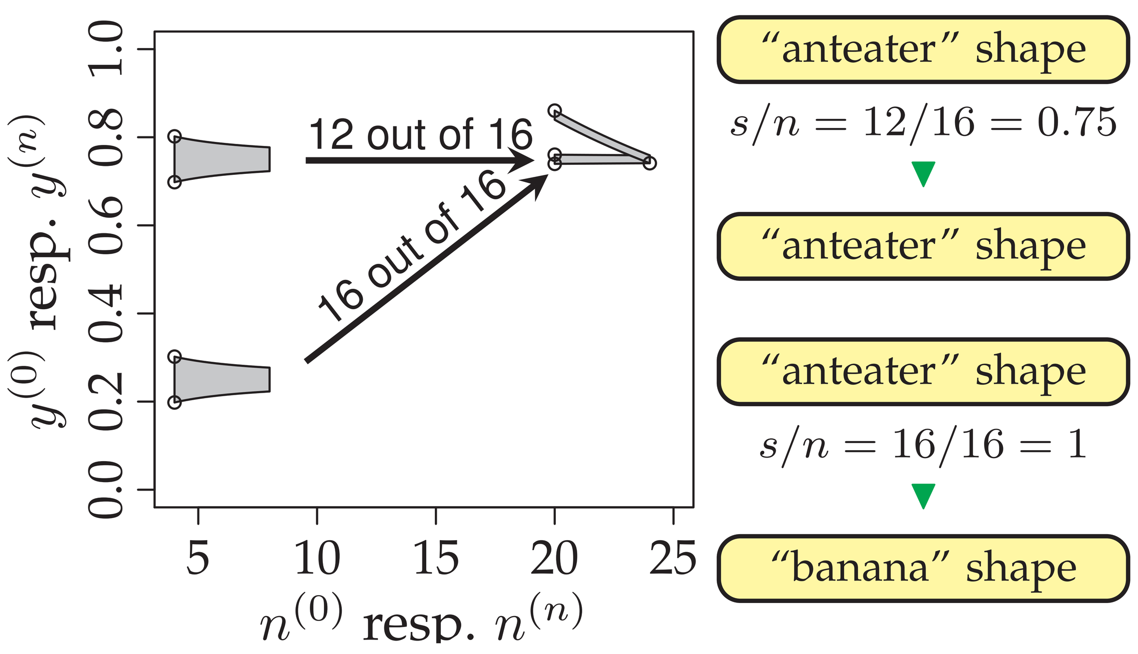
For $\alpha \in [0, 1]$, the same inference as for a pdc-IBBM result, where the prior parameter set consists of the union of the parameter sets of uninformative and informative model.

pdc-Imprecise BBM (Walley 1991)



posterior parameter set $= \{(n^{(n)}, y^{(n)}) \mid n^{(0)} \in [\underline{n}^{(0)}, \overline{n}^{(0)}], y^{(0)} \in [\underline{y}^{(0)}, \overline{y}^{(0)}]\}$

Anteater Shape



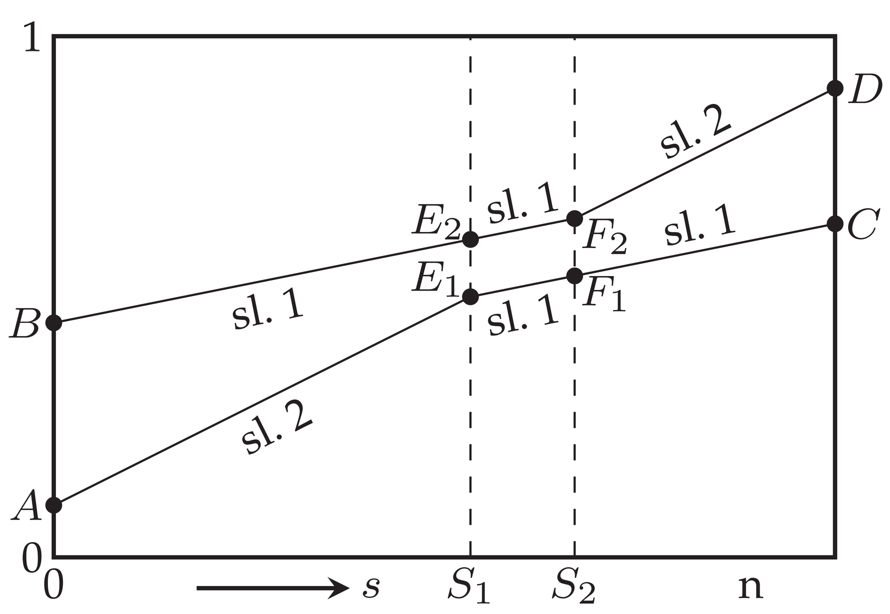
posterior param. set $= \{(n^{(n)}, y^{(n)}) \mid$ "anteater set"$\}$

Anteater sets are derived by a thought experiment on the effect of hypothetical observations.

Predictive Probability Plots (PPP)

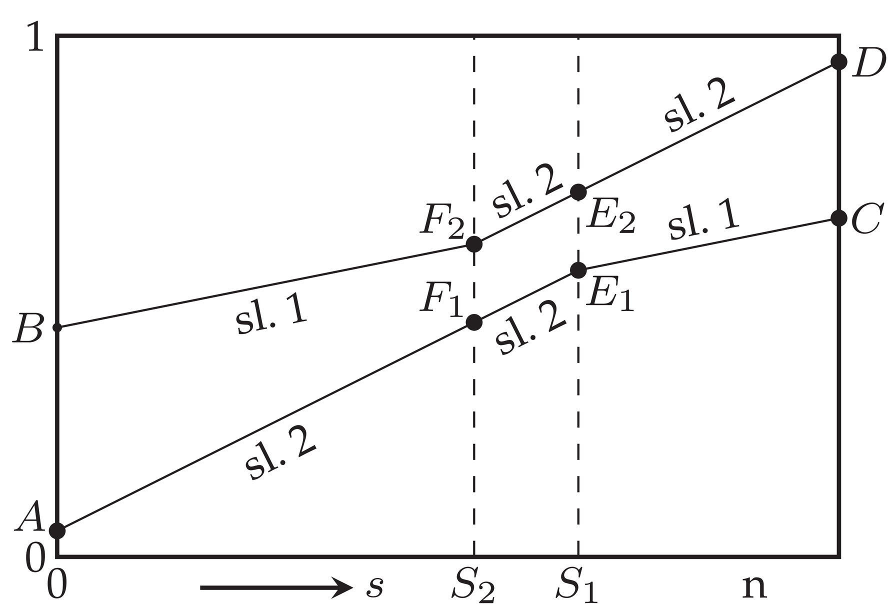
The Predictive Probability Plot (PPP) displays $[\underline{\mathbf{P}}, \overline{\mathbf{P}}]$ as a function of s, indicating how imprecision increases in case of prior-data conflict.

PPP for pdc-IBBM



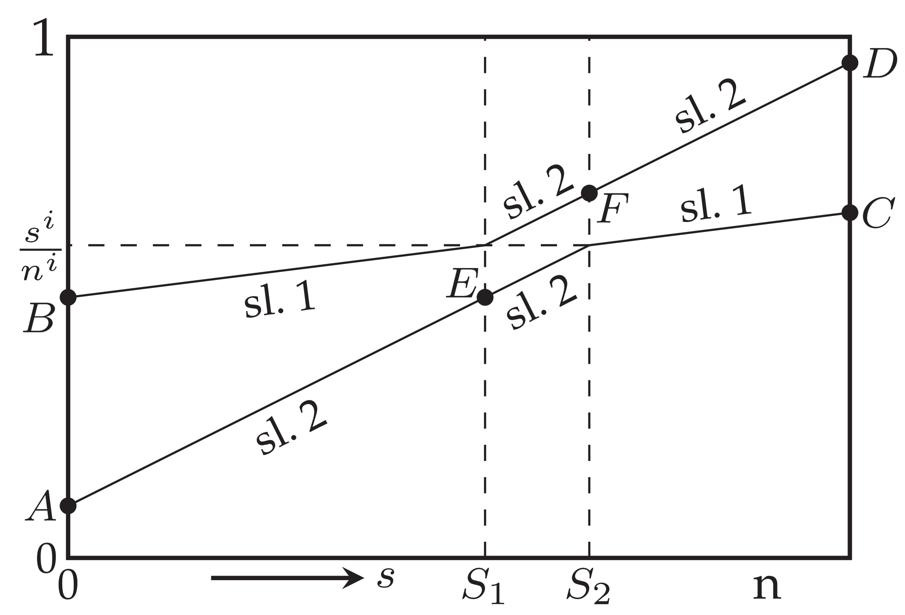
See Section 2.2 of the paper for the values A, B, C, etc.

PPP for the Anteater Shape



See Section 2.3 of the paper for the values A, B, C, etc.

PPP for Weighted Inference



See Section 3.2 of the paper for the values A, B, C, etc.

Generalizations

In principal, any prior set shape is possible, leading to different behaviour, e.g., a certain number of slopes in the PPP. Sets of BBMs can be generalized to any distribution from an exponential family, as those have the same weighted average structure for $y^{(n)}$ (Quaghebeur & de Cooman 2005). A generalization of pdc-IBBM along this lines was presented in (Walter & Augustin 2009).

Generalizations

The method can be used to combine any two predictive inferences, from any model, on any event of interest. A possible source for $\mathbf{P}^u/\mathbf{P}^i$ is, e.g., the NPI model (Coolen & Augustin 2009).

References

F.P.A. Coolen & T. Augustin. A nonparametric predictive alternative to the Imprecise Dirichlet Model: the case of a known number of categories. *International Journal of Approximate Reasoning*, 50:217–230, 2009.

E. Quaeghebeur & G. de Cooman. Imprecise probability models for inference in exponential families. In F.G. Cozman, R. Nau, and T. Seidenfeld, editors, *ISIPTA '05*, pp 287–296, 2005.

P. Walley. Inferences from multinomial data: learning about a bag of marbles. *JRSS B*, 58:3–57, 1996.

G. Walter & T. Augustin. Imprecision and prior-data conflict in generalized Bayesian inference. *Journal of Statistical Theory and Practice*, 3:255–271, 2009.

P. Walley. *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London, 1991.