Boat or Bullet: Prior Parameter Set Shapes and Posterior Imprecision

Gero Walter, Frank Coolen & Mik Bickis

Gero Walter, Department of Statistics
Ludwig-Maximilians-Universität München (LMU)
gero.walter@stat.uni-muenchen.de

June 29th, 2012

Introduction

- Bernoulli observations: 0/1 observations (team wins no/yes)

Introduction

- Bernoulli observations: 0/1 observations (team wins no/yes)
- given: a set of observations (team won 12 out of 16 matches)

Introduction

- Bernoulli observations: 0/1 observations (team wins no/yes)
- given: a set of observations (team won 12 out of 16 matches)
- additional to observations, we have strong prior information (we are convinced that $P(\text{win})$ should be around 0.75)

Introduction

- Bernoulli observations: 0/1 observations (team wins no/yes)
- given: a set of observations (team won 12 out of 16 matches)
- additional to observations, we have strong prior information (we are convinced that $P(\text{win})$ should be around 0.75)
- we are, e.g., interested in (predictive) probability P that team wins in the next match

Introduction

- Bernoulli observations: 0/1 observations (team wins no/yes)
- given: a set of observations (team won 12 out of 16 matches)
- additional to observations, we have strong prior information (we are convinced that $P(\text{win})$ should be around 0.75)
- we are, e.g., interested in (predictive) probability P that team wins in the next match
- standard statistical model for this situation: Beta-Bernoulli/Binomial Model

Introduction

- Bernoulli observations: 0/1 observations (team wins no/yes)
- given: a set of observations (team won 12 out of 16 matches)
- additional to observations, we have strong prior information (we are convinced that $P(\text{win})$ should be around 0.75)
- we are, e.g., interested in (predictive) probability P that team wins in the next match
- standard statistical model for this situation: Beta-Bernoulli/Binomial Model
- **prior-data conflict:** if $P(\text{win})$ is actually very different from our prior guess (prior information and data are in conflict), this should show up in the predictive inferences (probability P and, e.g., credibility intervals)

Beta-Bernoulli/Binomial Model (BBM)

- Beta prior on $p = P(\text{win})$
- here in parameterization used, e.g., by Walley (1991):

Beta-Bernoulli/Binomial Model (BBM)

- Beta prior on $p = P(\text{win})$
- here in parameterization used, e.g., by Walley (1991):

Data :	s	$\sim$	$\text{Binom}(p, n)$
conjugate prior:	p	$\sim$	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$p \mid s$	$\sim$	$\text{Beta}(n^{(n)}, y^{(n)})$

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

Beta-Bernoulli/Binomial Model (BBM)

- Beta prior on $p = P(\text{win})$
- here in parameterization used, e.g., by Walley (1991):

Data :	s	$\sim$	$\text{Binom}(p, n)$
conjugate prior:	p	$\sim$	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$p \mid s$	$\sim$	$\text{Beta}(n^{(n)}, y^{(n)})$

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

$$y^{(n)} = \text{E}[p \mid s] \qquad \text{Var}(p \mid s) = \frac{y^{(n)}(1 - y^{(n)})}{n^{(n)} + 1}$$

Beta-Bernoulli/Binomial Model (BBM)

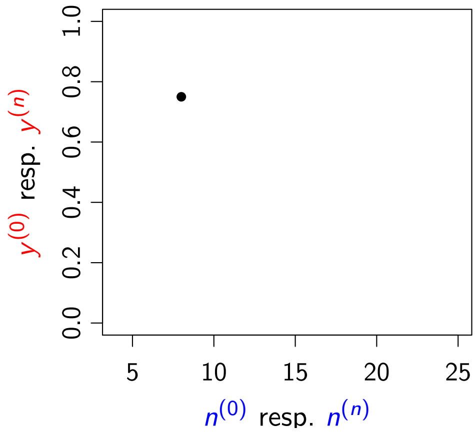
- Beta prior on $p = P(\text{win})$
- here in parameterization used, e.g., by Walley (1991):

Data :	s	$\sim$	$\text{Binom}(p, n)$
conjugate prior:	p	$\sim$	$\text{Beta}(n^{(0)}, y^{(0)})$
posterior:	$p \mid s$	$\sim$	$\text{Beta}(n^{(n)}, y^{(n)})$

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

$$y^{(n)} = \text{E}[p \mid s] = \text{P} \qquad \text{Var}(p \mid s) = \frac{y^{(n)}(1 - y^{(n)})}{n^{(n)} + 1}$$

Beta-Bernoulli/Binomial Model (BBM)

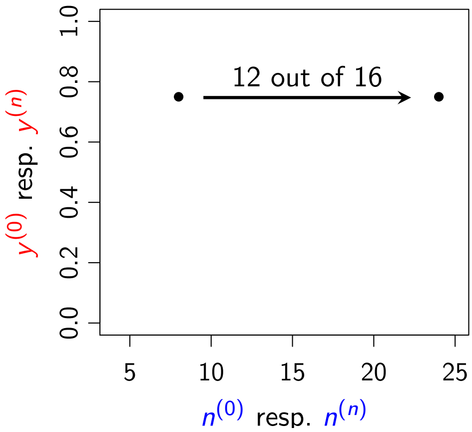


no conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.75$

data $s/n = 12/16 = 0.75$

Beta-Bernoulli/Binomial Model (BBM)

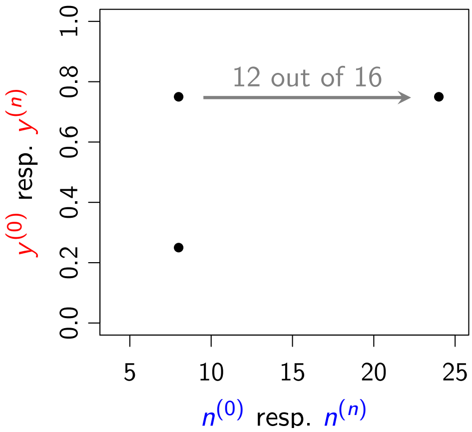


no conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.75$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24$, $y^{(n)} = 0.75$

Beta-Bernoulli/Binomial Model (BBM)



no conflict:

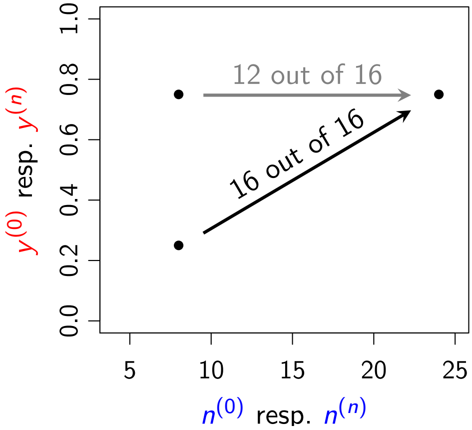
prior $n^{(0)} = 8$, $y^{(0)} = 0.75$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24$, $y^{(n)} = 0.75$

prior-data conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.25$
data $s/n = 16/16 = 1$

Beta-Bernoulli/Binomial Model (BBM)



no conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.75$
data $s/n = 12/16 = 0.75$

▼

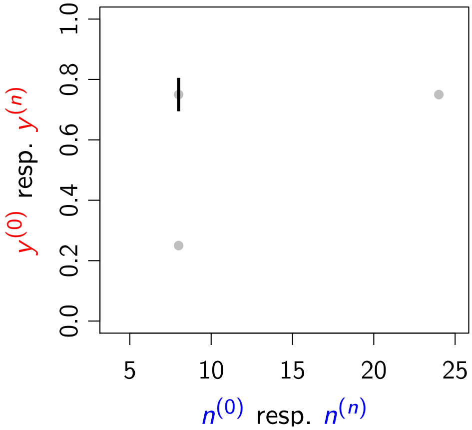
$n^{(n)} = 24$, $y^{(n)} = 0.75$

▲

prior-data conflict:

prior $n^{(0)} = 8$, $y^{(0)} = 0.25$
data $s/n = 16/16 = 1$

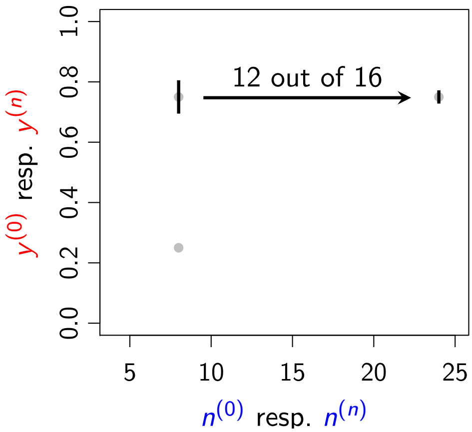
Imprecise BBM (IBBM) $\hat{=}$ IDM with prior information



no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

Imprecise BBM (IBBM) $\hat{=}$ IDM with prior information

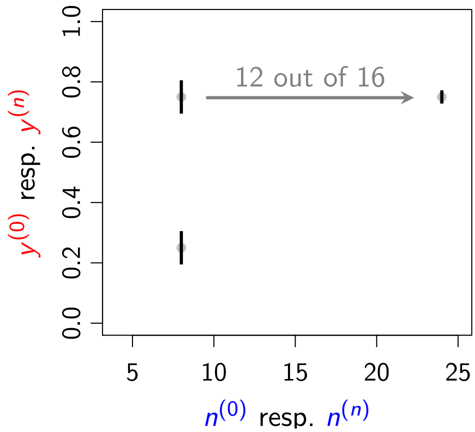


no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24$, $y^{(n)} \in [0.73, 0.77]$

Imprecise BBM (IBBM) $\hat{=}$ IDM with prior information



no conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

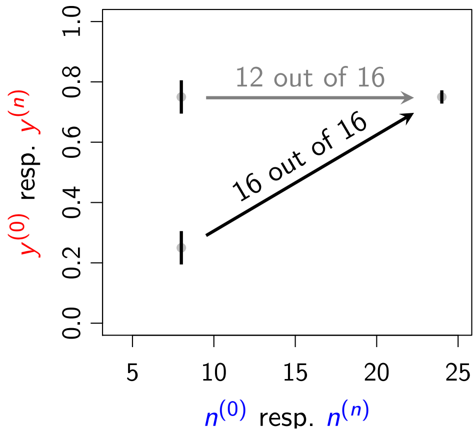
▼

$n^{(n)} = 24$, $y^{(n)} \in [0.73, 0.77]$

prior-data conflict:

prior $n^{(0)} = 8$, $y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

Imprecise BBM (IBBM) $\hat{=}$ IDM with prior information



no conflict:

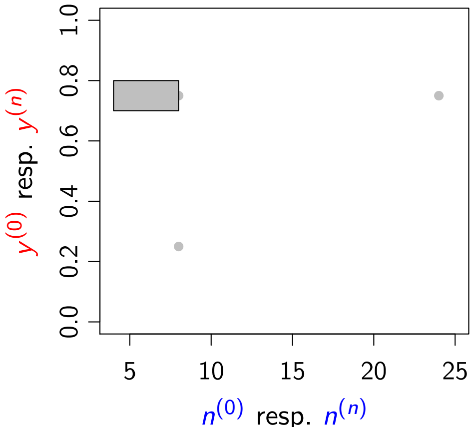
prior $n^{(0)} = 8,\ y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

$n^{(n)} = 24,\ y^{(n)} \in [0.73, 0.77]$

prior-data conflict:

prior $n^{(0)} = 8,\ y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

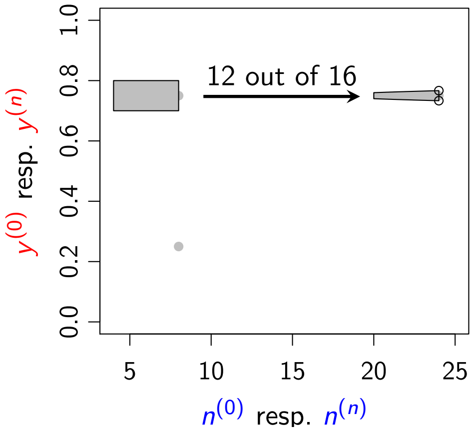
pdc-Imprecise BBM (pdc-IBBM): Walley 1991, Ch.5.4.3



no conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

pdc-Imprecise BBM (pdc-IBBM): Walley 1991, Ch.5.4.3

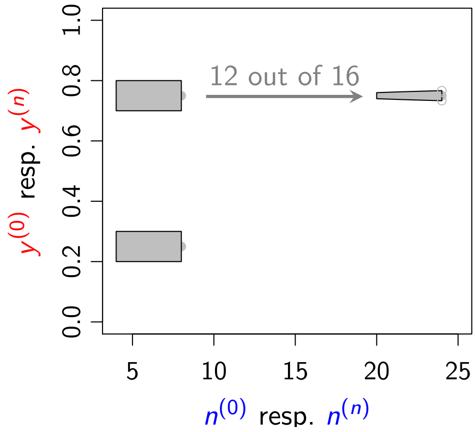


no conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

"spotlight" shape

pdc-Imprecise BBM (pdc-IBBM): Walley 1991, Ch.5.4.3



no conflict:

prior $n^{(0)} \in [4,8]$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

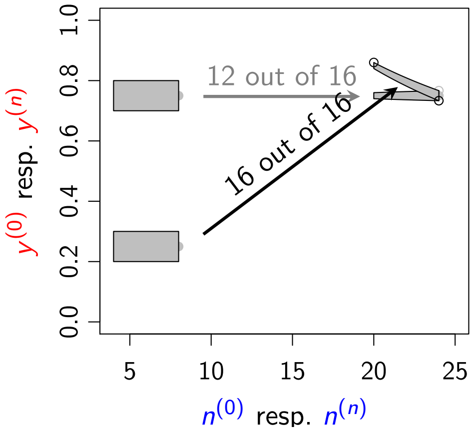
▼

"spotlight" shape

prior-data conflict:

prior $n^{(0)} \in [4,8]$, $y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

pdc-Imprecise BBM (pdc-IBBM): Walley 1991, Ch.5.4.3



no conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.7, 0.8]$
data $s/n = 12/16 = 0.75$

▼

"spotlight" shape

prior-data conflict:

prior $n^{(0)} \in [4, 8]$, $y^{(0)} \in [0.2, 0.3]$
data $s/n = 16/16 = 1$

▼

"banana" shape

What are the Credal Sets?

- credal set $\triangleq$ finite convex mixtures of parametric priors
  to ensure coherence of prior and posterior

What are the Credal Sets?

- credal set $\hat{=}$ finite convex mixtures of parametric priors to ensure coherence of prior and posterior
- sets of distributions are induced by sets of parameters (parameter sets need not be convex)

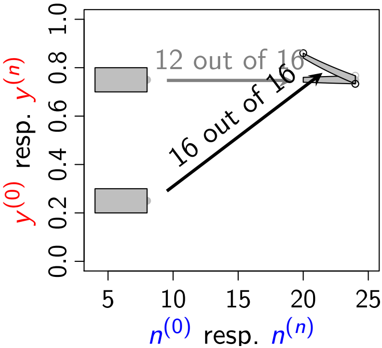
What are the Credal Sets?

- credal set $\hat{=}$ finite convex mixtures of parametric priors to ensure coherence of prior and posterior
- sets of distributions are induced by sets of parameters (parameter sets need not be convex)
- inferences should be linear in posterior distributions: then min/max are attained at the parametric distributions (these are the extreme points of the credal set)

What are the Credal Sets?

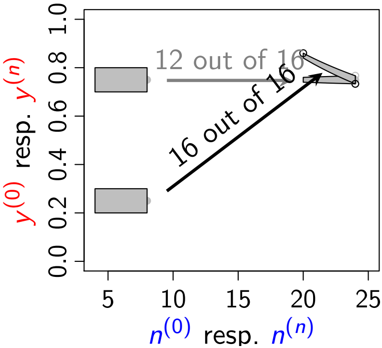
- credal set $\widehat{=}$ finite convex mixtures of parametric priors to ensure coherence of prior and posterior
- sets of distributions are induced by sets of parameters (parameter sets need not be convex)
- inferences should be linear in posterior distributions: then min/max are attained at the parametric distributions (these are the extreme points of the credal set)
- E, Var are linear in the parametric distributions

Update Step Properties



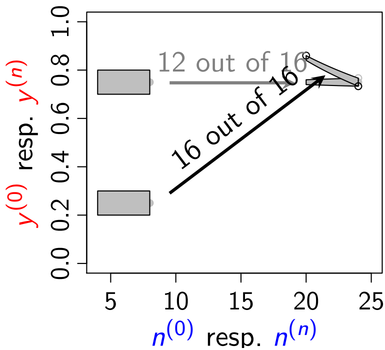
- additional imprecision in case of pdc due to “banana” shape

Update Step Properties



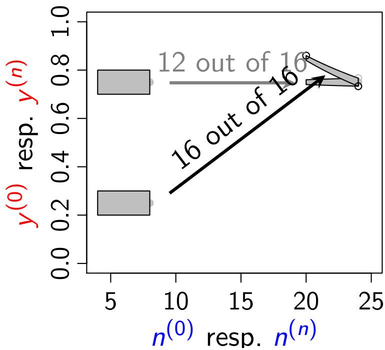
- additional imprecision in case of pdc due to "banana" shape
- $y^{(n)}$'s move faster for low $n^{(0)}$, and overtake $y^{(n)}$'s at high $n^{(0)}$ (flexible weights $n^{(0)}$ for the prior information encoded in $y^{(n)}$)

Update Step Properties



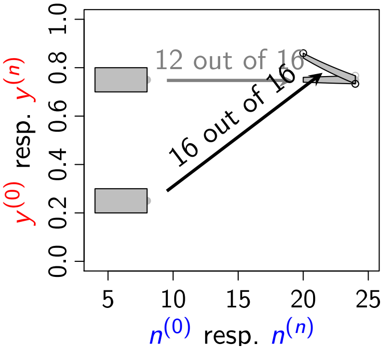
- additional imprecision in case of pdc due to “banana” shape
- $y^{(n)}$'s move faster for low $n^{(0)}$, and overtake $y^{(n)}$'s at high $n^{(0)}$ (flexible weights $n^{(0)}$ for the prior information encoded in $y^{(n)}$)
- $y^{(0)}$ stretch ▲ ⟷ $y^{(n)}$ stretch ▲

Update Step Properties



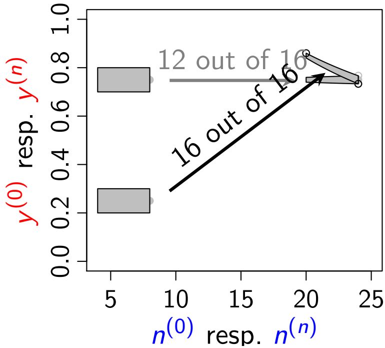
- additional imprecision in case of pdc due to "banana" shape
- $y^{(n)}$'s move faster for low $n^{(0)}$, and overtake $y^{(n)}$'s at high $n^{(0)}$ (flexible weights $n^{(0)}$ for the prior information encoded in $y^{(n)}$)
- $y^{(0)}$ stretch ▲ ⟷ $y^{(n)}$ stretch ▲
- $n^{(0)}$ ▲ ⟷ imprecision ▲ ($n^{(0)}$ governs size of posterior set!)

Update Step Properties



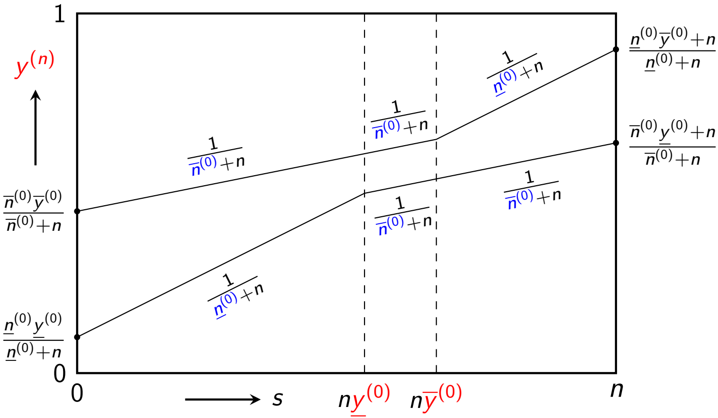
- additional imprecision in case of pdc due to "banana" shape
- $y^{(n)}$'s move faster for low $n^{(0)}$, and overtake $y^{(n)}$'s at high $n^{(0)}$ (flexible weights $n^{(0)}$ for the prior information encoded in $y^{(n)}$)
- $y^{(0)}$ stretch ▲ ⟷ $y^{(n)}$ stretch ▲
- $n^{(0)}$ ▲ ⟷ imprecision ▲ ($n^{(0)}$ governs size of posterior set!)
- $n^{(0)}$ stretch ▲ ⟷ stronger reaction to prior-data conflict

Update Step Properties



- additional imprecision in case of pdc due to "banana" shape
- $y^{(n)}$'s move faster for low $n^{(0)}$, and overtake $y^{(n)}$'s at high $n^{(0)}$ (flexible weights $n^{(0)}$ for the prior information encoded in $y^{(n)}$)
- $y^{(0)}$ stretch ▲ ⟷ $y^{(n)}$ stretch ▲
- $n^{(0)}$ ▲ ⟷ imprecision ▲ ($n^{(0)}$ governs size of posterior set!)
- $n^{(0)}$ stretch ▲ ⟷ stronger reaction to prior-data conflict
- contrary effect of $n^{(0)}$ on imprecision and variance: high $n^{(0)}$ gives a highly imprecise posterior set, but with low variance distributions within the set

Update Step Properties



Parameter Set Shapes

- rectangular prior set $[\underline{n}^{(0)}, \overline{n}^{(0)}] \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$ adds distributions with higher variance to the prior/posterior credal set as compared to the imprecise BBM ($\overline{n}^{(0)} \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$)

Parameter Set Shapes

- rectangular prior set $[\underline{n}^{(0)}, \overline{n}^{(0)}] \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$ adds distributions with higher variance to the prior/posterior credal set as compared to the imprecise BBM $(\overline{n}^{(0)} \times [\underline{y}^{(0)}, \overline{y}^{(0)}])$
- rectangular prior set (two-dimensional interval) seems natural, but for the model, generally any shape is possible

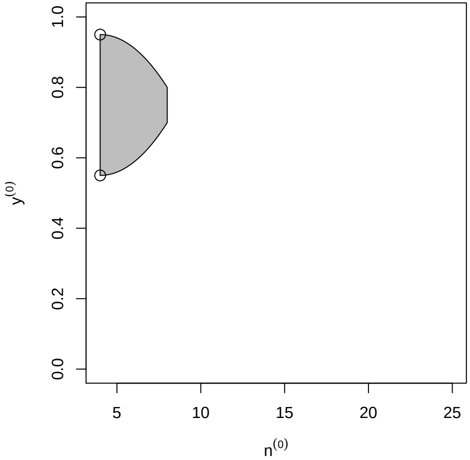
Parameter Set Shapes

- rectangular prior set $[\underline{n}^{(0)}, \overline{n}^{(0)}] \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$ adds distributions with higher variance to the prior/posterior credal set as compared to the imprecise BBM ($\overline{n}^{(0)} \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$)
- rectangular prior set (two-dimensional interval) seems natural, but for the model, generally any shape is possible
- posterior parameter sets are not rectangular anyway (but can be fully described by the four 'corners')

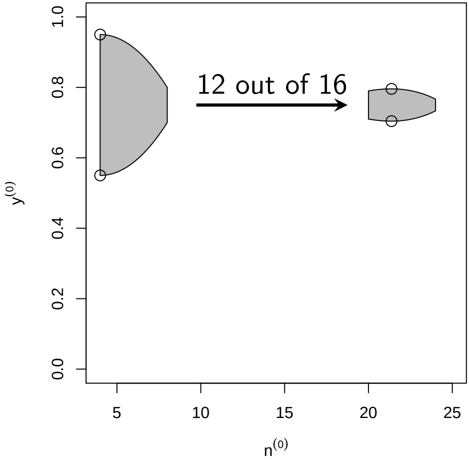
Parameter Set Shapes

- rectangular prior set $[\underline{n}^{(0)}, \overline{n}^{(0)}] \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$ adds distributions with higher variance to the prior/posterior credal set as compared to the imprecise BBM ($\overline{n}^{(0)} \times [\underline{y}^{(0)}, \overline{y}^{(0)}]$)
- rectangular prior set (two-dimensional interval) seems natural, but for the model, generally any shape is possible
- posterior parameter sets are not rectangular anyway (but can be fully described by the four 'corners')
- actual prior shape influences the posterior inferences (position of max / min $y^{(n)}$, other objective functions)

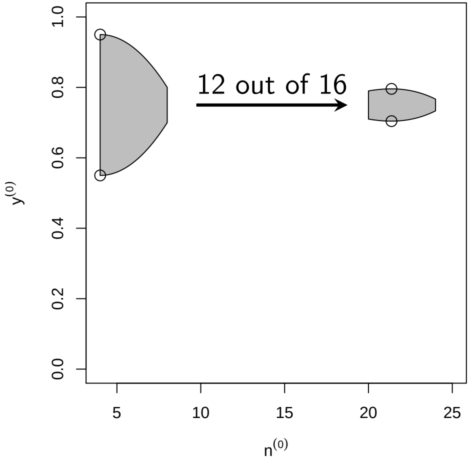
Parameter Set Shapes



Parameter Set Shapes

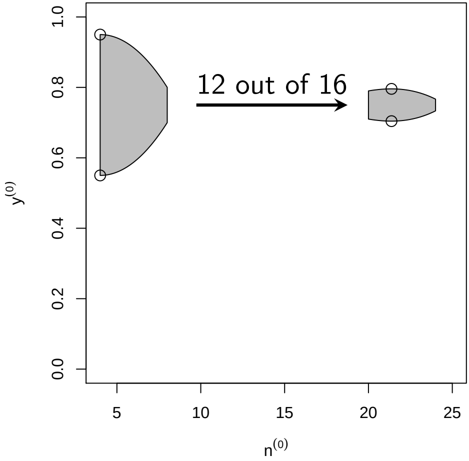


Parameter Set Shapes



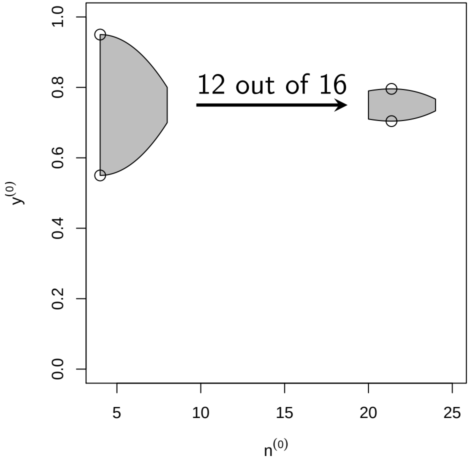
- the shape matters!

Parameter Set Shapes



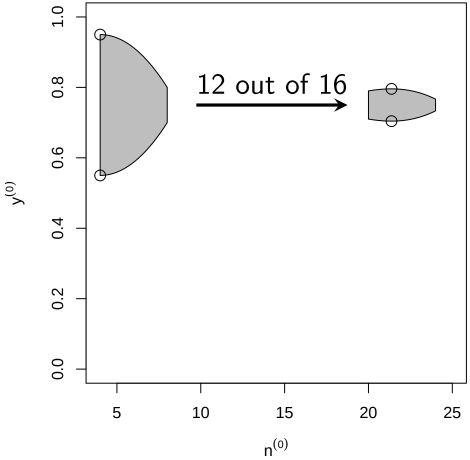
- the shape matters!
- shape could be tailored to enable desired inference properties

Parameter Set Shapes



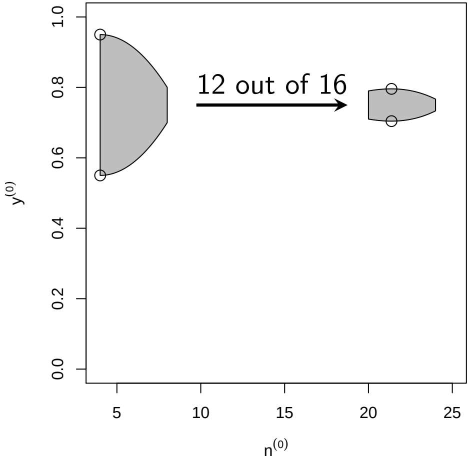
- the shape matters!
- shape could be tailored to enable desired inference properties
- however, difficult to elicit

Parameter Set Shapes



- the shape matters!
- shape could be tailored to enable desired inference properties
- however, difficult to elicit
- shape updating is quite difficult to grasp

Parameter Set Shapes



- the shape matters!
- shape could be tailored to enable desired inference properties
- however, difficult to elicit
- shape updating is quite difficult to grasp
- shape that has easy description for posterior set? (set description that's invariant under updating?)

Strong Prior-Data Agreement

- desired inference property:
  bonus precision if prior and data agree especially well

Strong Prior-Data Agreement

- desired inference property:
  bonus precision if prior and data agree especially well
- ideally also with high $n^{(n)}$ values at max / min $y^{(n)}$ (small variances!)

Strong Prior-Data Agreement

- desired inference property:
  bonus precision if prior and data agree especially well
- ideally also with high $n^{(n)}$ values at max / min $y^{(n)}$
  (small variances!)
- deus ex machina: Mik Bickis & his approach to imprec. BBM:
  change parametrization $(n^{(0)}, y^{(0)})$ to (η_0, η_1) (WPMSIIP'11)
  ➡ the parameter sets do not change shape during updating!

Mik's Parametrization

- $(\eta_0, \eta_1) \in \mathbb{R}^2$ such that $\eta_0 > 2$, $-1 - \frac{1}{2}\eta_0 < \eta_1 < 1 + \frac{1}{2}\eta_0$
- $\eta_0 = n^{(0)} - 2$
- Updating: win $\hat{=} + \left(1, \frac{1}{2}\right)$, not win $\hat{=} + \left(1, -\frac{1}{2}\right)$
- Distributions with the same expectation ($=$ P) lie along equidistant rays emanating from $(-2, 0)$

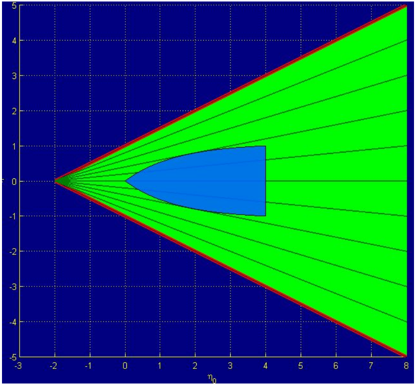
The Boat Shape

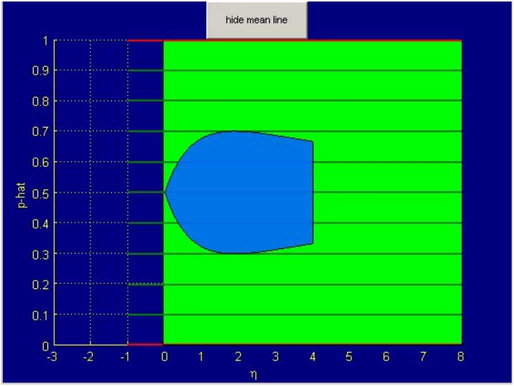
Our suggestion for a (η_0, η_1) shape that leads to

- additional imprecision in case of prior-data conflict
- bonus precision in case of strong prior-data agreement

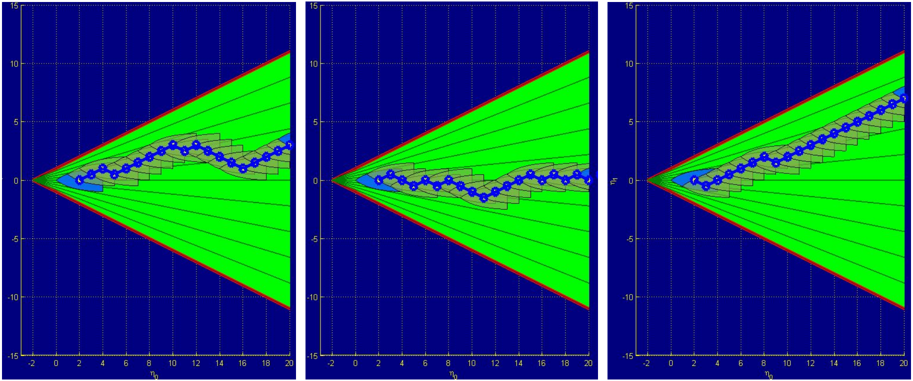
looks like a bullet, or a boat with a so-called transom stern.

The Boat Shape

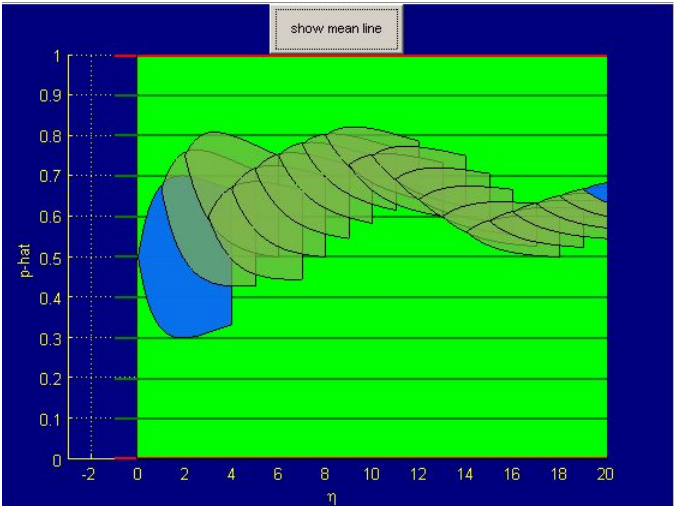




The Boat Shape

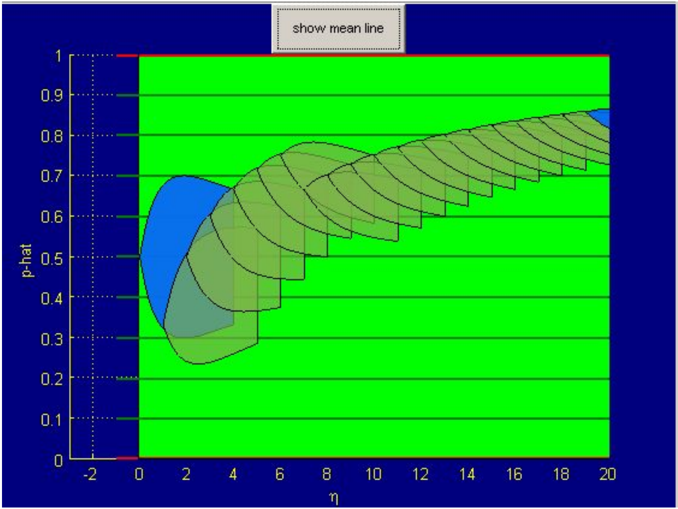


The Boat Shape

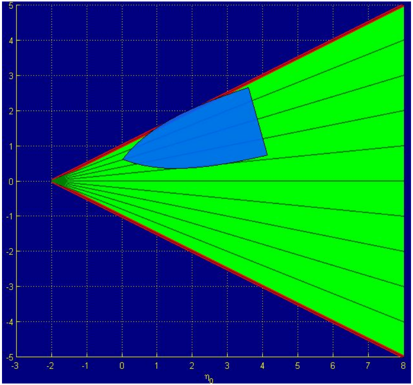


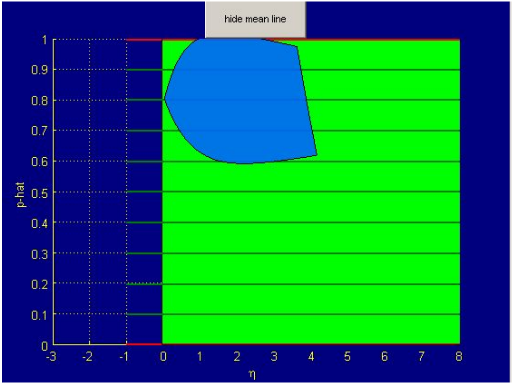
The Boat Shape

The Boat Shape

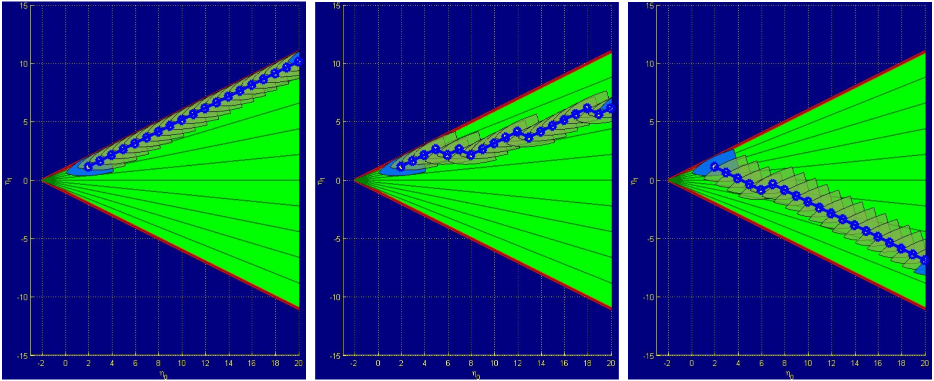


The Boat Shape

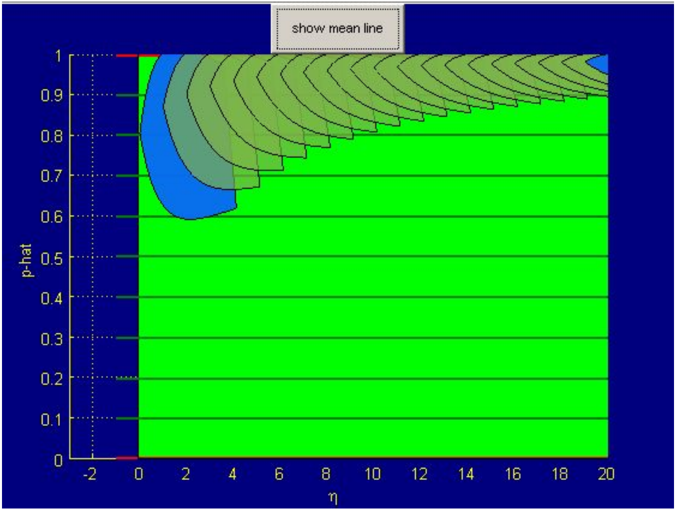




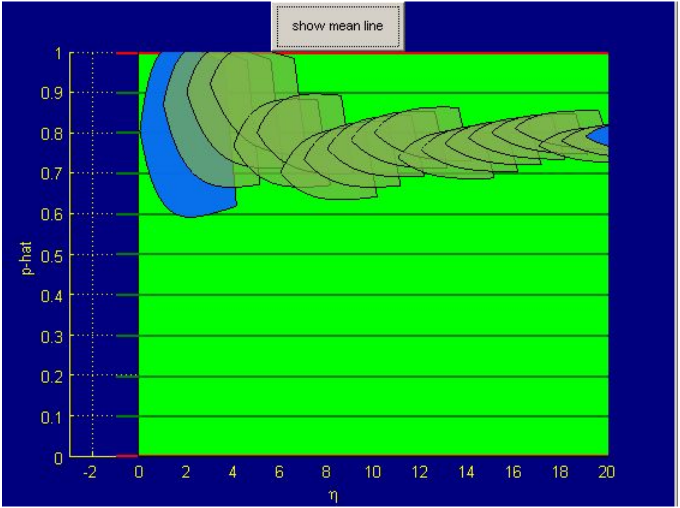
The Boat Shape



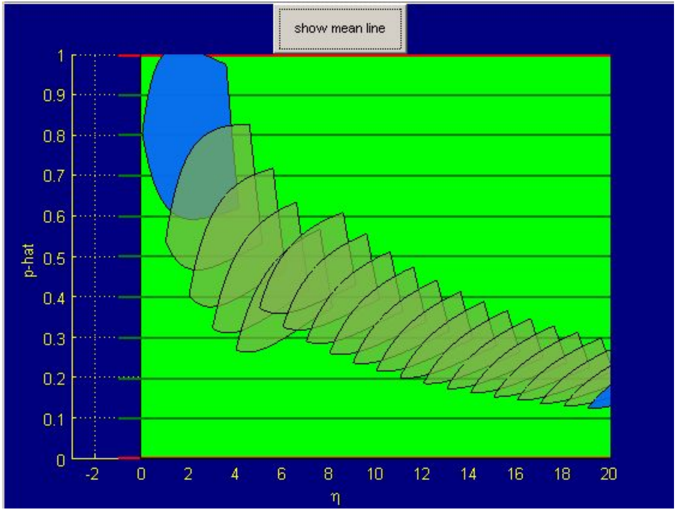
The Boat Shape



The Boat Shape



The Boat Shape



The Boat Shape

- interactive graph by Mik Bickis
- contours are exponential curves (mirrored at central ray)
- "touching rays" / "shadow" must be determined numerically
- detailed properties still to be explored
- suggestions?