Imprecise probability models and prior-data conflict

Gero Walter

Department of Statistics
University of Munich

11.12.2007

prior-data conflict

LUCK-models

$$p(\vartheta) \propto \exp\left\{ n^{(0)} \left[\langle \psi, y^{(0)} \rangle - \mathbf{b}(\psi) \right] \right\}$$

and

$$p(\vartheta \mid w) \propto \exp\left\{ n^{(1)} \left[\langle \psi, y^{(1)} \rangle - \mathbf{b}(\psi) \right] \right\},$$

with

$$n^{(1)} = n^{(0)} + q \qquad \text{and} \qquad y^{(1)} = \frac{n^{(0)} y^{(0)} + \tau(w)}{n^{(0)} + q}.$$

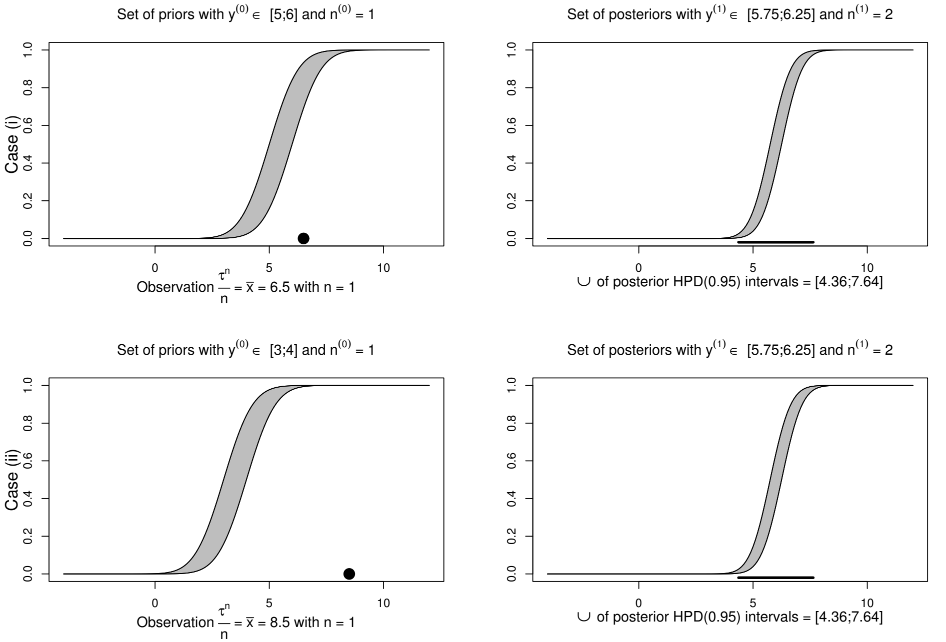
imprecise priors

Just like in the IDM: set $\mathcal{Y}^{(0)} \subset \mathcal{Y}$ and $n^{(0)}$ fix.

$$\mathcal{Y}^{(1)} = \left\{ \frac{n^{(0)} y^{(0)} + \tau(w)}{n^{(0)} + n} \;\middle|\; y^{(0)} \in \mathcal{Y}^{(0)} \right\} \subset \mathcal{Y}\,.$$

$\mathcal{Y}^{(1)}$ can actually be seen as a shifted and rescaled version of $\mathcal{Y}^{(0)}$:

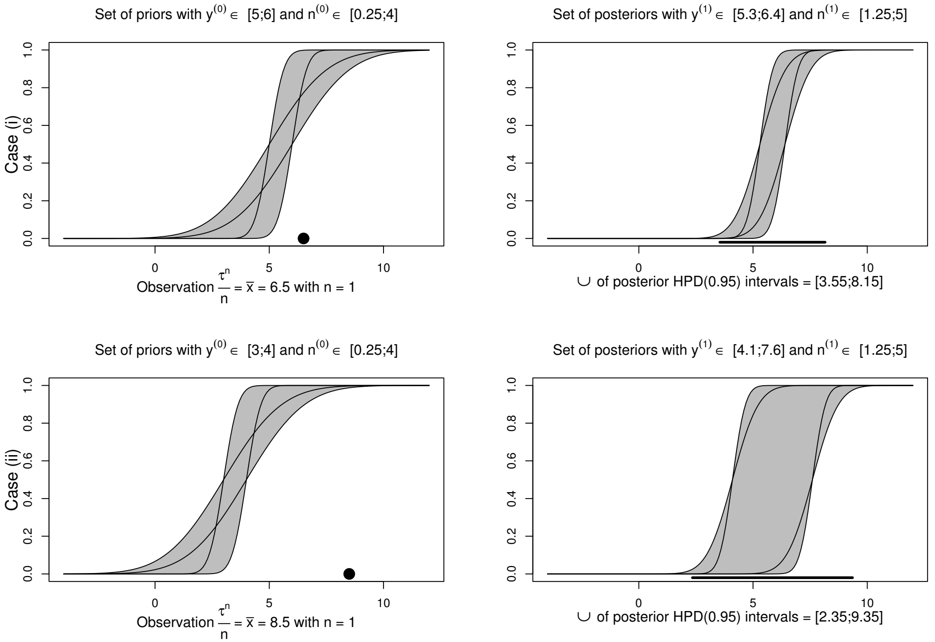
$$\mathcal{Y}^{(1)} = \frac{n^{(0)}}{n^{(0)} + n} \cdot \mathcal{Y}^{(0)} + \frac{n}{n^{(0)} + n} \cdot \frac{1}{n} \sum_{i=1}^{n} \tau(w_i)\,,$$


Set of priors with y(0) ∈ [5;6] and n(0) = 1
Case (i)
0.0
0.2
0.4
0.6
0.8
1.0
0
5
10
Observation τn/n = x̄ = 6.5 with n = 1
Set of posteriors with y(1) ∈ [5.75;6.25] and n(1) = 2
∪ of posterior HPD(0.95) intervals = [4.36;7.64]
Set of priors with y(0) ∈ [3;4] and n(0) = 1
Case (ii)
Observation τn/n = x̄ = 8.5 with n = 1
Set of posteriors with y(1) ∈ [5.75;6.25] and n(1) = 2
∪ of posterior HPD(0.95) intervals = [4.36;7.64]

Why that?

$$\overline{y}^{(1)} - \underline{y}^{(1)} \quad = \quad \frac{n^{(0)} \left(\overline{y}^{(0)} - \underline{y}^{(0)}\right)}{n^{(0)} + n} \quad \text{does not depend on } \tau^n(x)!$$

Idea: vary $n^{(0)}$ in some set $\mathcal{N}^{(0)}$ as well!


Set of priors with y(0) ∈ [5;6] and n(0) ∈ [0.25;4]
Case (i)
Observation τn/n = x̄ = 6.5 with n = 1
Set of posteriors with y(1) ∈ [5.3;6.4] and n(1) ∈ [1.25;5]
∪ of posterior HPD(0.95) intervals = [3.55;8.15]
Set of priors with y(0) ∈ [3;4] and n(0) ∈ [0.25;4]
Case (ii)
Observation τn/n = x̄ = 8.5 with n = 1
Set of posteriors with y(1) ∈ [4.1;7.6] and n(1) ∈ [1.25;5]
∪ of posterior HPD(0.95) intervals = [2.35;9.35]

Why does it work?

$$\underline{y}^{(1)} = \begin{cases} \dfrac{\overline{n}^{(0)}\underline{y}^{(0)} + \tau^n(x)}{\overline{n}^{(0)} + n} & \text{if} \quad \dfrac{\tau^n(x)}{n} \geq \underline{y}^{(0)} \\[2ex] \dfrac{\underline{n}^{(0)}\underline{y}^{(0)} + \tau^n(x)}{\underline{n}^{(0)} + n} & \text{if} \quad \dfrac{\tau^n(x)}{n} < \underline{y}^{(0)} \quad \iff \quad \text{prior-data conflict} \end{cases}$$

$$\overline{y}^{(1)} = \begin{cases} \dfrac{\overline{n}^{(0)}\overline{y}^{(0)} + \tau^n(x)}{\overline{n}^{(0)} + n} & \text{if} \quad \dfrac{\tau^n(x)}{n} \leq \overline{y}^{(0)} \\[2ex] \dfrac{\underline{n}^{(0)}\overline{y}^{(0)} + \tau^n(x)}{\underline{n}^{(0)} + n} & \text{if} \quad \dfrac{\tau^n(x)}{n} > \overline{y}^{(0)} \quad \iff \quad \text{prior-data conflict} \end{cases}$$

Now:

$$\overline{y}^{(1)} - \underline{y}^{(1)} = \frac{\overline{n}^{(0)}(\overline{y}^{(0)} - \underline{y}^{(0)})}{\overline{n}^{(0)} + n} + \Delta\left(\frac{\tau^n(x)}{n}; \underline{y}^{(0)}, \overline{y}^{(0)}\right) \frac{n(\overline{n}^{(0)} - \underline{n}^{(0)})}{(\overline{n}^{(0)} + n)(\underline{n}^{(0)} + n)},$$

where

$$\Delta\left(\frac{\tau^n(x)}{n}; \underline{y}^{(0)}, \overline{y}^{(0)}\right) = \inf\left\{\left|\frac{\tau^n(x)}{n} - y^{(0)}\right| : \underline{y}^{(0)} \le y^{(0)} \le \overline{y}^{(0)}\right\}$$

is the distance between the prior interval $\left[\underline{y}^{(0)}; \overline{y}^{(0)}\right]$ and the observation $\frac{\tau^n(x)}{n}$.

Why then not choose $\sup \mathcal{N}^{(0)} = n^{(0)}$?

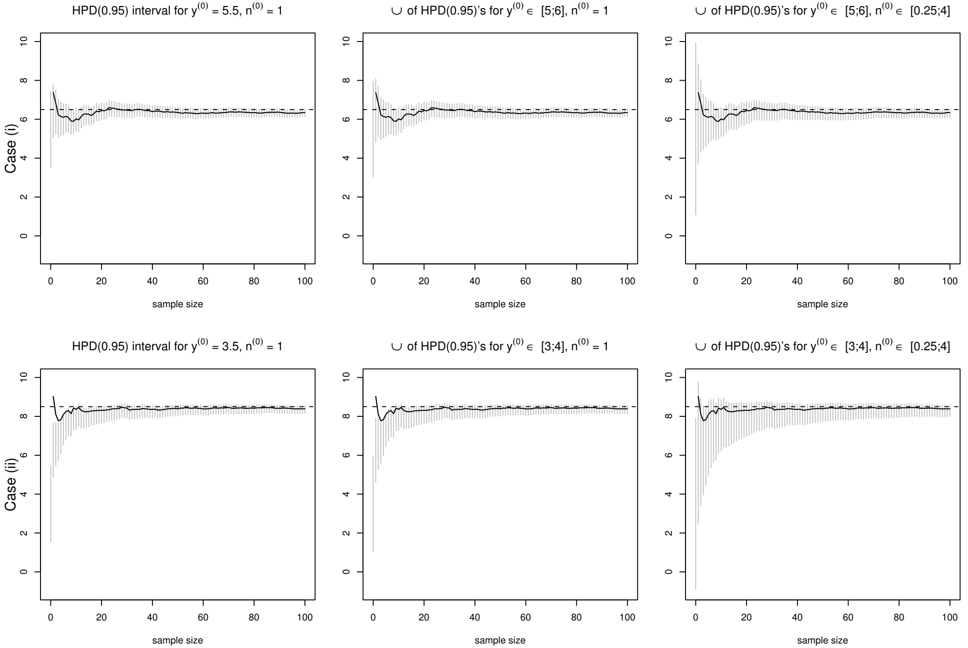
The factor to $\Delta(\ )$ gets maximal if $n = \sqrt{\underline{n}^{(0)}\overline{n}^{(0)}}$.

$\iff$ $\overline{y}^{(1)} - \underline{y}^{(1)}$ maximal for fixed $\Delta(\ )$.

$\iff$ Same weight on the prior and on the sample.

$\iff$ $\sqrt{\underline{n}^{(0)}\overline{n}^{(0)}}$ is the 'global' prior strength to be compared to $n^{(0)}$.

For $n \longrightarrow \infty$, it doesn't matter, as $\dfrac{\overline{n}^{(1)} - \underline{n}^{(1)}}{\overline{n}^{(1)}} \longrightarrow 0$:


HPD(0.95) interval for $y^{(0)} = 5.5$, $n^{(0)} = 1$
∪ of HPD(0.95)'s for $y^{(0)} \in [5;6]$, $n^{(0)} = 1$
∪ of HPD(0.95)'s for $y^{(0)} \in [5;6]$, $n^{(0)} \in [0.25;4]$
Case (i)
HPD(0.95) interval for $y^{(0)} = 3.5$, $n^{(0)} = 1$
∪ of HPD(0.95)'s for $y^{(0)} \in [3;4]$, $n^{(0)} = 1$
∪ of HPD(0.95)'s for $y^{(0)} \in [3;4]$, $n^{(0)} \in [0.25;4]$
Case (ii)
sample size

$\underline{\mathrm{P}}$ and $\overline{\mathrm{P}}$ for classical confidence intervals

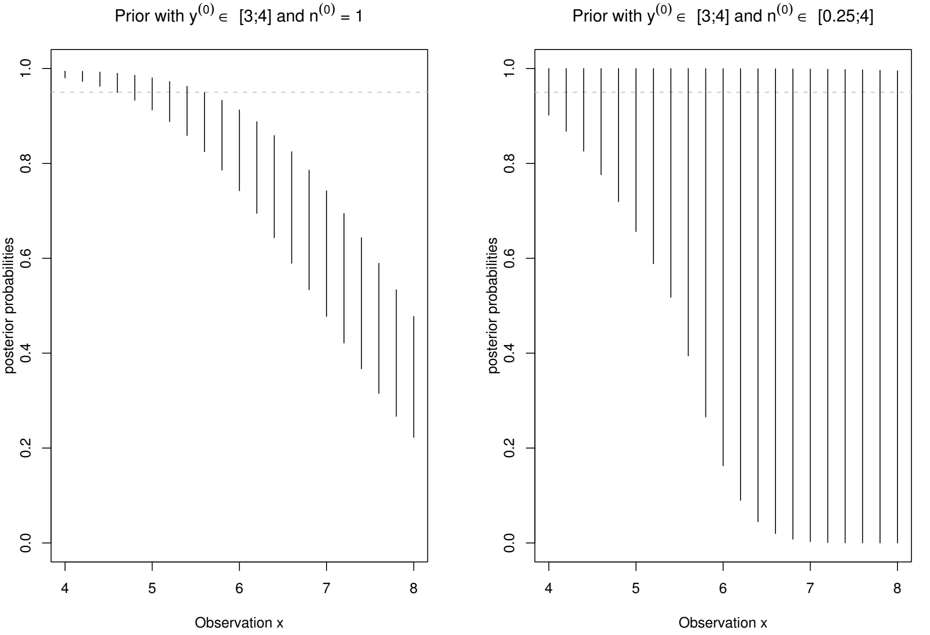
$$\mathrm{CI} = \bar{x} \pm z_{\frac{1+\gamma}{2}} \frac{1}{n} \qquad \text{(we have } \sigma^2 = 1\text{)}$$

$$\left[\underline{\mathrm{P}}(\{\mu \in \mathrm{CI}\}); \overline{\mathrm{P}}(\{\mu \in \mathrm{CI}\})\right]$$

should tend to vacuous the more severe the prior-data conflict is.

Probability weight of area around $\bar{x}$ should be very insecure:

- if data 'is right', CI should have high probability weight
- if prior 'is right', CI should have low probability weight


Prior with y(0) ∈ [3;4] and n(0) = 1
posterior probabilities
0.0
0.2
0.4
0.6
0.8
1.0
4
5
6
7
8
Observation x
Prior with y(0) ∈ [3;4] and n(0) ∈ [0.25;4]
posterior probabilities
0.0
0.2
0.4
0.6
0.8
1.0
4
5
6
7
8
Observation x

dataset: Mietspiegel

Summary