Using historical data for Bayesian sample size determination

Author: Fulvio De Santis,

J. R. Statist. Soc. A (2007) **170**, *Part* 1, *pp.* 95–113

Harvard Catalyst Journal Club: December 7th 2016

Kush Kapur, PhD

- **Inferential** and **decision theoretic (not discussed)**

- Bayesian statistician –

  a. Sequential procedure

  b. Bayesian philosophy (antecedents: prior and likelihood and hence in principle no need for preplanned sample sizes?)

What does computing sample size mean in Bayesian context?

Secondary: What computational challenges do we face in the implementation?

- Frequentist Approach (Neyman-Pearson) - local optimality problem

Hybrid Frequentist-Bayesian method –

- Prior predictive distribution with prior incorporated (marginal density of data obtained by integrating out the prior) – Sample size

- Posterior predictive distribution with non-informative priors – Inference

Is this approach justified?

Sample size obtained based on controlling some aspect of the posterior distribution

- Average posterior variance is smaller than some value

- Average length criterion (ALC) :

a. The main idea is that the coverage probability (1-alpha) is fixed and the HPD interval length varies depending on the sample.
b. N is chosen to be min integer which satisfies average width less than equal to L.

- Average coverage criterion (ACC):

a. Similar approach, hold the HPD length L as fixed and allow the coverage probability (1-alpha) to vary with data.
b. N is chosen to obtain average coverage over the data equal to (1-alpha).

The ACC and ALC are based on averages over all possible samples.

Inferences are conditional on the observed sample - result in coverage or length small or large

- Worst outcome criterion (WOC): Conservative approach ensures a maximum length L and a minimum coverage probability (1-alpha) regardless of the data.

- However, this approach fails to incorporate the marginal distribution of the historical data and hence the prior!!

- All the aforementioned approaches - Where is the concept of Power??

- Prior specification ⇔ Alternate hypothesis

- Controlling the width of the posterior interval of variance ⇔ Type I error rate

Unconditional Power approximately Classical power (Conditional power) integrated over by the prior

Power Priors

- Defined Hierarchically by combining the prior and the likelihood of historical data
- The likelihood is scaled by the exponential (or discounting) factor “a_0”

- Initial prior assumed non-informative
- Discounting factor is assumed to be deterministic in this manuscript

Power Priors

“It can be interpreted as a posterior distribution that is associated with a sample whose informative content on *the parameter* is *qualitatively* the same as that of the historical data set, but *quantitatively* equivalent to that of a sample of size $r = a_0 n_0$.”

a_0	*Discount*
<0.2	Severe
0.2–0.5	Substantial
0.5–0.8	Moderate
>0.8	Slight

Power Priors

- Except for few standard distribution, the implementation is simulation based

- Computing resources are becoming cheaper!!
- But iterative steps during the design stage can require some pre-planning
- Implementation might not be always easy for different criteria (see Wang and Gelfand, 2002)

Examples

- *Sample size for inference on the normal mean*

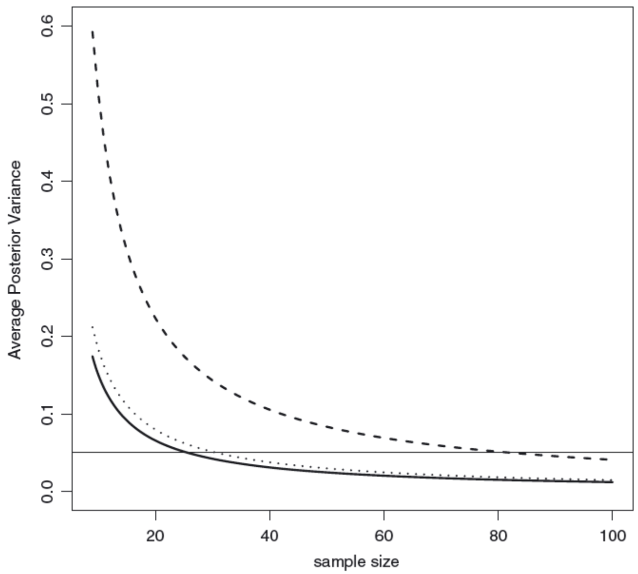

Fig. 1. Average posterior variance criterion for the normal mean, $n_0 = 20$ and $s_0 = 1$, using the power prior with $a_0 = 1$ (———), $a_0 = 0.5$ (·······) and $a_0 = 0.2$ (– – –)

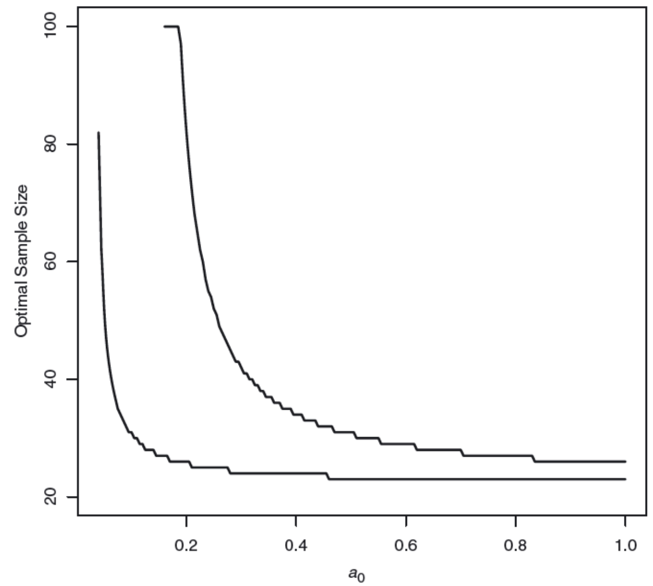

Fig. 2. Optimal sample sizes from the average posterior variance criterion for the normal mean, $s_0 = 1$, using the power prior, as a_0 varies in (0,1), $n_0 = 20$ (upper curve) and $n_0 = 100$ (lower curve)

Examples

- *Sample size for inference on the exponential parameter*

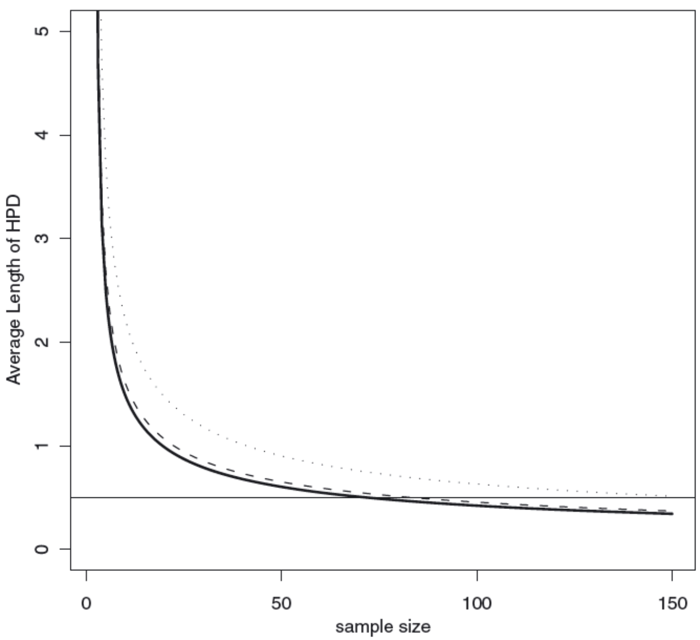

Fig. 3. ALC for the normal mean, $n_0 = 20$, $s_0 = 1$, and $\alpha = 0.05$, using the power prior with $a_0 = 1$ (———), $a_0 = 0.2$ (·······) and $a_0 = 0.5$ (– – –)

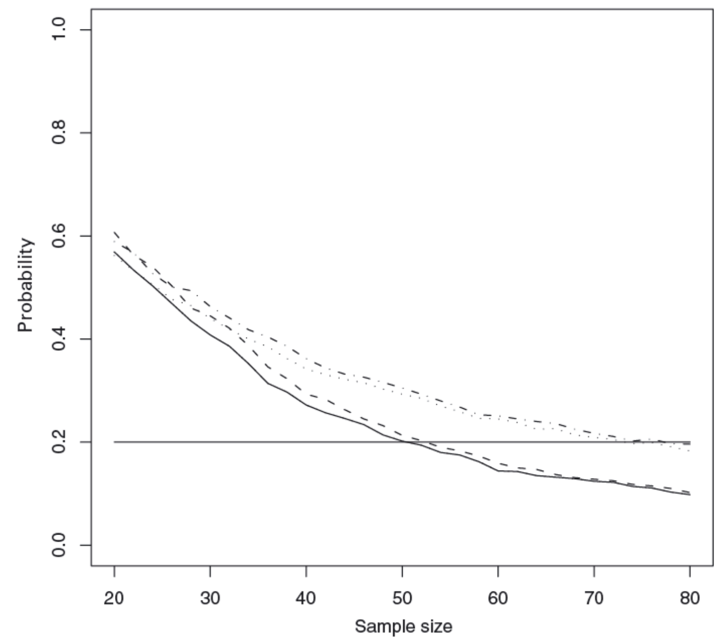

Fig. 4. Survival analysis example: $\Pr\{L_\alpha(\mathbf{X}_n) \geqslant 30\}$ as a function of n when $a_0 = 1$ (———, HPD; – – –, equal tails) and when $a_0 = 0.5$ (·······, HPD; ·–·–·, equal tails)

Examples

- *Sample size for inference on a proportion*

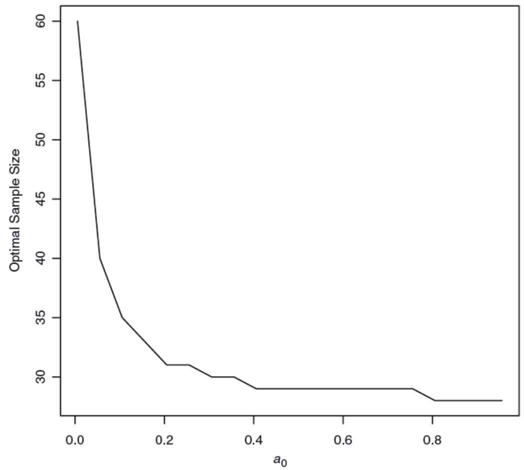

Fig. 5. Optimal sample size as a function of a_0, using ALC for 95% equal-tails intervals of the Bernoulli parameter ($n_0 = 176$ and $s_0 = 12$)

- **Multiple-power priors: sample size determination for proportions**
- **What about multiple historical trials providing opposite results?**
- **How do we choose the discounting factors?**
- **Should we build in budget constraint in these procedures?**

- Does not depend on specification of alternate hypothesis (could be an overenthusiastic belief on the parameter of interest)
- Includes specification of Prior (use of historical information)
- Easily accommodates composite hypothesis (for example treatment effect > 0 vs <0)

- Power priors can be sensitive to the discounting factor
- Distribution of the discounting factor can lead another layer of complexity
- Power priors cannot be used for inference
- Implementation not so easy every time

Spiegelhalter (2004, Bayesian approaches to clinical trials and health care evaluation) –

"It could be argued that **elicitation of prior beliefs and demands from abroad community of stakeholders** is necessary not only in order to undertake a specifically Bayesian approach to design and analysis, **but also more generally as part of good research practice**."

Any thoughts?