Assessment of Bayesian methods for constructing prior distributions (MAP & Power prior)

Context: Evaluating the superiority of a treatment compared to a control group in "Proof-of-Concept" study (early phase clinical trial) for Normal-Gamma distribution case

Léa POULMARCH

Tutor: Karine FLORIN

Journée Nationale Biopharmacie et Santé

SFdS – November 28th 2016

Content

Contents

Introduction

- Context: « Proof-of-concept » PoC study
  - **Failures of development in late phase (IIb-III) influence R&D costs**
  - **Need a proof of concept as soon as possible (phase IIa)**
    - ⇨ Decision-making (Go/No Go)
- Study: Superiority of treatment compared to control group, parallel design
- Historical data from different sources: publications, intern studies, expert advice

- Improve the probability of success to make the right decision

- Evaluation of Bayesian methods for constructing prior distributions from historical data in case of PoC studies compared to frequentist analysis and classical Bayesian analysis
  - **Normal-gamma distribution case**
  - **Implemented using SAS 9.4**

Contents

Different methods

- **Meta-Analytic-Predictive (MAP) Prior:** Based on hierarchical models
  [1] Neuenswchander, 2010

- **Power prior:** Based on weighting the likelihood [4] Ibrahim *et al.*, 2015

- **Methods common principle**
  - **First Step:** Constructing the prior distribution
    - Summarize information from historical data $(\boldsymbol{Y_1}, \dots, \boldsymbol{Y_H})$
  - **Second Step:** Use this distribution as prior with the new data $\boldsymbol{Y}^*$

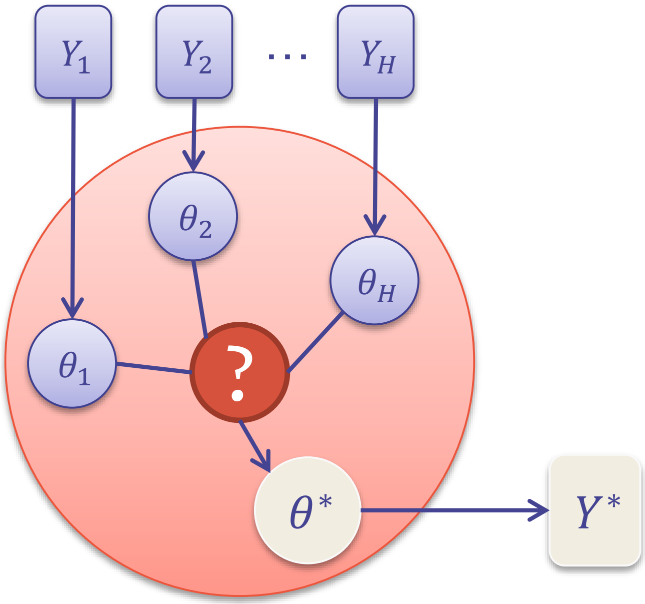
    $$\pi(\theta^*|\boldsymbol{Y}^*) \propto \pi(\boldsymbol{Y}^*|\theta^*) \boxed{\pi(\theta^*|\boldsymbol{Y_1}, \dots, \boldsymbol{Y_H})}$$
  - **Third Step:** Use the posterior distribution of the parameter θ^*
    - Compute the probability from the posterior distribution e.g. $\boldsymbol{P}(\theta_{\boldsymbol{T}} - \theta_{\boldsymbol{C}} < \boldsymbol{0})$

Meta-Analytic-Predictive (MAP) prior

- **First step:** Constructing the prior distribution
  - A standard random-effect meta-analysis of historical data (hierarchical model)

$$\forall\, i = 1, \dots, H \text{ and } \forall\, j = 1, \dots, n_i$$
$$Y_{ij} | \theta_i \sim \mathcal{F}(\theta_i)$$
$$\theta_i | \phi, \tau^2 \sim N(\phi, \tau^2)$$
$$\phi \sim M \qquad \tau^2 \sim V$$

e.g. $\phi \sim N(0, 10^{-6})$ and $\tau^2 \sim IG(10^{-3}; 10^{-3})$



  - Generate the predictive distribution of θ^*
    - Several methods exist
    - Choice: an **normal approximation** of the the prediction distribution is used as in Walley *et al.*, 2016 [11]

- **Second step:** Use predictive distribution of θ^* as prior distribution with the new data $\boldsymbol{Y}^*$

SANOFI

Robust MAP prior

- **First step:** Use of the distribution $\pi_{MAP}(\theta^*|Y_1, \ldots, Y_H)$ obtained previously, to which is added a non-informative distribution

$$\pi_{Robust\ MAP}(\theta^*|Y_1, \ldots, Y_H) = w \times \pi_{MAP}(\theta^*|Y_1, \ldots, Y_H) + (1 - w) \times \pi_Z(\theta^*)$$

  - For example: a distribution $Z \sim N(0,10^6)$ for normal case can be taken
  - w represents the probability that the new trial will be similar to the historical data

- **Second step:** Exactly the same as for the classical MAP prior

$$\pi(\theta^*|Y^*) \propto \pi(Y^*|\theta^*)\pi_{Robust\ MAP}(\theta^*|Y_1, \ldots, Y_H)$$

- Drawback of the robust MAP prior: arbitrary choice of the w parameter
  - **0.9 often used in studies** [3] Schmidli *et al.*, 2014
  - **Large heterogeneity between studies: 0.5 is chosen**

Power Prior

- **First step: Constructing the prior distribution**
  - **Introduction of the a_0 parameter** $0 < a_0 < 1$
  - **Application of the Bayes formula on historical data** $\pi_{Power}(\theta|Y_1, \dots, Y_{\mathrm{H}}) \propto \pi(Y_1, \dots, Y_{\mathrm{H}}|\theta)^{a_0}\pi(\theta)$
    - $a_0 = 1 \rightarrow$ Classical Bayes formula
    - $a_0 = 0 \rightarrow$ No historical data
    - $\pi(\theta) \rightarrow$ Non informative prior

- **Second step: Use the Power prior**
  - **Use the posterior distribution on historical data as a prior with the PoC data**

    $$\pi(\theta^*|Y^*) \propto \pi(Y^*|\theta^*)\pi_{Power}(\theta^*|Y_1, \dots, Y_H)$$

  - **Use the posterior distribution of the interest parameter** θ^* **to calculate the probability** $P(\theta_T^* - \theta_C^* < 0)$

Power prior/ Pros & Cons

- Pros
  - **Possible weighting of the historical data with the a_0 parameter**

- Cons
  - **Arbitrary choice of a_0**
  - **If prior on a_0, violation of the likelihood principle**
    - If we multiply the likelihood by a constant K: The prior on a_0 is modified and becomes $K^{a_0} \times \pi(a_0)$, which changes estimations

Normalized power prior

- Solution: the normalized Power prior

$$\pi_{Normalized\ Power}(\theta|Y_1,\dots,Y_{\mathrm{H}}) \propto \frac{\pi(Y_1,\dots,Y_{\mathrm{H}}|\theta)^{a_0}\pi(\theta)}{\int \pi(Y_1,\dots,Y_{\mathrm{H}}|\theta)^{a_0}\pi(\theta)d\theta}\pi(a_0)$$

  - **The K^{a_0} value disappears**
  - **There is still the choice of the prior for a_0**

Prior Effective Sample Size (ESS)

- Definition
  - **ESS is a hypothetical sample size brought by the prior, in terms of patients added by borrowing historical data**
  - **Distance between the posterior with non informative prior and the informative prior**

- Approximations of prior ESS
  - **Morita *et al.*, 2008 [16]**
    - Normal case: Ratio of the known variance in the likelihood ($\boldsymbol{\sigma^2}$) to the prior variance of θ^* ($\boldsymbol{b^2}$)

$$\boldsymbol{ESS = \frac{\sigma^2}{b^2}}$$

Contents

Simulated dataset context

- Purpose: Comparison of methods and evaluate the operating characteristics (frequentist properties as the type I error and the power) of Bayesian methods (recommended by FDA)

- Classical project
  - **Primary endpoint: normal distribution**
  - **Groups: CONTROL vs TREATMENT**
  - **Aim of PoC study: evaluate the superiority of TREATMENT arm compared to CONTROL arm**
    - Decision-making rule for the superiority: if **P(treatment – Control>0) > 0,95**
  - **Historical data**
    - **3 Standard publications**
    - **1 Internal study**

SANOFI

Simulated dataset context

Historical data		New study
Publications Control arm (mean and SD)	**Internal study** (TDR) Individual data	**PoC** Simulated data
Control Publication 1: Nc_1 mean_1 et sd_1 ⋮ Publication H: Nc_H mean_H et sd_H	**Control**: $\mathrm{Nc}_{\mathrm{H+1}}$ **Treatment**: $\mathrm{Nt}_{\mathrm{H+1}}$	**Control**: $\mathrm{Nc} = 26$ $N \sim (\mathrm{mean}_\mathrm{C}\ ;\ SD = \boldsymbol{\sigma})$ **Treatment** : $\mathrm{Nt} = 52$ $N \sim (\mathrm{mean}_\mathrm{C} + \boldsymbol{\Delta}\ ;\ SD = \boldsymbol{\sigma})$

Note : Assumption of variance homogeneity $\sigma^2 = \sigma_C^2 = \sigma_T^2$
1000 simulated datasets

- **Aim**: evaluate the impact of the difference in response between PoC study and historical data
  - **$\boldsymbol{\Delta}$ difference between the treatment and the control mean**
    - $\Delta = 0$ to control type I error and $\Delta = -2$ to control power
    - Also $\Delta = -3$ and $\Delta = -1$ to understand the power progress in extreme cases
  - **$\boldsymbol{\sigma}$ standard deviation of the treatment and the control arm:**
    - $\sigma = 2.5\ ;\ 3.5\ ;\ 7$

⟶ **8 scenarios**

Simulations carried out

- First, according to the methods: Impact evaluation of key parameters (with $\sigma = 3.5$)
  - **MAP Prior: choice of prior distribution for τ^2**
    - $HN(1) - HN(0.5) - IG(e,e)$ **avec** $e = 0.1 \quad 0.01$ **and** 0.001

    $\Longrightarrow$ Choice of $\boldsymbol{IG\ (0.001, 0.001)}$
      - **Best compromise (power, type I error)**
      - **Recommended by Viele et al. (2014)**
  - **Robust MAP Prior: impact evaluation of $\boldsymbol{w}$**
    - **0.3, 0.5, 0.7, 0.9**
      - **Close results**

    $\Longrightarrow$ Choice of $\boldsymbol{w = 0.5}$
  - **Power Prior: impact of a_0**
    - **Value of a_0 from 0 to 1 in 0.1 step**

    $\Longrightarrow$ Choice of $\boldsymbol{a_0 = 0.8}$
      - Best equilibrium between power and type I error
      - Estimate of $a_0 \approx 0.88$ from Normalized power prior method

SANOFI

Simulations carried out

Comparison of 7 methods

Bayesian method with non-informative or weakly informative prior
Equivalent to frequentist methods

1. Non-informative prior/ No historical data taken into account
2. Non-informative prior / Historical data taken into account
   Likelihood on Pool data (historical dataset + PoC)

Bayesian method with informative prior
Estimated parameters based on historical data

3. Classical informative prior (Elicitation)
4. MAP prior (prior for τ^2 chosen)
5. Robust MAP prior (fixed w)
6. Power prior (fixed a_0)
7. Normalized Power prior

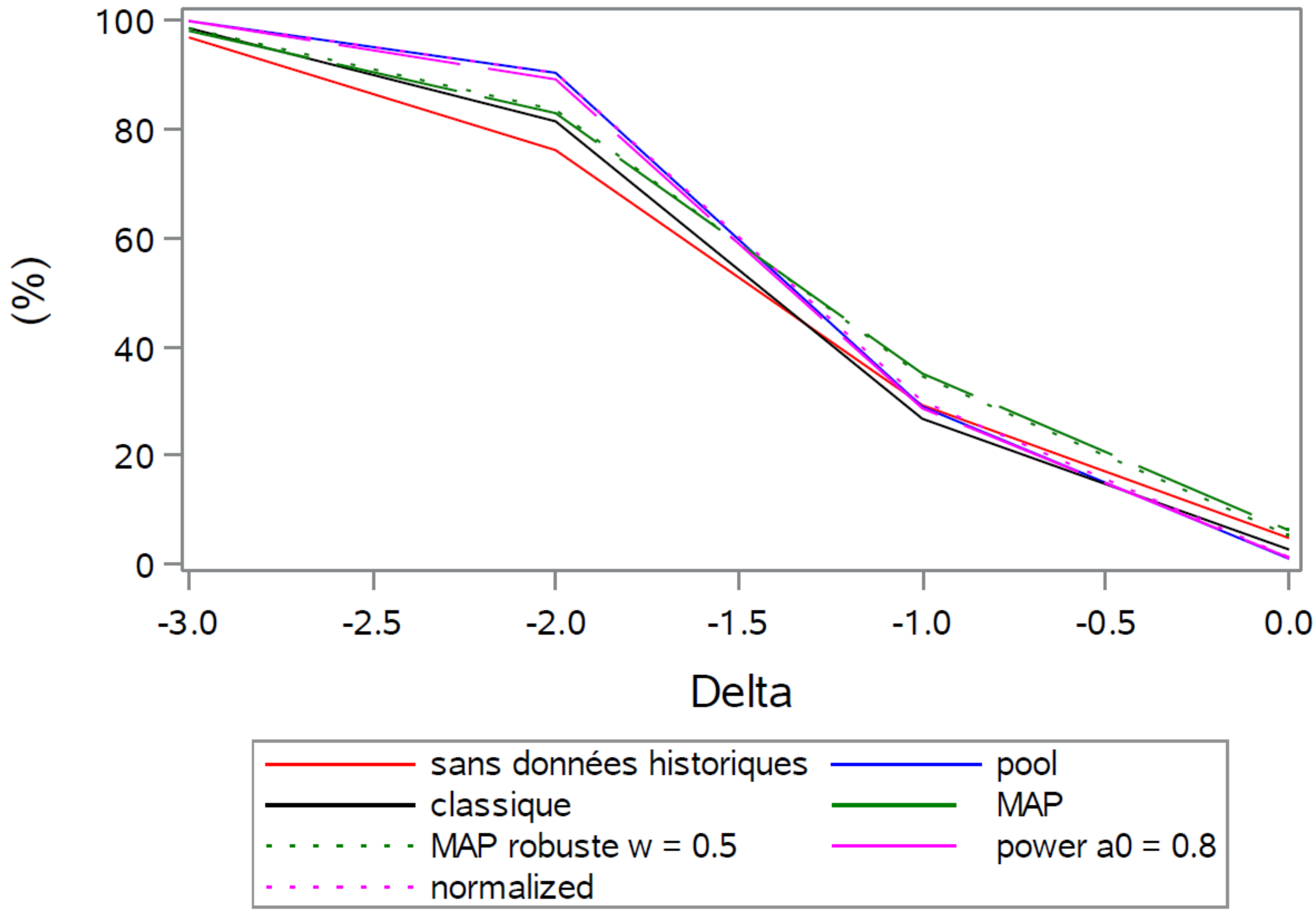
$\Longrightarrow$ **7 methods * 8 scenarios = 56 test cases**

Simulations Results

Method	Power			Type I error		
$\sigma =$	**2.5**	**3.5**	**7**	**2.5**	**3.5**	**7**
No historical data	95.4	76.1	29.6	4.5	4.9	4.8
Pool	99.2	90.3	48.7	0.2	1.1	4.0
Classical	93.7	81.4	46.5	0.4	2.8	7.3
MAP	95.5	82.9	48.5	3.3	6.3	10.3
Robust MAP prior $(w = 0.5)$	95.2	83.5	47.6	2.8	5.4	9.2
Power prior $(a_0 = 0.8)$	98.8	89.1	45.7	0.2	1.3	3.6
Normalized Power prior	99.4	90.3	46.5	0.9	1.4	3.6

MAP and robust MAP: prior for τ^2 : IG (0.001,0.001) / robust MAP: w=0.5
Power prior: a_0 =0.8
Pool: Bayesian analysis on pooled data (historical + POC) with a non-informative prior distribution e.g. $\mu_C \sim N(0; 10^6)$
No historical data: Bayesian analysis on PoC data with non informative prior
Classical: only on PoC study with a informative prior
Normalized: median of $a_0 = 0.88$
Number of simulation: 1000

Power



SANOFI

Contents

Conclusion
Simulations

- Good contribution of historical data
  - **Higher power**
- MAP prior and its Robust version
  - **Close results**
  - **But type I error slightly more controlled with robust MAP prior**
- Power prior and its Normalized version
  - **Expected: few differences between Power prior and Normalized Power prior**

- Comparison between MAP and Power prior
  - **Power prior (particularly Normalized) :**
    - Higher power, smaller type I error,
    - Higher impact of historical data on estimated Δ, RMSE lower
  - **MAP:**
    - Lower impact of historical data on estimated Δ and a better coverage when $\boldsymbol{\Delta} < -1$ and stable whatever $\boldsymbol{\Delta}$
    - Bad control of type I error with MAP prior compared to Power prior or Normalized

SANOFI

Conclusion
Recommendations

- Take into account that the frequentist properties have to be met (recommendations of FDA)
  - **In our context: the power prior is the best method to use**
    - If we have an idea of the power a_0: Power prior
    - Else: Normalized power prior

Remark: these recommendations are the same as the 2015 internship on binomial data

- Extension:
  - **More important between study variance on historical data**
  - **Going further : Classical analysis, Power prior and Normalized Power prior**
    - If we have an important number of publications
    - Impact of historical data sample size

Conclusion
Theoretical comparison

	Pros	Cons
MAP prior	Incorporation of the between study variance τ^2 Bayesian hierarchical model	Choice of prior on τ^2
Robust MAP prior	In case of conflict between historical data	Choice of w
Power prior	Weighting by the power a_0 No compute predictive distribution Easy to implement Different weighting according to the study/publication	Arbitrary choice of a_0 Non-respect of the principle likelihood $\Rightarrow$ Solution: normalized power prior
Normalized Power prior	Prior on a_0 keep the principle likelihood Measure of the degree to which the historical and current data are commensurate	Choice of prior distribution on a_0

References

[1] Beat Neuenschwander, Gorana Capkun-Niggli, Michael Branson, and David J Spiegelhalter. *Summarizing historical information on controls in clinical trials.* Clinical Trials, 7(1) :5–18,2010.

[2] Stuart J Pocock. *The combination of randomized and historical controls in clinical trials*. Journal of chronic diseases, 29(3) :175–188, 1976.

[3] Heinz Schmidli, Sandro Gsteiger, Satrajit Roychoudhury, Anthony O'Hagan, David Spiegelhalter, and Beat Neuenschwander. *Robust meta-analytic-predictive priors in clinical trials with historical control information*. Biometrics, 70(4) :1023–1032, 2014.

[4] Joseph G Ibrahim, Ming-Hui Chen, Yeongjin Gwon, and Fang Chen. *The power prior : theory and applications*. Statistics in medicine, 34(28) :3724–3749, 2015.

[5] Yuyan Duan. *A modified bayesian power prior approach with applications in water quality evaluation*. PhD thesis, Virginia Polytechnic Institute and State University, 2005.

[6] Neelon Brian. Bayesian analysis using power priors with application to pediatric quality of care. Journal of Biometrics & Biostatistics, 2010.

[7] Christian Robert. *Le choix bayésien : Principes et pratique*. Springer Science & Business Media, 2006.

[8] Kert Viele, Scott Berry, Beat Neuenschwander, Billy Amzal, Fang Chen, Nathan Enas, Brian Hobbs, Joseph G Ibrahim, Nelson Kinnersley, Stacy Lindborg, et al. *Use of historical control data for assessing treatment effects in clinical trials.* Pharmaceutical statistics, 13(1) :41–54, 2014.

[9] Andrew Gelman et al. *Prior distributions for variance parameters in hierarchical models* (comment on article by browne and draper). Bayesian analysis, 1(3) :515–534, 2006.

References

[10] L Di Scala, J Kerman, and B Neuenschwander. *Collection, synthesis, and interpretation of evidence : a proof-of-concept study in copd.* Statistics in medicine, 32(10) :1621–1634, 2013.

[11] Rosalind Walley, John Sherington, Joe Rastrick, Eric Detrait, Etienne Hanon, and Gillian Watt. *Using bayesian analysis in repeated preclinical in vivo studies for a more effective use of animals*. Pharmaceutical Statistics, 2016.

[12] Éric Parent and Jacques Bernier. *Le raisonnement bayésien : modélisation et inférence.* Springer Science & Business Media, 2007.

[13] Jean-Jacques Boreux, Eric Parent, and Jacques Bernier. *Pratique du calcul bayésien.* Springer Science & Business Media, 2010.

[14] US Food, Drug Administration, et al*. Guidance for the use of bayesian statistics in medical device clinical trials*. Silver Spring, Maryland : FDA, 2010.

[15] Dorian Bankovic and Karine Florin. *Évaluation de méthodes bayésiennes de construction de prior dans le cadre de l'évaluation de la supériorité d'un effet traitement par rapport à un groupe contrôle dans les études "proof-of-concept" (poc) des essais cliniques en phase précoce.* Sanofi, 2015.

[16] Morita, S., Thall, P. F., & Müller, P. (2008). *Determining the effective sample size of a parametric prior*. Biometrics, 64(2), 595-602.