Use of a 3 step Bayesian approach for the Behrens-Fisher problem in research experiments

Non-Clinical Statistics Conference
25 September 2012, Potsdam, Germany

Karine Florin
Research and CMC Biostatistics, Sanofi, Montpellier, France

Jean-Michel Marin
Institute of Mathematics and Mathematical Modelling (I3M), University Montpellier 2, France

Antoine Barbieri & Marouane Seffal
Master's degree in Biostatistic, University of Montpellier

SANOFI

Introduction
Why use Bayesian STATISTICS in Research?

- **Specificity of research experiments**
  - Experiments are routinely performed using the same protocol
    - Historical data available
  - Small sample size per experiment
- **Current methods : Frequentist methods**
- **Necessity to explore Bayesian methods**
  - Historical data taken into account
    - More precise : solid conclusion
    - More powerful
  - Small sample inference in the same manner as large sample

SANOFI

Introduction
Context

- **Experimental context**
  - **One Research experiment**
    - Objective: Evaluation of a treatment effect vs control

$$C \sim \mathcal{N}(\mu_c, \sigma_c^2) \quad and \quad T \sim \mathcal{N}(\mu_t, \sigma_t^2)$$

  - **Specifics**
    - Several previous experiments available using the same protocol
- **Behrens-Fisher problem**
  - Comparison of treated and control means normally distributed
    - without assuming the homogeneity of variance hypothesis
- **Current frequentist method applied**
  - **T-Test with Satterthwaite correction**

$$\begin{cases} H_0 : \mu_c = \mu_t = \mu \\ H_1 : \mu_c \neq \mu_t \end{cases}$$

SANOFI

Introduction
Contents

- **Classical Bayesian method**
  - **Delta and credible intervals**
- **Model choice Bayesian method**
  - **Calculation of the posterior probabilities and bayes factor**
  - **Proposition of a three step Bayesian method**
    - Robust choice of objective and subjective combined priors
- **Application**
  - **Real data**
  - **Simulated data**
- **Conclusion**

SANOFI

Classical Bayesian approach
Delta and credible intervals

- **Classical Bayesian approach** $\delta = \mu_c - \mu_t$
  - Choice of the prior distribution
  - Estimation of the Posterior distribution according to the prior
  - Estimate the credibility interval of δ
  - Rule: Reject the equality between means if zero is outside the credibility interval
- **Need to explore an other approach**
  - To do inference Bayesian testing
    - Using the model choice theory
  - To estimate the **posterior probability of H0 and H1 hypotheses**
    - Probability of the difference between means
    - Probability of the equality between means

Formal Bayesian approach
Model choice theory

- **Bayes factor**
  - BF = posterior odds ratio/ prior odds ratio

$$B_{1,0}(y) = \frac{P(M_1|y)/P(M_0|y)}{P(M_1)/P(M_0)}$$

- **Scale of decision for Bayes factor**
  - Jeffrey's scale (1961)
  - More recently: Kass& Raftery scale (1995)

$2\log_e(B_{10})$	(B_{10})	Evidence against H_0
0 to 2	1 to 3	Not worth more than a bare mention
2 to 6	3 to 20	Positive
6 to 10	20 to 150	Strong
>10	>150	Very strong

Formal Bayesian approach
Model choice theory

- **Choice between two models M0 and M1**

$$M_0 : \{\mu_c = \mu_t = \mu\} \Longleftrightarrow \left\{ \begin{array}{l} C \sim \mathcal{N}(\mu, \sigma_c^2) \\ T \sim \mathcal{N}(\mu, \sigma_t^2) \end{array} \right. \quad M_1 : \{\mu_c \neq \mu_t\} \Longleftrightarrow \left\{ \begin{array}{l} C \sim \mathcal{N}(\mu_c, \sigma_c^2) \\ T \sim \mathcal{N}(\mu_t, \sigma_t^2) \end{array} \right.$$

$$\theta_0 = (\mu, \sigma_c, \sigma_t) \qquad \theta_1 = (\mu_c, \mu_t, \sigma_c, \sigma_t)$$

- **Posterior probability of each model**

Prior probabilities

$$P(M_j|y) = \frac{P(M_j)P(y|M_j)}{P(M_0)P(y|M_0) + P(M_1)P(y|M_1)}, \; j = 0,1$$

$$P(y|M_j) = \int \underbrace{\ell_j(\theta_j|y)}_{\text{Likelihood}} \underbrace{\pi_j(\theta_j)}_{\text{Prior}} \, d\theta_j, \; j = 0,1$$

Proposition of a 3 steps Bayesian method

- Interest of the Bayesian methods: Prior !
  - Improve precision and power of analyses
- Drawback of the Bayesian methods: Prior!
  - Choice of prior can be controversial
- Idea of the proposed sequential Bayesian method
  - Robust choice of combined priors
    - Non informative prior
    - Informative prior
      - Incorporation of informations based on historical data
  - 3 steps are necessary to estimate posterior probabilities

Proposition of a 3 steps Bayesian method

- **Step 1:**
  - Prior : Jeffreys' prior (improper!)
  - Likelihood: Data of experiment 1
  - Result : Posterior distribution / Model posterior probabilities not defined
- **Step 2:**
  - Prior : Step 1 posterior distribution / $P(M_0)$ & $P(M_1)$ =1/2
  - Likelihood: Data of experiment 2
  - Result : Posterior distribution / Model posterior probabilities
- **Step 3:**
  - Prior : Step 2 posterior distribution/ Model Step 2 posterior probabilities
  - Likelihood: Data of experiment 3
  - Result : Model Posterior probability & Bayes factor

Application on real data

Proposition of a 3 steps Bayesian method Development of the method

- Under M1: Explicit
  - **Posterior distributions (for each step)**
    - Normal distribution for mean parameters
    - Inverse-Gamma for variance parameters
  - **Calculation of integral of the posterior distributions**
- Under M0: Non explicit
  - **Posterior distribution**
    - Estimation of the variance posterior parameters distribution
      - Use of sampling methods (MCMC methods through WinBUGS)
      - Estimation of inverse-Gamma parameters for each sampling
  - **Approximation of integral by numerical methods**
      - Adaptative integration from sampling of parameters
- 3 step Bayesian method results

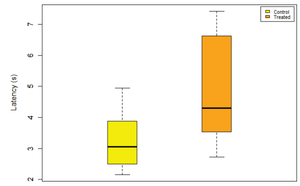
  Ratio of integrals $\longrightarrow$ **Bayes factor and posterior probabilities**

Application on real data

- Description of the CFA protocol
  - **Aim of the study:**
    - Evaluate potential anti-inflammatory product after intra plantar administration of CFA (Freund's Complete Adjuvant) in mice
  - **Description of the thermal test:**
    - A radiant heat source was focused on the paw
  - **Measured parameter:**
    - Latency (s) from the initiation of the radiant heat until paw withdrawal
  - **Normality and homogeneity of variance hypotheses:**
    - Previous statistical studies (realized with Sample Size estimation) have been done. The normality is satisfying but there is a problem of heterogeneity of variance on this protocol

Application on real data
Ibuprofen effect

- Objective of the study : Ibuprofen effect versus Vehicle



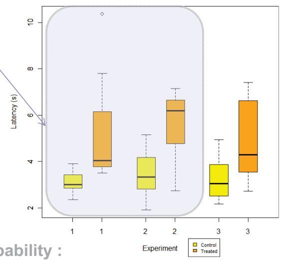
- Results of frequentist approach
  - Rejection of the null (H0) at the 5%
  - P-value near to the threshold

Test de Student	
P-value :	0.04^{*}

Application on real data
Ibuprofen effect

- Bayesian approach : three steps method
  - Choice of two prior experiments (in agreement with scientist)

Bayesian Method			
Posterior probabilities		Bayes Factor	
$\Pr(M_0 \mid y)$	$\Pr(M_1 \mid y)$	$B_{1,0}$	Evidence against M_0
0.006	0.994	9.63	*Positive***



- Direct interpretation of the Posterior probability :
  - The probability that the ibuprofen has no effect in comparison to the vehicle group is 0.6%
  - The probability that the ibuprofen is different from the control is 99.4%

Application on simulated data

APPLICATION TO SIMULATED DATA

- Objectives :
  - **Verify the good frequentist properties of the 3 step Bayesian method**
    According to the FDA guideline "Guidance for the Use of Bayesian Statistics in Medical Device Clinical Trials"
    - Control of Type I error
    - Evaluation of power (the converse of type II error rate)
  - **Compare the power of three steps method & current frequentist method used**
    - Three steps Bayesian method :
      - Bayes factor power : number of time (%) that interpretation concludes at least "positive evidence" (Kass's scale)
      - Posterior probability power : number of time (%) that posterior probability is greater than 0.8
    - Frequentist approach :
      - T-test power : Number of times that p-value is less than 0.05 (%)

APPLICATION TO SIMULATED DATA How?

- 3 experiments simulated using normal distribution
  - **Distribution parameters from CFA historical data**
    - Control group (quite stable)
      - Mean &Sd: Median of CFA Vehicle values
    - Treated group:
      - Mean: **Four sizes effect : 0%, 30%, 40% and 50%**
      - Sd: **Min** (0.6), **median** (1.5), **max** (3)
  - **Size per group : 10** (max N used on the protocol)
  - **Number of simulations : each experiment 1000 times**

APPLICATION TO SIMULATED DATA RESULTS

Experiment 1	10%	10%	30%	50%
Experiment 2	10%	10%	30%	30%
Experiment 3	40%	50%	10%	10%
% T.test	59.00	78.00	9.25	8.2
% $\Pr(M_1 \mid y) > 0.8$	32.65	45.75	27.90	26.8
% $BF_{(M_1/M_0)}$	50.20	67.05	13.85	15.7

- Bayesian method power is affected by the previous experiment effect
  - **No effect in prior experiment decreases the Bayesian power**
  - **Effect in prior experiment increases the Bayesian power**

APPLICATION TO SIMULATED DATA RESULTS

$N = 1000$, $n = 10$, $\mu_c = 2.94$ and $\sigma_{t_1} = \sigma_{t_2} = 1.5$

Effect	0%			30%			40%			50%		
σ_{t_3}	min	median	max	min	median	max	min	median	max	min	median	max
% T.test	4.8	5.25	5.4	87.1	36.20	14.3	98.8	56.30	20.4	99.9	75.25	29.3
% $\Pr(M_1 \mid y) > 0.8$	1.2	1.95	1.1	78.7	58.10	27.6	94.9	81.85	43.0	99.3	93.7	57.7
% $BF_{(M_1/M_0)}$	0.7	3.20	4.7	88.8	51.75	19.1	98.9	71.65	26.2	99.8	82.95	27.7

- Whatever the variability of the last experiment :
  - **Control of the type 1 error for three steps Bayesian method**

- Three steps Bayesian method
  - More powerful than current approach
- In the case of the high variance heterogeneity :
  - **Posterior probability : more powerful**

Conclusion

CONCLUSION

- Three step Bayesian method developped for the Behrens-Fisher problem
  - **Robust choice of prior**
    - Combination of non informative and informative priors
  - **Estimation of the posterior probability of each hypothesis**
    - Direct interpretation of the probabilities
- According to FDA, correct frequentist properties need to be verify
    - Control of type 1 error
    - Sufficient Power
  - **OK for CFA protocol with N=10**
- As expected, when compared with actual frequentist methods used on real & simulated data:
  - **Be more powerful**

PRINCIPAL BIBLIOGRAPHY

- FDA (2010). Guidance for the use of Bayesian statistics in medical device clinical trials
- Jeffreys, H., (1961). Theory of probability (3rd ed.), Oxford, U.K.: Oxford University Press.
- Kass, R. E., & Raftery, A. E., (1995). Bayes factors, Journal of the American Statistical Association, 90 (430), 773-795.
- O'Hagan, T., (2006). Bayes factors, Significance, 3, 184–186.
- Marin, J-M., & Robert, C. P., (2007). Bayesian Core: A Practical Approach to Computational Bayesian Statistics, Springer.
- Albert, J., (2009). Bayesian Computation with R (2nd Ed.), Springer.
- Ghosh&Kim, (2001). The Behrens-Fisher problem revisited: a bayes-frequentist synthesis, The canadian Journal of statistics, vol. 29, pp 5-17
- Scott, J. G. and Berger, J. O. (2006) An exploration of aspects of Bayesian multiple testing, Journal of Statistical Planning and Inference, 136, 2144 – 2162

Thank you !!!