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 (UM), University Montpellier 2, France

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

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

Introduction

Context

● Experimental context
  - One Research experiment
    - Objective: Evaluation of a treatment effect vs control
      - $C \sim \mathcal{N}(\mu_c, \sigma_C^2)$ and $T \sim \mathcal{N}(\mu_t, \sigma_T^2)$
  - 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
    \[
    \begin{align*}
    H_0 : \mu_c &= \mu_t = \mu \\
    H_1 : \mu_c &\neq \mu_t
    \end{align*}
    \]
Classical Bayesian approach
Delta and credible intervals

- Classical Bayesian approach
  \( \delta = \mu_e - \mu_t \)
  - Choice of the prior distribution
  - Estimation of the posterior distribution according to the prior
  - Rule: Reject the equality between means if zero is outside the credibility interval

- Need to explore another approach
  - To do inference Bayesian testing
  - Using the model choice theory

Formal Bayesian approach
Model choice theory

- Bayes factor
  \( BF = \frac{P(M_1|y)/P(M_0|y)}{P(M_1)/P(M_0)} \)

- Scale of decision for Bayes factor
  - Jeffrey’s scale (1961)

<table>
<thead>
<tr>
<th>( 2 \log(B_{10}) )</th>
<th>( (BF) )</th>
<th>Evidence against ( H_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 to 2</td>
<td>1 to 3</td>
<td>Not worth more than a bare mention</td>
</tr>
<tr>
<td>2 to 6</td>
<td>3 to 20</td>
<td>Positive</td>
</tr>
<tr>
<td>6 to 10</td>
<td>20 to 150</td>
<td>Strong</td>
</tr>
<tr>
<td>&gt; 10</td>
<td>&gt; 150</td>
<td>Very strong</td>
</tr>
</tbody>
</table>

- 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_1) & P(M_2) = 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

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

Description of the protocol
- 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
    - Adaptive integration from sampling of parameters
  - 3 step Bayesian method results
    - Ratio of integrals • Bayes factor and posterior probabilities

Aim of the study:
- Evaluate potential anti-inflammatory product after intra plantar administration of CFA (Freund’s Complete Adjuvant) in mice

Description of the protocol:
- A radiant heat source was focused on the paw
- Latency (s) from the initiation of the radiant heat until paw withdrawal

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

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

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
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

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

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

- 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

<table>
<thead>
<tr>
<th>Effect</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{Pr}(M_2</td>
<td>\theta) &gt; 0.8)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(\text{BF}_{M_2(M_1)})</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experiment</th>
<th>10%</th>
<th>10%</th>
<th>30%</th>
<th>30%</th>
</tr>
</thead>
<tbody>
<tr>
<td>% (L_1p)</td>
<td>50.00</td>
<td>78.00</td>
<td>2.25</td>
<td>8.2</td>
</tr>
<tr>
<td>% (\text{Pr}(M_1</td>
<td>\theta) &gt; 0.8)</td>
<td>32.65</td>
<td>45.75</td>
<td>27.90</td>
</tr>
<tr>
<td>(\text{BF}_{M_2(M_1)})</td>
<td>30.20</td>
<td>67.05</td>
<td>13.83</td>
<td>15.7</td>
</tr>
</tbody>
</table>

Conclusion

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

- Three step Bayesian method developed 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