Bayesian Sample Size Calculations

Tony O'Hagan, University of Sheffield
John Stevens, AstraZeneca R&D Charnwood

The Problem

• A trial is to be designed comparing Treatment 1 with Treatment 2
• Both efficacy and cost will be measured for each patient in the trial
• How large a sample is needed?
  • \( n_1 \) patients in group 1 receiving Treatment 1
  • \( n_2 \) patients in group 2 receiving Treatment 2
Cost-Effectiveness

- We will judge cost-effectiveness by the net monetary benefit $\beta(K)$ of switching from Treatment 1 to Treatment 2.
- If the true cost difference is $\Delta_c$ and the true efficacy difference is $\Delta_e$, then
  $$\beta(K) = K \Delta_e - \Delta_c$$
- Treatment 2 is more cost-effective than Treatment 1 if $\beta(K) > 0$.

Willingness to Pay

- The factor $K$ is the willingness to pay.
- It represents what the health provider is prepared to pay in order to obtain an increase of one unit in efficacy.
- In any formulation of the sample size problem, $K$ must be specified.
  - The sample size will depend on it.
Two Stages

- There are two stages to the study
  - Design stage. Plan the study in order to have a good chance of the desired outcome.
  - Analysis stage. Analyse data from the study, to see whether we can report the desired outcome.
- These two stages are reflected in
  - setting objectives
  - how we do the sample size calculations

Objectives

- Analysis objective
  - The desired outcome is to be able to report a probability of at least $\omega$ that Treatment 2 is more cost-effective than Treatment 1
- Design objective
  - We want a probability at least $\delta$ of being able to report the desired outcome
Analysis Objective

• Frequentist formulation
  • We wish to reject the null hypothesis that $\beta(K) = 0$, at the $100(1 - \omega)\%$ level of significance
  • e.g. $\omega = 0.95$ corresponds to usual 5% test

• Bayesian formulation
  • We wish to have at least a $100\omega\%$ posterior probability that $\beta(K) > 0$

Design Objective

• Frequentist formulation
  • We want sample sizes large enough to give $100\delta\%$ power to reject the null hypothesis that $\beta(K) = 0$, when the true value of $\beta(K)$ has some assumed alternative value

• Bayesian formulation
  • We want sample sizes large enough to give a $100\delta\%$ prior probability of achieving the desired posterior probability that $\beta(K) > 0$
Understanding the Bayesian Formulation - Analysis

- At the Analysis stage, the two objectives are similar
  - particularly if we employ weak prior information in the analysis of the data
- The Bayesian statement is how the p-value is usually (mis-)interpreted anyway
- So the meaning of $\omega$ is clear
  - Two sided formulations are also possible

Understanding the Bayesian Formulation - Design

- At the Design stage, however, the two objectives are quite different
- The frequentist approach fixes the parameters at more or less arbitrary values
- The Bayesian objective requires a probability $\delta$ of the desired analysis outcome *averaged* over the prior distribution of the parameters
Putting it Together

• Putting the two frequentist objectives together:
  “We wish to have 70% power to reject the null hypothesis that $\beta(K) = 0$ at the 5% level, if the true value is $\beta(K) = £2800$.”

• Putting the two Bayesian objectives together:
  “We wish to be 70% sure of obtaining a 95% posterior probability that $\beta(K) > 0$”

The Prior Distribution

• When advocating a Bayesian approach, the usual question arises - what about the prior distribution?

• Various possibilities
  • The company’s genuine prior beliefs
  • Realistic beliefs of the wider community
  • Noninformative
  • Sceptical
  • Regulator’s or decision-maker’s beliefs
Which Prior?

<table>
<thead>
<tr>
<th>Department</th>
<th>Analysis</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Company</td>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>Community</td>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>Noninformative</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Sceptical</td>
<td>✓</td>
<td>?</td>
</tr>
<tr>
<td>Regulator</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

Two Priors

- We will allow different prior distributions at the two stages
- Analysis prior will typically be noninformative, sceptical or some kind of consensus
- Design prior should generally be based on the company’s best knowledge
Example - 1

- Briggs and Tambour (1998):
  - Frequentist formulation
  - Assumed efficacy difference $\Delta_e = 0.8$
  - Assumed cost difference $\Delta_c = 1200$
  - $K = 5000$, hence true $\beta(K) = \beta(5000) = 2800$
  - 5% two-sided test, 70% power
  - Equal sample sizes of 762

Example - 2

- Bayesian formulation
  - Analysis prior noninformative
  - Design prior expectation of $\beta(5000)$ set to 6300, standard deviation 8367
  - $\omega = 0.975$ (1-sided, equates to 5% 2-sided test)
  - $\delta = 0.7$ (corresponding to 70% power)
  - Equal sample sizes of 1048
Example - Design Prior

Probability that $\beta(5000) > 0$ is 0.774
Probability that $\beta(5000) > 2800$ is 0.662

Power curves

Power curves for three different sample sizes.
Power is also Bayesian conditional assurance with a noninformative analysis prior
Example - Assurance

Assurance $\delta$ is expected power with respect to prior.

Curves give $\delta = 0.15, 0.49, 0.70$.

Example - Conclusion

- The Bayesian analysis has given a larger sample size
- It assures a 70% chance overall of demonstrating that Treatment 2 is more cost-effective than treatment 1
- But it recognises the chance (22.6%) that Treatment 2 is really not more cost-effective
Some Extreme Cases

• If the Analysis prior is too strong, we may find that zero sample size is enough
  • The Analysis prior alone gives probability $\omega$ that Treatment 1 is more cost-effective

• If the Design prior is too weak, we may find that infinite sample size is not enough
  • The Design prior does not give probability $\delta$ that Treatment 1 is more cost-effective

Discussion

• The Bayesian approach allows considerable flexibility to represent the real problem
• The Design objective is expressed as an overall assurance instead of requiring an assumed, more or less arbitrary, true effect
• This talk is based on a paper accepted for publication in Medical Decision Making - see http://www.shef.ac.uk/~st1ao
  (Email: a.ohagan@sheffield.ac.uk)