A Bayesian model for the selection of sample size in clinical trials

Sebastian Jobjörnsson

Chalmers University of Technology

June 5, 2014

Introduction

- IDEAL = Integrated Design and Analysis of small population group trials.
- Overview
    1. Basic introduction to Bayesian decision theory.
    2. Description of the decision situation.
    3. Description of a statistical model.
    4. Application of backward induction to the model.
    5. Conclusions and discussion of extensions to the model.

Bayesian decision theory

$$\begin{aligned}\text{BDT} &= \text{Subjective Bayesian statistics}\\ &+ \text{Principle of Maximum Expected Utility}\end{aligned}$$

This can be expanded somewhat:

- Probabilites are degrees of belief.
- Priors on unknown parameters.
- Actions should maximize expected utility.
- Statistical model + Utility function $\implies$ Optimal policy.

Decision situation for a clinical trial

- Two treatments: one new, one old.
- Assumption: mean effect for old is 0.
- Assumption: variance for old effect is very small.

Let N be the size of the total population.
Let X_k be the effect observed for subject k if given the new treatment.
The agent should do the following:

1. Choose a sample size n in a clinical trial, $0 \leq n \leq N$, resulting in observed effects $X_1, \ldots, X_n$.
2. Decide on treatment d (1 for new and 0 for old), for $N - n$ remaining subjects, resulting in observed effects $X_{n+1}, \ldots, X_N$ or $0, \ldots, 0$.

This is a two-stage decision problem, with n and d as decision variables.

Statistical model

- $X_1, \ldots, X_N$ i.i.d., with $X_k \mid \mu \sim \mathrm{N}(\mu, \sigma^2)$.
- Assign a conjugate prior distribution $\mu \sim \mathrm{N}(m, s^2)$ to the unknown population mean.
- The variance σ^2 is assumed known.
- m and s^2 are assumed known.

Some notation

Notation for sequences of random variables and realized values of those random variables:

- x_k denotes a realized value of the random variable X_k.
- $X^n = (X_k)_{k=1}^n$ and $x^n = (x_k)_{k=1}^n$.
- $X^N = (x_k)_{k=n+1}^N$ and $x^N = (x_k)_{k=n+1}^N$.

Notation for sequence means:

$$\overline{X^n} = \frac{\sum_{k=1}^n X_k}{n}$$

$$\overline{x^n} = \frac{\sum_{k=1}^n x_k}{n}$$

$$\overline{X^N} = \frac{\sum_{k=n+1}^N X_k}{N-n}$$

$$\overline{x^N} = \frac{\sum_{k=n+1}^N x_k}{N-n}$$

Utility function

The utility function for the agent is defined as

$$u(n, \overline{x^n}, d, \overline{x^N}) = \begin{cases} \sum_{k=1}^{n} x_k & \text{if d = 0} \\ \sum_{k=1}^{N} x_k & \text{if d = 1} \end{cases}$$
$$= n * \overline{x^n} + d * (N - n) * \overline{x^N}.$$

$u(n, \overline{x^n}, d, \overline{x^N})$ is the utility corresponding to

1. A choice of sample size n, resulting in $X^n = x^n$.
2. A choice d of treatment for remaining subjects, resulting in $X^N = x^N$ (if $d = 1$).

How to choose an optimal n?

Use backward induction.
For our specific problem, we had:

1. Choose a sample size n ...
   ... and observe effects $X_1, \ldots, X_n$.
2. Give the $N - n$ remaining subjects the new ($d = 1$) or old ($d = 0$) treatment ...
   ... and observe effects $X_{n+1}, \ldots, X_N$ (if $d = 1$).

After all steps, we have $u(n, \overline{x^n}, d, \overline{x^N})$.
Now go backwards in time:

- Compute $u(n, \overline{x^n}, d) = \mathbb{E}[u(n, \overline{x^n}, d, \overline{X^N}) \mid n, \overline{x^n}, d]$.
- Choose optimal d by maximizing $u(n, \overline{x^n}, d)$.
- Compute $u(n) = \mathbb{E}[\max_d u(n, \overline{X^n}, d) \mid n]$.

Computation of $u(n, \overline{x^n}, d)$

Recall the definition of the utility function as the sum of all observed effects:

$$u(n, \overline{x^n}, d, \overline{x^N}) = n * \overline{x^n} + d * (N - n) * \overline{x^N}.$$

This gives

$$\begin{aligned} u(n, \overline{x^n}, d) &= \mathbb{E}[n * \overline{X^n} + d * (N - n) * \overline{X^N} \mid n, \overline{x^n}, d] \\ &= n * \overline{x^n} + d * (N - n) * \mathbb{E}[\overline{X^N} \mid n, \overline{x^n}, d] \\ &= n * \overline{x^n} + d * (N - n) * m_n, \end{aligned}$$

where $m_n = \mathbb{E}[\overline{X^N} \mid n, \overline{x^n}, d]$ is the updated mean for the posterior distribution of the population mean μ.

Computation of $u(n) = \mathbb{E}[\max_d u(n, \overline{X^n}, d) \mid n]$

From the previous slide:

$$u(n, \overline{x^n}, d) = n * \overline{x^n} + d * (N - n) * m_n.$$

This gives:

$$\begin{aligned} u(n) &= \mathbb{E}[\max_d (n * \overline{X^n} + d * (N - n) * m_n) \mid n] \\ &= \mathbb{E}[n * \overline{X^n} + \max_d (d * (N - n) * m_n) \mid n] \\ &= \mathbb{E}[n * \overline{X^n} + (N - n) * \max(0, m_n) \mid n] \\ &= \mathbb{E}[n * \overline{X^n} \mid n] + (N - n) * \mathbb{E}[\max(0, m_n) \mid n] \\ &= n * m + (N - n) * \mathbb{E}[\max(0, m_n) \mid n]. \end{aligned}$$

Set $t_n^2 = s^2(n/\sigma^2)s_n^2$, where s_n^2 is the updated variance for the population mean μ. Then:

$$u(n) = n * m + (N - n) * \left(t_n^2 * \mathrm{N}(0 \mid m, t_n^2) + m * \left(1 - \Phi(0 \mid m, t_n^2)\right)\right).$$

Example 1

Consider the form of $u(n)$ for the following parameter values:

- Mean m of the prior for the population mean μ equals 0.
- Variance s^2 of the prior for the population mean μ equals 1.
- The population variance σ^2 equals 1.

Insertion of these parameter values results in

$$u(n) = \frac{(N-n)}{\sqrt{2\pi}} \sqrt{\frac{n}{n+1}}.$$

Plot of $u(n)$ for example 1

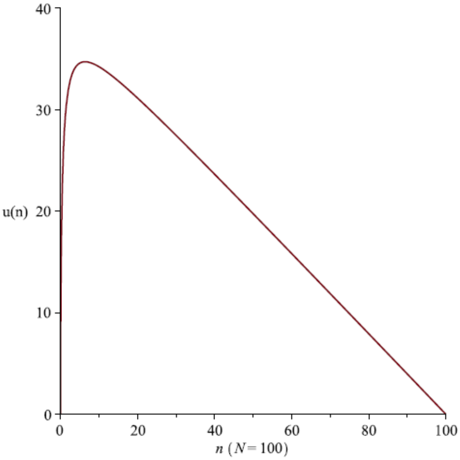


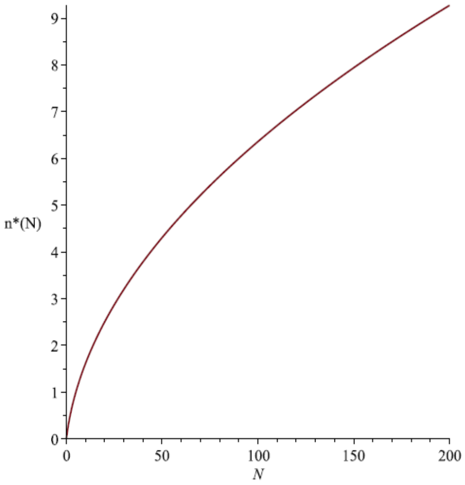
Figure: Plot of $u(n)$ versus n, for $N = 100$.

Optimal sample size for a specific population size

Let $n^*(N)$ be the optimal sample size for a specific N.
To find $n^*(N)$, differentiate $u(n)$ w.r.t. n and set the derivative to 0.
This leads to the formula:

$$n^*(N) = -\frac{3}{4} + \frac{1}{4}\sqrt{9 + 8N}$$

Plot of $n^*(N)$ for example 1



Addition of a regulatory requirement

Introduce a binary regulatory function $r(n, \overline{x^n})$.
The agent may choose $d = 1$ if and only if $r(n, \overline{x^n}) = 1$.
Assume the form:

$$r(n, \overline{x^n}) = \begin{cases} 1 & \text{if } \overline{x^n} > \frac{\sigma z_\alpha}{\sqrt{n}} \\ 0 & \text{if } \overline{x^n} \leq \frac{\sigma z_\alpha}{\sqrt{n}} \end{cases},$$

where $\mathbb{P}(Z \geq z_\alpha) = \alpha$ if $Z \sim \mathrm{N}(0, 1)$.
This corresponds to testing $H_0 : \mu \leq 0$ against $H_1 : \mu > 0$, at significance level α.
$u_r(n)$ may be derived as in the case without any requirement.

Example 2: Formulas

Compute $u(n)$ and $u_r(n)$ for the same parameter values as in example 1:
$m = 0$, $\sigma^2 = s^2 = 1$. Insertion into the general formulas yields

$$u(n) = \frac{(N-n)}{\sqrt{2\pi}}\sqrt{\frac{n}{n+1}}$$

$$u_r(n) = u(n)e^{-\frac{1}{2}\left(\frac{z_\alpha^2}{n+1}\right)}.$$

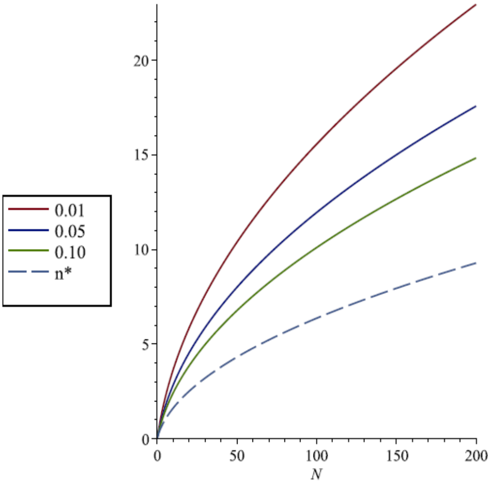
Example 2: Optimal sample sizes

Let

- $n^*(N)$ = Optimal sample size without r.
- $n_r^*(N)$ = Optimal sample size with r.

The next slide shows a plot of these functions for three different significance levels: $\alpha = 0.01$, $\alpha = 0.05$ and $\alpha = 0.10$.

Example 2: Plot of optimal sample sizes



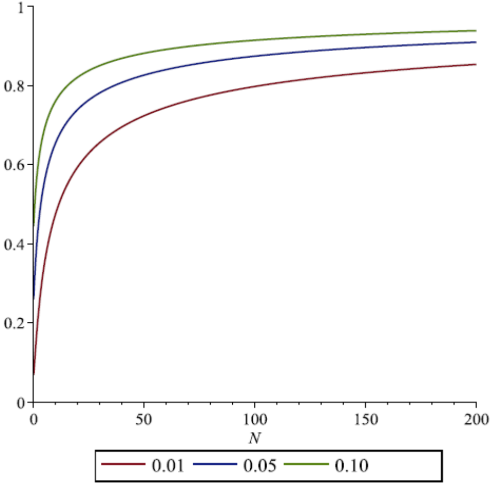
Example 2: Optimal utilities

Let

- $u^*(N) = u(n^*(N))$ = Optimal utility without r.
- $u_r^*(N) = u_r(n^*(N))$ = Optimal utility with r.

In order to compare the two cases,
we plot the ratio $\frac{u_r^*(N)}{u^*(N)}$ against N.

Example 2: Plot of $u_r^*(N)/u^*(N)$ versus N



Conclusions

For our simple model:

1. Significant impact of regulatory requirements for small population sizes N.
2. The precise loss in expected utility may be computed.

But the model is too simple!

Possible extensions

For the statistical model:

- Drop the assumption of known σ^2 (and use prior).
- Generalize to other distributions for the effect.

For the utility function:

- Include cost of trial in the utility function.
- Change the utility function entirely? (For example, if the agent is a company).

For the structure of the decision process:

- More stages, there could be several clinical trials for the same treatment.
- More agents? At least two seems appropriate: Sponsor and Regulatory Agency.
- Go from constant N to random process N_t?

Last slide

Thanks for listening!