Iranian Int. J. Sci. ***6****(2), 2005, p.203-212*

A Fully Bayesian Sample Size Determination for Inference on the Difference of two Binomial Proportions

H. Pezeshk[1], V. Maroufy[2]
[1]*Center of Excellence of Biomathematics, School of Mathematics, Statistics, and Computer Science, University College of Science, University of Tehran, Tehran, Iran - e-mail: pezeshk@khayam.ut.ac.ir*
[2]*Department of Statistics, Faculty of Economics, University of Allameh Tabatabaii, Tehran, Iran - e-mail: maroofy180@yahoo.com*
(*received: 6/5/2005 ; accepted: 17/9/2005*)

Abstract

A fully Bayesian approach for sample size determination for a clinical trial is presented in which the final decision whether to use the new treatment is taken by potential users and their medical advisers on the basis of the strength of the evidence provided by the trial. Data are assumed to come from two independent binomial distributions and the parameter of interest is $p = p_1 - p_2$, where p_1 and p_2 are two independent proportions. The optimal size is obtained by maximizing the expected net benefit function, which is the expected benefit from subsequent use of the new treatment minus the cost of the trial.

Keywords*: Clinical Trials, Sample Size Determination, Bayesian Approach, Binomial Distribution, Expected Net Benefit*

1. Introduction

An important question in the planning of medical experiments to assess the performance of new drugs or treatments is how big to make the trial. There have been a number of papers on the subject, from both the frequentist and Bayesian points of view. In the frequentists' approach sample sizes are usually determined either by absolute error criterion or by power and size control rule (Desu and Raghavarao, 1990; Vol. 12 No. 4, 2002 of Journal of Biopharmaceutical Statistics). Bayesian approach to sample size question may be divided into two groups: inferential (Adcock, 1988; Joseph *et al*., 1997) and fully

Bayesian or decision theoretic approach (Lindley, 1997; Pezeshk and Gittins, 2002).

The first paper on the fully Bayesian approach was by Grundy *et al* (1956). A major problem in following this approach is that the ultimate decision on whether or not to use the new treatment is taken by large number of patients and their medical advisers, and does not depend on the outcome of the trial in any clear-cut way. Some authors (Pezeshk and Gittins, 1999, 2002; Gittins and Pezeshk, 2000, 2002) return to the decision theoretic analysis of Raiffa and Schlaifer (1961), with the modification that instead of a utility maximizing terminal decision a plausible model is assumed for the way patients and their medial advisers respond to the evidence from a trial. Inferential Bayesian methods without utility functions are reviewed by Pezeshk (2003).

Bayesian sample size determination for estimating the success probability in binomial sampling has received considerable attention (Adcock, 1992, 1995; Joseph *et al.,* 1995; Pham-Gia and Turkkan, 1992, 2003; Pezeshk and Gittins, 2002).

In this work we consider a fully Bayesian approach to sample size determination in which the number of subsequent users of the therapy under investigation, and hence also the total benefit resulting from the trial, depends on the strength of the evidence provided by the trial. We model the subsequent usage by plausible assumptions for actual behaviour, rather than assuming that this represents decisions which are in some sense optimal. For this reason the procedure may be called "Behavioural Bayes" (or BeBay for short).

Using the central limit theorem to justify the assumption of normality for the case with binary responses, Gittins and Pezeshk (2000) discussed how one may apply the BeBay methodology to binomially distributed data. Pezeshk and Gittins (2002) extend the BeBay methodology to binomial data for trials where there is no control group. This work applies the BeBay methodology to the Bernoulli sampling of two independent populations. The procedure is applicable to trials for drugs applied to new classes of patients and phase III clinical trials for which there are control groups. The optimal sample size is obtained by maximizing the expected net benefit, which

is the benefit from subsequent use of the new treatment minus the cost of conducting the trial.

The work is organized as follows. In the next section we introduce the notation. In section 3 the function *m* representing the number of subsequent users of the new treatment is introduced. Section 4 works out the expected net benefit function and sets out the results of running the BeBay program for a trial.

2. Notation

Let X_1 and X_2 be the total number of "successes" out of n_1 and n_2 trials from independent binomial experiments with parameters p_1 and p_2, respectively. Let p_1 and p_2 denote the probabilities of a favorable outcome for individuals in the treated group and the control group, respectively. To formulate our prior knowledge about $p_i\,(i=1,2)$, we assume that they have prior densities. With a binomial likelihood it is mathematically convenient, and often reasonably realistic, to make the assumption that p_i has a prior distribution π_i which is beta with parameters α_i and $\beta_i\,(i=1,2)$. Thus

$$\pi_i(p_i)=\frac{1}{B(\alpha_i,\beta_i)}\,p^{\alpha_i-1}(1-p_i)^{\beta_i-1}\quad \alpha_i,\beta_i>0,\quad 0\le p_i\le 1,\quad i=1,2 \quad (1)$$

($B(\alpha_i,\beta_i)$ is the beta function with parameters α_i and β_i). We may write this as $p_i \sim Beta(\alpha_i,\beta_i)$. If X_i takes the value x_i, then the posterior distribution of p_i is a beta distribution with parameters α_i+x_i and $\beta_i+n_i-x_i\,(i=1,2)$.

The means and the variances of the prior and posterior distributions are then, respectively,

$$\mu_i = \frac{\alpha_i}{\alpha_i + \beta_i}, \qquad \tau_i^2 = \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i)^2(\alpha_i + \beta_i + 1)}, \quad i = 1,2,$$

and (2)

$$\mu_i' = \frac{\alpha_i + x_i}{\alpha_i + \beta_i + n_i}, \qquad \tau_i'^2 = \frac{(\alpha_i + x_i)(\beta_i + n_i - x_i)}{(\alpha_i + \beta_i + n_i)^2(\alpha_i + \beta_i + n_i + 1)}, \quad i = 1,2.$$

Under the assumption of independence, the mean and the variance of the posterior density for $p = p_1 - p_2$ are, respectively:

$$\mu' = \mu_1' - \mu_2' = \frac{\alpha_1 + x_1}{\alpha_1 + \beta_1 + n_1} - \frac{\alpha_2 + x_2}{\alpha_2 + \beta_2 + n_2}$$

and

$$\tau'^2 = \tau_1'^2 + \tau_2'^2$$
$$= \frac{(\alpha_1 + x_1)(\beta_1 + n_1 - x_1)}{(\alpha_1 + \beta_1 + n_1)^2(\alpha + \beta_1 + n_1 + 1)} + \frac{(\alpha_2 + x_2)(\beta_2 + n_2 - x_2)}{(\alpha_2 + \beta_2 + n_2)^2(\alpha_2 + \beta_2 + n_2 + 1)}$$

Note that the predictive distribution of X_i is a beta-binomial distribution with parameters α_i, β_i and n_i and density function

$$f_{X_i}(x_i) = \binom{n_i}{x_i} \frac{B(\alpha_i + x_i, \beta_i + n_i - x_i)}{B(\alpha_i, \beta_i)}, \quad x_i = 0,1,2,\ldots,n_i, \quad i = 2,1. \tag{3}$$

So, under the assumption of independence of X_i's, the joint density function of (X_1, X_2) is

$$f_{X_1,X_2}(x_1,x_2) = \prod_{i=1}^{2} \binom{n_i}{x_i} \frac{B(\alpha_i + x_i, \beta_i + n_i - x_i)}{B(\alpha_i, \beta_i)} \tag{4}$$

Before taking a sample of fixed size n_i, pre-posterior analysis studies the possible outcomes concerning the posterior distribution. Since X_i is a random variable, it follows that the posterior mean, μ_i', and the

posterior variance, $\tau_i'^2$, for p_i are both random variables because of their dependence on X_i, $(i = 1,2)$.

As noted by Pham-Gia and Turkkan (1992), we may use (3) to obtain the means and the variances of μ_i' and $\tau_i'^2$. For instance, for $i = 1,2,$ the means of the posterior mean and posterior variance are, respectively,

$$E(\mu_i') = \frac{\alpha_i}{\alpha_i + \beta_i} = \mu_i, \qquad E\tau_i'^2 = \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i)(\alpha_i + \beta_i + 1)(\alpha_i + \beta_i + n_i).} \quad i = 1,2$$

Using these parameters and the variances of μ_i' and τ_i' one may calculate the sample size required to satisfy precision conditions on criteria related to the posterior distribution (Pham-Gia and Turkkan, 1992). Here we follow the fully Bayesian methodology and obtain the optimal sample size by maximizing the expected net benefit from conducting the trial minus its cost.

3. Number of subsequent users of the new treatment

Following Gittins and Pezeshk (2002) we assume that the number of subsequent users of the new treatment will not be high, unless it is, in the statistical sense, significantly better than the current one. There must also be a reasonable expectation that the new treatment achieves a sufficiently large improvement to justify a switch.

These considerations lead us to assume that the number of subsequent users of the new treatment, *m*, depends on the mean μ' and the standard deviation τ' of the posterior distribution for $p = p_1 - p_2$ as shown in Figure 1. Here *M* is the expected total number of users, given a substantial improvement in performance. *A* and *B* are two parameters which must be estimated before running the trial. Their values depend on the severity of the condition to be treated and on the expected cost of the new treatment.

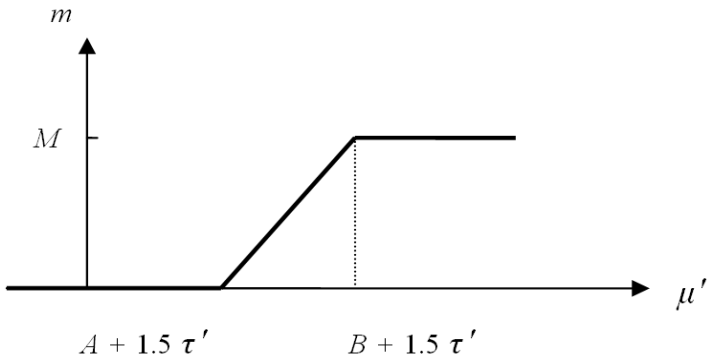

Figure 1- The number of subsequent users.

This function corresponds to assuming that each individual has a personal threshold difference, and will switch to the new treatment provided that the probability of the difference between the two treatments exceeding this threshold is at least 0.93. This is on the basis of normal approximation; the probability of a standard normal variate less than 1.5 is 0.93. Other figures, fairly near to 1.5 and 0.93 are equally plausible.

Using relationship (2) between μ_i' and x_i, m may be expressed as a function of $x_i\,(i=1,2)$.

4. The Objective Function

To keep our representation as straightforward as possible our discussion will be for the case when the number of patients on the treated group equals the number of patients on the control group, $n_1 = n_2 = n$. There are no essential difficulties in extending the methodology to the case when $n_1 \neq n_2$.

Using (4), it may be shown that the expected net benefit of conducting the trial may be written as:

$$r(n) = \sum_{x_1=0}^{n}\sum_{x_2=0}^{n} mb\prod_{i=1}^{2}\binom{n}{x_i}\frac{B(\alpha_i + x_i, \beta_i + n_i - x_i)}{B(\alpha_i, \beta_i)} - (c_1 n + c_2 n). \qquad (5)$$

In which *b* is the total benefit per user and $c_i\,(i=1,2)$ is the total cost per user in each group.

The computer program has been written to calculate the expected net benefit function $r(n)$ and the maximizing value n^* of n. Extensive numerical calculation shows that for all values of the various parameters expected net benefit $r(n)$ has a local maximum which is also the overall maximum.

For the case when $n_1 \neq n_2$ the computer program will obtain (n_1^*, n_2^*) which maximizes $r(n_1, n_2)$. This is, once again, the expected benefit of resulting change in the number of subsequent users of the new treatment minus the cost of carrying out two trials.

$$r(n_1,n_2)=\sum_{x_1=0}^{n_1}\sum_{x_2=0}^{n_2} mb\prod_{i=1}^{2}\binom{n_i}{x_i}\frac{B(\alpha_i+x_i,\beta_i+n_i-x_i)}{B(\alpha_i,\beta_i)}-(c_1n_1+c_2n_2). \tag{6}$$

A numerical routine written in C++ has been used to find the maximum of the above function.

4.1. A Clinical Trial

Losing hair is a common problem for those patients under chemotherapy for treating cancer. Without any treatment, 80% of patients lose hair. Consider a randomized clinical trial designed to detect a reduction to 33% in the proportion of patients losing their hair if they are treated with a certain cream. The question is what sample size is needed to maximize the expected net benefit of conducting the trial?

A trial along these lines was carried out by a United Kingdom-based pharmaceutical company, whose name has been omitted for commercial reasons. Since no direct information on prior distributions was available it was decided, after consulting the company, to make the following reasonably realistic assumptions.

The clinically relevant reduction is 47%. The mean μ_1 of the prior distribution for p_1 was assumed to be 0.33. The prior standard deviation, τ_1, for p_1 was assumed to be $\min(\mu_1, 1-\mu_1)/2 = 0.165$.

This is simply a convenient representation of a considerable degree of prior ignorance.(it follows that $\alpha_1 = 2.35$ and $\beta_1 = 4.77$). The mean μ_2 of the prior distribution for p_2 was assumed to be 0.20. The prior standard deviation, τ_2, for p_2 was assumed to be $\min(\mu_2, 1-\mu_2)/2 = 0.10$ (it follows that $\alpha_2 = 3$ and $\beta_2 = 12$). It was also assumed, and this seemed realistic, that the maximum critical difference for no sale is $A = 0.8\mu = 0.264$, and the minimum critical difference for maximum sale is $B = 1.2\ \mu = 0.396$. Since the trial was for an existing cream applied to a new class of patients, the current sales and profit margin were known and it was estimated that the expected net benefit, *Mb*, would be £ 5,000,000 (around $ 8,000,000) provided the trial was successful. The total cost c_1 per patient in the treatment group including hospital expenses and administration costs, was assumed to be £ 4,000 (around $ 6,400) and the cost c_2 per patient in the control group including hospital expenses and administration was assumed to be £ 2,000 (around $ 3,200).

The function $r(n)$ was maximized and the optimal (i.e. maximizing) sample size n^* for this trial was found to be 73, resulting in an expected net benefit $r(n^*)$ of £ $1.28M$ (around $ 2.01 M).

In fact the company concerned based the sample size on calculations of the *size* of the test (the chance of wrongly deciding that the new treatment is better) and the *power* of the test (the chance of detecting a clinically relevant difference when it is present). These lead to a sample size of 20.

It is noticeable that the optimal sample sizes based on our analysis are larger than the size used in practice. It should be noted here that the typical requirements for calculating the sample size in classical or frequentist framework are a size of 5% and a power of 80%. A trial designed along these lines has a certain scientific validity, in a rather narrow sense. There is no guarantee at all, however, that a trial of the size calculated actually should be carried out. Since there is no explicit attempt to balance the cost of carrying out the trial against the possible benefits of the new therapy, the classical procedure gives, in particular, no indication of those cases when the likely benefits do not

justify carrying out a trial at all, or of those cases when the possible benefits justify a lager trial, so as to increase the confidence of potential users in the new therapy, and thereby persuade more of them to use it.

References

Adcock, C.J. (1988) *A Bayesian Approach to Calculating Sample Sizes*. The Statistician, **37**, 433 – 439.

Adcock, C.J. (1992) *Bayesian approaches to the determination of sample sizes for binomial and multinomial sampling-some comments on the paper by Pham-Gia and Turkkan*. The Statistician, **41**, 399-404.

Adcock, C.J. (1995) *The Bayesian approach to the determination of sample size-some comments on the paper by Joseph, Wolfson, and Berger.* The Statistician, **44**, 155-161.

Desu, M. and Raghavarao, D. (1990) *Sample size methodology.* Academic Press, Boston.

Gittins, J.C. and Pezeshk H. (2000) *How large should a clinical trial be?* The Statistician, **49**, 177-187.

Gittins, J.C. and Pezeshk H. (2002) *A decision theoretic approach to sample size determination in clinical trials*. Journal of Biopharmaceutical Statistics, **12**(4), 535-551.

Grundy, P. M., Healy, M. J. R. and Rees D.H. (1956) *Economic choice of the amount of experimentation*. Journal of the Royal Statistical Society *A,* **18**, 32-38.

Joseph, L., Wolfson, D.B., and Du Berger R. (1995) *Sample size calculations of binomial proportions via highest posterior density intervals.* Statistics in Medicine, **14**, 143-154.

Joseph, L., Du Berger, R., and Belisle, P. (1997) *Bayesian and mixed Bayesian/likelihood criteria for sample size determination*. Statistics in Medicine**, 16**, 769-781.

Lindley, D.V. (1997) *The choice of sample size*. The Statistician**, 46**, 129-138.

Pezeshk, H. and Gittins, J.C. (1999) *Sample size determination in clinical trials*. Student, **3**, 19-26.

Pezeshk, H. and Gittins, J.C. (2002) *A fully Bayesian approach to calculating sample sizes for slinical trials with binary responses*. Drug Information Journal**, 36**, 143-150.

Pezeshk, H. (2003) *Bayesian techniques for sample size determination in clinical trials; a short review*. Statistical Methods in Medical research**, 12**(6), 489-504.

Pham-Gia, T. and Turkkan, N. (1992) *Sample size determination in Bayesian analysis*. The Statistician, **41**, 389-392.

Pham-Gia, T. and Turkkan, N. (2003) *Determination of exact sample sizes in the Bayesian estimation of the difference of two proportions*. The Statistician**, 52**, 131-150.

Raiffa, H., and Schlaifer, R. (1961) *Applied statistical decision theory.* Cambridge, MA, USA: Harvard University Graduate, School of Business Administration, Division of Research.