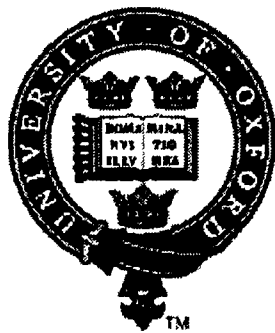


# How Large Should A Clinical Trial Be?

Hamid Pezeshk  
*Keble College*

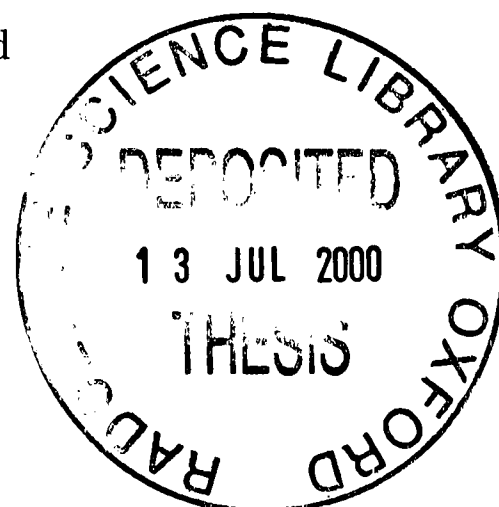


University of Oxford

Thesis submitted to the Faculty of Mathematical Sciences for the degree of  
Doctor of Philosophy of the University of Oxford.

Hillary Term 2000

Department of Statistics  
1 South Parks Road, Oxford



# How Large Should A Clinical Trial Be?

Hamid Pezeshk  
Keble College, Oxford

DPhil Thesis  
Hillary Term 2000

One of the most important questions in the planning of medical experiments to assess the performance of new drugs or treatments, is how big to make the trial. The problem, in its statistical formulation, is to determine the optimal size of a trial. The most frequently used methods of determining sample size in clinical trials is based on the required  $p$ -value, and the required power of the trial for a specified treatment effect. In contrast to the Bayesian decision theoretic approach there is no explicit balancing of the cost of a possible increase in the size of the trial against the benefit of the more accurate information which it would give. In this work we consider a fully Bayesian (or decision theoretic) 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, depend on the strength of the evidence provided by the trial. Our procedure differs from the usual Bayesian decision theory methodology, which assumes a single decision maker, by recognizing the existence of three decision makers, namely: the pharmaceutical company conducting the trial, which decides on its size; the regulator, whose approval is necessary for the drug to be licenced for sale; and the public at large, who determine the ultimate usage. Moreover, 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), the word Bayes referring to the optimization of the sample size. In the BeBay methodology the total expected benefit from carrying out the trial minus the cost of the trial is maximized. For any additional sales to occur as a result of the trial it must provide sufficient evidence both to convince the regulator to issue the necessary licence and to convince potential users that they should use the new treatment. The necessary evidence is in the form of a high probability after the trial that the new treatment achieves a clinically relevant improvement compared to the alternative treatment. The regulator is assumed to start from a more sceptical and less well-informed view of the likely performance of the treatment than the company carrying out the trial. The total benefit from a conclusively favourable trial is assessed on the basis of the size of the potential market and aggregated over the anticipated life-time of the product, using appropriate discounting for future years.

## Acknowledgements

I would like to thank my supervisor Professor John Gittins for initially suggesting this project to me and for his time, patience, encouragement, criticism and valuable advice and diligent supervision throughout the course of this work.

I am grateful to Mr. S. Day for his time and helpful discussions, and to Professor S. Senn for his interesting and stimulating comments on some parts of this work. In addition, thanks are due to Professor C. Adcock, Professor D. Altman, Dr. C. Baigent, Professor D. Lindley, Dr. M. Lunn, Dr. M. Parmar and Professor J. Whitehead for their many helpful comments.

I thank members of the Department of Statistics, Oxford, for their support and willingness to help. It is a particular pleasure to thank Ms Susan Hutchinson, departmental computing manager, for her prompt responses to all my computing queries. My thanks also go to Mrs Vivien Stchedroff for providing an efficient and friendly working environment. I will never forget her and her husband, Marcel, for their kindness to me and my family during our stay in Oxford.

I am also grateful to my parents and family. This dissertation would never have been conducted without the support and patience of my wife over the past three years. I am indebted to her and to my two little daughters, Golnoush and Mehrnoush, who have grown up during the preparation of this work.

The financial support of the Ministry of Culture and Higher Education of the Islamic Republic of Iran is gratefully acknowledged. The department of Mathematics and Computer Science at the University of Tehran kindly granted me the leave of absence from my teaching duties.

# Contents

<b>1</b>	<b>Introduction, Examples, Background, and Summary</b>	<b>6</b>
1.	Introduction . . . . .	6
2.	Examples . . . . .	8
3.	Background . . . . .	13
4.	Summary . . . . .	14
<b>2</b>	<b>A Review</b>	<b>16</b>
1.	Introduction . . . . .	16
2.	Frequentist methods of sample size determination; a review . . .	17
2.1.	Inference about the Variance of a Normal Distribution . .	22
2.2.	Multivariate Normal Distribution . . . . .	23
2.3.	Binomial Distribution . . . . .	24
2.4.	Multinomial Distribution . . . . .	26
2.5.	An Example . . . . .	28
3.	Bayesian methods of sample size determination; a review . . . .	29
3.1.	Bayesian and Mixed Bayesian/Likelihood Criteria . . . .	30
3.2.	Bayes Factor or Weiss’s Method . . . . .	32
3.3.	Methods of moments . . . . .	33
3.4.	Fully Bayesian Approach (Decision Theory) . . . . .	35
3.5.	An Example . . . . .	36
<b>3</b>	<b>A Simplified Introductory Model</b>	<b>41</b>
1.	Introduction and An Executive Summary . . . . .	41
2.	Sample Size Problem . . . . .	42
3.	A Bayesian Solution . . . . .	44
3.1.	Maximizing the Expected Net Benefit . . . . .	44
3.2.	The Parameters Determining Optimal Sample Size . . . .	50
<b>4</b>	<b>A Behavioural Bayes Approach to the Sample Size Question; Normal Distribution, Known Variance</b>	<b>53</b>
1.	Introduction . . . . .	53

2. The Number of Subsequent Users of the New Treatment . . . . .	54
3. A General Model . . . . .	57
3.1. Public Health Benefit Function . . . . .	58
3.2. Commercial Benefit Function . . . . .	61
4. Distribution of the Number of Subsequent Users of the New Treatment . . . . .	63
5. Range of Application . . . . .	66
6. Registration of the New Treatment . . . . .	69
6.1. Public Health Benefit Function . . . . .	70
6.2. Commercial Benefit Function . . . . .	72
7. The Regulator as a Third Decision-Maker . . . . .	73
7.1. The Regulator and the Public Health Benefit Function . . . . .	74
7.2. The Regulator and the Commercial Benefit Function . . . . .	79
8. Two Case Studies . . . . .	81
 <b>5 A Behavioural Bayes Approach to the Sample Size Question; Normal Distribution, Unknown Variance</b>	 <b>89</b>
1. Introduction . . . . .	89
2. Preliminaries . . . . .	90
3. Public Health Benefit Function . . . . .	92
4. Commercial Benefit Function . . . . .	95
5. Distribution of the Number of Subsequent Users of the New Treatment . . . . .	96
6. The Regulator as a Third Decision-Maker . . . . .	97
6.1. The Regulator and the Public Health Benefit Function . . . . .	98
6.2. The Regulator and the Commercial Benefit Function . . . . .	100
7. Convergence to the Known Variance Case . . . . .	103
7.1. An Example . . . . .	107
 <b>6 A Behavioral Bayes Approach to the Sample Size Question; Binomial Distribution</b>	 <b>109</b>
1. Introduction . . . . .	109
2. Notation . . . . .	110
3. Number of Subsequent Users of the New Treatment . . . . .	112
4. Utility Functions . . . . .	112
4.1. Public Health Benefit Function . . . . .	112
4.2. Commercial Benefit Function . . . . .	114

4.3. Example . . . . .	114
5. Distribution of the Number of Subsequent Users of the New Treatment . . . . .	116
6. The Regulator as a third decision-Maker . . . . .	118
6.1. The Regulator and the Public Health Benefit Function . . . . .	119
6.2. The Regulator and the Commercial Benefit Function . . . . .	119
6.3. Example . . . . .	120
<b>7 BeBay; A User-Guide</b>	<b>122</b>
1. Introduction . . . . .	122
2. Normally Distributed Data . . . . .	123
2.1. Data Input; Known Variance . . . . .	125
2.2. Results; Known Variance . . . . .	126
2.3. Illustration I . . . . .	127
2.4. Data Input; Unknown Variance . . . . .	132
2.5. Results; Unknown Variance . . . . .	133
2.6. Illustration II . . . . .	133
3. Binomially Distributed Data . . . . .	136
3.1. Results; Binary Data . . . . .	137
3.2. Illustration III . . . . .	138
4. Assessing the Truncation Error . . . . .	140
5. Numerical Precision . . . . .	146
6. A Summary of the BeBay Program Files . . . . .	149
<b>8 Desirable Further Work</b>	<b>151</b>
1. More General Assumptions . . . . .	152
2. A Note On the Cost Structure . . . . .	154
3. Numbers of Patients on Each of the Two Treatments; Known Variance . . . . .	155
3.1. Discussion . . . . .	157
4. Frequentist Regulatory Authorities . . . . .	157
4.1. Known Variance . . . . .	157
4.1.1. Public Health Benefit Function . . . . .	158
4.1.2. Commercial Benefit Function . . . . .	162
<b>Bibliography</b>	<b>163</b>

# Chapter 1

## Introduction, Examples, Background, and Summary

### 1. Introduction

When a new clinical treatment is being considered trials are carried out to estimate the increase in performance which is likely to result if the new treatment were to replace the treatment in current use. Many authors have looked at this problem and many procedures have been introduced to solve it. An important feature of the analysis in this work is that account is taken of the fact that only if it turns out that the new treatment, in a statistical sense, is clearly better than the previous treatment will the number of subsequent users of the new treatment be high.

In the planning of medical experiments to assess the performance of new drugs or treatments the question of how big to make the trial is an important issue. Traditional classical methods of selecting sample sizes are based on the required  $p$ -value, and the required power of the trial for a specified treatment effect. The most frequently used sample size formulae arise from the relationship between the standard error of the estimator of the parameter of interest and the sample size (Lachin (1981), Desu and Raghavarao (1990) Lemeshow *et al* (1990), and Lipsey (1990)).

Since the formulae can be highly sensitive to the choice of inputs, careful selection of the parameter estimates and target criteria are essential steps in determining the sample size. Classical or frequentist methods are unable to take into account uncertainty in point specifications. Bayesian methods are ideally suited for design since they provide a tool for specifying uncertainty,

and how it changes in response to further information.

In this work I consider a Bayesian approach to sample size determination for clinical trials in which the number of subsequent users of the therapy under investigation, and hence also the total benefit resulting from the trial, depend on the strength of the evidence provided by the trial.

The statistical significance of a difference between the performance of two treatments is one factor which may influence a decision to switch from one treatment to the other, but by itself it is not sufficient. What is really required is significant evidence that the performance of the new treatment improves on the performance of the old treatment by at least some minimum, and in general positive, difference.

Our objective function is based on the cost of the trial and the total benefit from the resulting change in the number of patients using the new treatment. The number of subsequent users of the new treatment is assumed to depend on the apparent difference between the performance of the new treatment and that of the treatment in current use. A general model with different features which allow for various contingencies is introduced.

Our procedure differs from the usual fully Bayesian or decision theoretic methodology, which assumes a single decision maker, by recognizing the existence of three decision makers, namely: the pharmaceutical company conducting the trial, which decides on its size; the regulator, whose approval is necessary for the drug to be licenced for sale; and the public at large, who determine ultimate usage. Moreover, we model the subsequent usage by plausible assumptions for actual behaviour, rather than assuming that it represents decisions which are in some sense optimal. For this reason the procedure may be called *Behavioural Bayes* (or BeBay for short), the word Bayes referring to the optimization of the sample size.

The following list gives the distinctive features of our BeBay procedure.

- The total expected benefit from carrying out the trial minus the cost of the trial is maximized.
- For any additional sales to occur as a result of the trial it must provide sufficient evidence both to convince the regulator to issue the necessary licence and to convince potential users that they should use the new treatment. The necessary evidence is in the form of a high probability



after the trial that the new treatment achieves a clinically relevant improvement compared to the alternative treatment.

- The regulator is assumed to start from a more sceptical and less well-informed view of the likely performance of the treatment than the company carrying out the trial.
- The number of patients who switch to the new treatment depends on the strength of the evidence provided by the trial, from zero up to some maximum number if the evidence is fairly conclusive in favour of the new treatment.
- The total benefit following a conclusively favourable trial is assessed on the basis of the size of the potential market, and aggregated over the anticipated life-time of the product, using appropriate discounting for future years.

## 2. Examples

Current practice for the determination of the size of a clinical trial is largely based on the theory for testing the hypothesis that the new treatment being tested is no better than the alternative treatment (or placebo) against the hypothesis that the new treatment is better by some specified, and clinically relevant, amount. Typical requirements are: at most a 5% chance of wrongly deciding that the new treatment is better; and at least an 80% chance of detecting the clinically relevant difference when it is present.

A test designed along these lines may turn out to be of a reasonably appropriate size, but there is no guarantee whatever that this will be so. Since there is no explicit attempt to balance the cost of carrying out the trial against the possible sales 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 larger trial, so as to increase the confidence of potential users of the new therapy, and thereby persuade more of them to use it. Our subsequent analysis of six different cases provide examples of both of these outcomes.

The following tables and figures set out the results of recalculating the sample sizes for six trials which were recently carried out, with sample sizes determined in the manner outlined in the first paragraph of this section. Trials 1, 2 and 3 were all for the same drug, with different groups of potential

users. Trials 4, 5 and 6 were all for different drugs. Trial 1 was for normally distributed data, and trials 2 to 6 for binomially distributed data. For the binomial data inference was based on a normal approximation to the distribution of the log-odds ratio of the cure rate with the new treatment to the cure rate with a placebo; this leads to the tabulated standard deviation of 2 for each observation. The cost per patient in the trial was assumed, on the basis of experience, to be £600 for out-patients and £4000 for hospitalized patients. For each trial four calculations of sample size were carried out, the results of which are shown in table 2, together with the expected power  $(1 - \beta)$  of an hypothesis test with 5% probability of detecting a non-existent difference in performance between the new treatment and placebo. The four calculations allow for high and low estimates for the strength of the evidence required by the regulator, and high and low estimates of the total benefit from a conclusively favourable trial. The low estimate for the evidence required by the regulator amounts to an assumption that the regulator's decision will be positive even when the apparent improvement from using the new treatment is barely enough to persuade potential customers to use it. This may well be a realistic assumption for at least some of the six cases, as in every case the drug concerned was already on the market, so it is plausible that approval for the additional use may have been obtainable more easily than would have been the case for a new drug.

The results, not surprisingly, show that the optimal sample size depends strongly on the expected benefit from a conclusively favourable outcome, and on the strength of the evidence required by the regulator. In practice it should be possible to estimate both of these much more accurately than the wide intervals which have been used here for illustrative purposes.

It is also noticeable that the optimal sample size based on our analysis is frequently much larger than the size used in practice. For large sample sizes an interim analysis is very desirable, as it has the potential both to reduce sampling costs and to advance the launch date of the new therapy. The results for trial 2 show that our analysis can indicate quite small sample sizes.

All the calculations reported here assume that there is a control group with the same number of patients as the group receiving the treatment.

Table 1: Data for Six Recent Trials

Trial Number:	1	2	3	4	5	6
s.d. of each observation	0.3	2	2	2	2	2
Company's prior mean	0.15	2.09	1.1	1.35	0.41	0.64
Company's prior s.d.	0.075	1.045	0.55	0.7	0.21	0.32
Regulator's prior mean	0	0	0	0	0	0
Regulator's prior s.d.	0.15	2.09	1.1	1.35	0.41	0.64
Critical Difference: max for no sale	0.12	1.67	0.9	1.1	0.33	0.51
Critical Difference: min for max sale	0.18	2.51	1.3	1.6	0.49	0.77
Critical Difference: min for licence(low)	0.1	1.50	0.8	1.0	0.27	0.48
Critical Difference: min for licence(high)	0.15	2.09	1.1	1.35	0.41	0.64
Cost per patient in trial, £	600	4000	600	4000	4000	600
Benefit from max sales(low), £M	25	5	25	15	15	15
Benefit from max sales(high),£M	250	25	250	150	150	150

In table 2 the following notation is used.

- $n$  = actual sample size used – calculated in the standard way.
- $n^*$  = optimal sample size calculated by BeBay.
- $R^*$  = maximized net expected benefit, calculated by BeBay.
- $1 - \beta$  = probability of detecting an improvement which is equal to the mean of the pharmaceutical company's prior distribution for the amount of improvement. This might be called expected power.

Table 2: Sample Sizes for the Six Trials

Trial 1		
$n = 100$		
	(Low) Licence Minimum	(High) Licence Minimum
Low Benefit	$n^* = 1319$ $R^* = £10.11M$ $1 - \beta = 1.00$	$n^* = 1350$ $R^* = £9.09M$ $1 - \beta = 1.00$
High Benefit	$n^* = 6165$ $R^* = £113.9M$ $1 - \beta = 1.00$	$n^* = 6270$ $R^* = £103.9M$ $1 - \beta = 1.00$

**Trial 2**  
 **$n = 20$**

	<i>(Low) Licence Minimum</i>	<i>(High) Licence Minimum</i>
<i>Low Benefit</i>	$n^* = 76$ $R^* = \pounds 1.57M$ $1 - \beta = 1.00$	$n^* = 78$ $R^* = \pounds 1.35M$ $1 - \beta = 1.00$
<i>High Benefit</i>	$n^* = 228$ $R^* = \pounds 9.57M$ $1 - \beta = 1.00$	$n^* = 233$ $R^* = \pounds 8.71M$ $1 - \beta = 1.00$

**Trial 3**  
 **$n = 270$**

	<i>(Low) Licence Minimum</i>	<i>(High) Licence Minimum</i>
<i>Low Benefit</i>	$n^* = 1243$ $R^* = \pounds 10.25M$ $1 - \beta = 1.00$	$n^* = 1270$ $R^* = \pounds 9.32M$ $1 - \beta = 1.00$
<i>High Benefit</i>	$n^* = 5807$ $R^* = \pounds 114.54M$ $1 - \beta = 1.00$	$n^* = 5892$ $R^* = \pounds 105.45M$ $1 - \beta = 1.00$

**Trial 4**  
 **$n = 36$**

	<i>(Low) Licence Minimum</i>	<i>(High) Licence Minimum</i>
<i>Low Benefit</i>	$n^* = 207$ $R^* = \pounds 4.94M$ $1 - \beta = 1.00$	$n^* = 213$ $R^* = \pounds 4.36M$ $1 - \beta = 1.00$
<i>High Benefit</i>	$n^* = 981$ $R^* = \pounds 63.03M$ $1 - \beta = 1.00$	$n^* = 1001$ $R^* = \pounds 57.58M$ $1 - \beta = 1.00$

**Trial 5**  
 **$n = 470$**

	<i>(Low) Licence Minimum</i>	<i>(High) Licence Minimum</i>
<i>Low Benefit</i>	$n^* = 403$ $R^* = £2.06M$ $1 - \beta = 0.74$	$n^* = 399$ $R^* = £1.39M$ $1 - \beta = 0.73$
<i>High Benefit</i>	$n^* = 2128$ $R^* = £48.58M$ $1 - \beta = 1.00$	$n^* = 2187$ $R^* = £42.40M$ $1 - \beta = 1.00$

**Trial 6**  
 **$n = 400$**

	<i>(Low) Licence Minimum</i>	<i>(High) Licence Minimum</i>
<i>Low Benefit</i>	$n^* = 1245$ $R^* = £5.22M$ $1 - \beta = 1.00$	$n^* = 1278$ $R^* = £4.58M$ $1 - \beta = 1.00$
<i>High Benefit</i>	$n^* = 5885$ $R^* = £64.37M$ $1 - \beta = 1.00$	$n^* = 6007$ $R^* = £58.21M$ $1 - \beta = 1.00$

In the above tables the *Low Benefit* and the *High Benefit* refer, respectively, to the last two rows of table 1. *(Low) Licence Minimum* represents a low regulator's requirement while *(High) Licence Minimum* stands for the regulator with higher requirement concerning the apparent difference between the performances of the two treatments.

Figure 1.1 illustrates the variation of the expected net benefit as a function of the size of the trial.

For each trial we provide two graphs representing *low benefit from maximum sale* and *high benefit from maximum sale* of the new treatment. In each case there are two curves corresponding to the low and intermediate regulator requirements. Not surprisingly, with a lower regulator requirement the expected net benefit is higher.

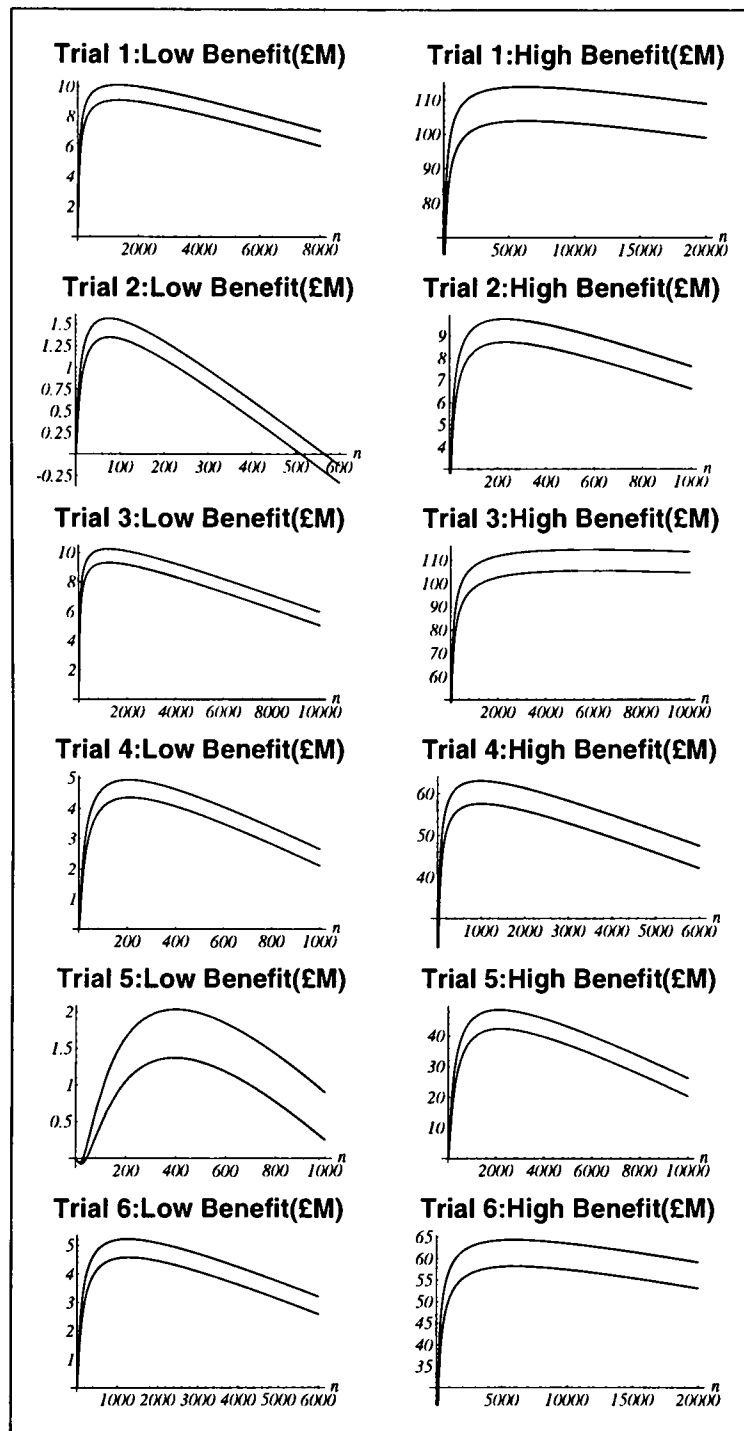


Figure 1.1: Expected net benefit.

### 3. Background

Suppose the random variable  $X$  has a probability distribution function  $F(\cdot | \theta)$  of known form, and dependent on an unknown parameter  $\theta$ . Let the observations  $X_1, X_2, \dots, X_n$  be independently and identically distributed with the probability distribution function  $F(\cdot | \theta)$ .

**Definition:** The sample size determination problem is to choose the sample

size  $n$  so that inferences and decisions about the parameter  $\theta$  may be made.

Sample size determination is directly related to the cost of a survey or data gathering exercise. The quality of the inference about  $\theta$  and all decisions made clearly have potential implications for subsequent events related to the system of interest. In general terms, the aim of all sample size determination is to allow the investigator to determine the sample size by specifying a target level of accuracy or by maximizing a specified objective function. However, as noted by Adcock (1997), there are several complications. First, nuisance parameters may be present. This is the case, for example, if the estimation of the mean of a normal distribution is of interest, but the variance is unknown. Secondly, the problem under investigation may have more than one parameter of interest. For example, it may be of interest to estimate both the mean and the variance of a normal distribution. Thirdly, we may have more than one variable, and hence many unknown parameters. Fourthly, the probability distribution of the random quantity  $X$  may not be known. In short, the general sample size determination problem is potentially complicated and many dimensional.

The different criteria which are used for sample size determination may be classified into two groups: Bayesian and frequentist. Bayesian methods may be further divided into two groups of techniques; methods which are inferential and employ power calculations, hypothesis testing, and confidence intervals (see, for example, Joseph *et al* (1997), Spiegelhalter and Freedman (1986), and Gardner and Altman (1986)) and fully Bayesian or decision theoretic methods which are based on a loss or utility function. The fully Bayesian methodology was first introduced by Grundy *et al* (1956), and set out in detail by Raiffa and Schlaifer (1961). Lindley (1997) provides a recent example of its use. In the frequentist approach sample sizes may be determined by specifying the length of the resulting confidence interval, the probability of a given absolute error, or the power and size of the resulting hypothesis test (see, for example, Desu and Raghavarao (1990)). The book by Senn (1997) gives an extensive account of terminology and statistical methods, including sample size determination methods, used in clinical trial and drug development issues.

## 4. Summary

The thesis is organized as follows

In chapter 2 some of the frequentist and Bayesian approaches to the sample size determination problem are briefly reviewed.

Chapter 3 addresses a simplified introductory version of BeBay with no regulator. An exact solution is derived for the case where the number of subsequent users of the new treatment is proportional to the number of standard deviations by which the posterior mean of the unknown performance parameter exceeds the performance of the current treatment.

The following publication is based on chapter 3.

- Sample Size Determination in Clinical Trials. *Student* (1999), **3**, 19–26.

Chapters 4 and 5 discuss the full model for normally distributed observations. In chapter 4 the experimental variance of the observations is assumed to be known, and in chapter 5 it is assumed to be unknown. In chapter 5 a joint prior density for the parameter of interest and the variance of the observations is assumed.

Sections 1 to 3 and section 5 of chapter 4 have formed the basis for the following publication

- How Large Should a Clinical Trial be? *The Statistician* (2000), **49**, part 2, pp. 177–187.

The contents of chapters 1 and 5 together with section 7 of chapter 4 have formed the basis for the following publication

- A Behavioural Bayes Method for Determining the Size of a Clinical Trial. *Drug information Journal* (2000), Vol. 34, pp. 355–363.

Chapter 6 discusses the case of binomially distributed observations.

Chapter 7 is a user-guide for the *Mathematica*-based routines which we have used in this work. The accuracy of the calculations is also discussed in this chapter.

Chapter 8 outlines the way forward for desirable further work.



# Chapter 2

## A Review

### 1. Introduction

The determination of sample size is of key importance in designing medical studies. There have been a number of papers, from both the frequentist and Bayesian points of view, on this subject, a representative selection of which will be reviewed in this chapter. Not all of these papers were written with medical applications primarily in mind.

In section 2 some of the outlines of the frequentist methods of sample size determination are reviewed. As we shall see these methods could be classified into two groups: absolute error and power and size control rules. Current practice for the determination of the size of a clinical trial is largely based on the power and size control rules (see, for example Lemeshow *et al* (1990)). There is a comprehensive summary of power and size control rules in Desu and Raghavarao (1990) and Cohen (1977). Both Kraemer and Thiemann (1987) and Desu and Raghavarao (1990) present overviews and methods for numerous problems. There is an extensive treatment of sample size determination for regression in Muller *et al* (1992).

Section 3 briefly reviews the available Bayesian sample size determination methods. As already mentioned these methods may be classified into two groups: inferential and decision theoretic rules. Some of the key papers related to these two groups of techniques will be reviewed. Apart from the paper by Joseph and Bélisle (1997) which discusses some inferential rules for normally distributed data, there are other papers on Bayesian inference (see, for example, Adcock (1987), Pham-Gia and Turkkan (1992), Spiegelhalter *et al* (1994)). Adcock (1997) reviews both inferential and fully Bayesian (or

decision theoretic) rules. There is a detailed description of decision theoretic rules for the determination of sample sizes in Raiffa and Schlaifer (1961).

It should be noted that this chapter does not seek to contribute to the debate between Bayesians and frequentists. Our aim is simply to review the available methodology. There are so many papers on the subject that it is almost impossible to achieve a comprehensive review. Efforts have been made not to leave out any important approach. But in the words of Smith (1976), ‘no review can ever replace a thorough reading of a fundamental paper, since what is left out is often as important as what is included’.

## 2. Frequentist methods of sample size determination; a review

Sample size determination for inference about the mean of a normal distribution with known variance is the most simple case of sample size determination. If  $X \sim N(\mu, \sigma^2)$  then, following Desu and Raghavarao (1990), p.3, we say that the mean  $\mu$  may be estimated with absolute error  $e$  and probability  $1 - \alpha$  by the sample mean  $\bar{X}_n$  if

$$Pr[|\bar{X}_n - \mu| \leq e | \mu] \geq 1 - \alpha. \quad (2.1)$$

Since  $n^{1/2}(\bar{X}_n - \mu)/\sigma \sim N(0, 1)$  it follows that inequality (2.1) is satisfied when the sample size  $n$  satisfies

$$n \geq \sigma^2 z_{\alpha/2}^2 / e^2, \quad (2.2)$$

where  $z_\alpha$  denotes the upper  $100(1 - \alpha)\%$  point of the standard normal distribution. Alternatively, if a  $100(1 - \alpha)\%$  confidence interval for  $\mu$  is required of length  $l$ , then provided that the variance  $\sigma^2$  is known, the sample size must satisfy the following inequality

$$n \geq 4\sigma^2 z_{\alpha/2}^2 / l^2, \quad (2.3)$$

(see, for example, Joseph and Bélisle (1997)). An extension to the rule given by inequality (2.2) comes from the theory of hypothesis testing.

Suppose that we have

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu = \mu_1 \neq \mu_0, \end{cases}$$

and that the test is to be of size  $\alpha$  and with a type II error equal to  $\beta$  when  $|\mu_1 - \mu_0| = e$ . For a two-sided test the critical region is

$$|\bar{X}_n - \mu_0| > z_{\alpha/2} \left( \frac{\sigma}{\sqrt{n}} \right).$$

The required sample size  $n$  is chosen so that

$$Pr(|\bar{X}_n - \mu_0| > z_{\alpha/2} \left( \frac{\sigma}{\sqrt{n}} \right) | \mu = \mu_1) = 1 - \beta.$$

This can be written as

$$\begin{aligned} Pr\left(\frac{\bar{X}_n - \mu_1}{\sigma/\sqrt{n}} < -z_{\alpha/2} - \frac{(\mu_1 - \mu_0)}{\sigma/\sqrt{n}}\right) \\ + 1 - Pr\left(\frac{\bar{X}_n - \mu_1}{\sigma/\sqrt{n}} < z_{\alpha/2} - \frac{(\mu_1 - \mu_0)}{\sigma/\sqrt{n}}\right) = 1 - \beta. \end{aligned} \quad (2.4)$$

If  $\mu_1 > \mu_0$ , then the first probability on the left-hand side of (2.4) is close to 0, so that an approximate value of  $n$  may be found by solving the equation

$$z_{\alpha/2} - \frac{(\mu_1 - \mu_0)}{\sigma/\sqrt{n}} = -z_{\beta}.$$

On the other hand, if  $\mu_1 < \mu_0$ , then the second probability term on the left-hand side of (2.4) is close to 1, so that an approximate value of  $n$  may be found by solving the equation

$$-z_{\alpha/2} - \frac{(\mu_1 - \mu_0)}{\sigma/\sqrt{n}} = z_{\beta}.$$

These two equations can be combined into one equation

$$z_{\alpha/2} + z_{\beta} = \frac{|\mu_1 - \mu_0|}{\sigma/\sqrt{n}} = \frac{e}{\sigma\sqrt{n}},$$

giving

$$n = (z_{\alpha/2} + z_{\beta})^2 \sigma^2 / e^2. \quad (2.5)$$

The above is an example of a power and size control rule. Such rules are in general constructed as follows. Null and alternative hypotheses are both

specified for the parameter of interest  $\theta$ . The size and power parameters  $\alpha$  and  $1 - \beta$  are also specified. The sample size  $n$  and a rejection value for the null hypothesis are then chosen simultaneously to satisfy the two equations

$$\begin{aligned} Pr(\text{reject } H_0 \mid H_0 \text{ is true}) &= \alpha, \\ Pr(\text{reject } H_0 \mid H_1 \text{ is true}) &= 1 - \beta. \end{aligned} \quad (2.6)$$

These rules, which are used for the *one-sample problem*, may be extended to other common situations, such as the difference of two means, for which the problem of sample size determination is called the *two-sample problem*.

Formulae (2.2) and (2.5) are also applicable as approximations for distributions other than normal, for cases where the central limit theorem ensures that  $\bar{X}_n$  has an approximately normal distribution and an estimate for  $\sigma^2$  is available.

When  $\sigma^2$  is unknown one possibility is to specify the acceptable error in units of  $\sigma$ . Thus replacing  $e$  by  $e\sigma$ , formulae (2.2) and (2.5) become, respectively,  $n \geq z_{\alpha/2}^2/e^2$  and  $n = (z_{\alpha/2} + z_{\beta})^2/e^2$ .

In many applications it is not appropriate to measure the acceptable error in units of  $\sigma$ , and in such circumstances it is necessary to use an estimate of the unknown variance. Stein (1945) developed the following two-stage procedure. Take an initial sample of size  $n_0$  from the distribution and estimate the variance in the usual way by the sample variance,  $S_0^2$ . If

$$n_0 \geq S_0^2 t_{\alpha/2}^2 (n_0 - 1) / e^2 \quad (2.7)$$

then stop sampling. Otherwise draw a further sample of size

$$n = S_0^2 t_{\alpha/2}^2 (n_0 - 1) / e^2 - n_0, \quad (2.8)$$

where  $t_{\alpha}(n_0 - 1)$  denotes the upper  $100(1 - \alpha)\%$  point of the Student  $t$ -distribution with  $n_0 - 1$  degrees of freedom. The same analysis may be applied to show that for hypothesis testing the required size of trial is

$$n = n_0 + \max\{0, S_0^2 [t_{\alpha/2}(n_0 - 1) + t_{\beta}(n_0 - 1)]^2 / e^2 - n_0\}, \quad (2.9)$$

(see, for example, Desu and Raghavarao (1990), p.9).

The above procedure could be interpreted as a sequential scheme for which

initially  $n_0$  observations are made and thereafter observations are made one at a time and the sample variance is updated, until the stopping rule (2.7) is satisfied with the current estimate of variance and the current sample size. A variation of this rule is described by Liu (1997).

It should be noted that the choice of the initial sample size  $n_0$  is very important. If it is very small, then it is very likely that a second sample will have to be drawn. If it is very large, then unnecessary expenditure is occurred. Seelbinder (1953) presented a method for determining the size of the first part of a two-stage sample,  $n_0$ , for estimating the population mean of a normal distribution with a given precision. His approach, which uses the absolute error criterion, assumes  $e/\sigma$  is known. He then minimizes the expectation of the total sample size given  $n_0$  to obtain  $n$ . Moshan (1957) extends this method and proposes a guide to the selection of  $n_0$  by using an upper percentage point of the distribution of total sample size,  $n$ , in conjunction with its expectation.

The above power and size control methods use a pivot function for sample size calculations. This is a function of the observations and of the unknown parameter(s) whose distribution does not depend on the unknown parameter(s). In the above example the quantity  $Z = n^{1/2}(\bar{X}_n - \mu_0)/\sigma$  is a pivot, because it depends on  $\mu_0$  but its distribution,  $N(0, 1)$ , does not depend on  $\mu_0$ . When no pivot function exists some approximation procedure is typically needed. A more general problem is that this method of sample size determination requires users to carry out computations that are specific to the problem under consideration. Potentially this is a serious practical obstacle. Kraemer and Thiemann (1987) avoid this in an elegant way by reducing a large number of problems to a small number of standard forms. The basis of their method is twofold. First, they work out the problem of sample size determination for the correlation coefficient  $\rho$  in a bivariate normal distribution. Secondly, they present a series of methods for different sample size determination problems and transform each to the procedure for sample size determination for the correlation coefficient.

**Remark:** If a study is designed to test the difference between two normal means e.g.,  $H_1 : \mu_1 - \mu_2 = 0$  vs  $H_2 : \mu_1 - \mu_2 \neq 0$  then following a strategy similar to that employed in developing (2.5), it follows that

$$n = 2(z_{\alpha/2} + z_{\beta})^2 \sigma^2 / e^2.$$

As already mentioned, this is called *the two-sample problem* (see, for example, Lemeshow *et al* (1990)). It should be noted that here we assume the variances are known and **homoskedastic** (i.e.,  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ) and, in

addition, the same number of observations is selected from each of the two populations i.e.,  $n_1 = n_2 = n$ .

**Remark:** For the case of unknown variances, where  $\sigma_1^2 \neq \sigma_2^2$ , Koopmans and Qualls (1971) suggested a two-stage procedure using the absolute error criterion for estimating  $\mu_1 - \mu_2$ , to meet the requirement

$$P_r\{ |(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)| \leq e \} \geq 1 - \alpha, \quad (2.10)$$

where  $\bar{X}$  and  $\bar{Y}$  are the means of samples taken, respectively, from  $N(\mu_1, \sigma_1^2)$  and  $N(\mu_2, \sigma_2^2)$ .

Their procedure is described as follows

- (i) Take first stage independent samples of sizes  $n_{11}$  ( $\geq 2$ ) and  $n_{21}$  ( $\geq 2$ ) from the two populations. Let  $s_{11}^2$  and  $s_{21}^2$  be the two variances of these samples.
- (ii) Calculate upper  $100(\sqrt{1 - \alpha/2})\%$  confidence limits on  $\sigma_1^2$  and  $\sigma_2^2$ . Let  $\sigma_{1u}^2$  and  $\sigma_{2u}^2$  be these limits.
- (iii) Let  $C = c_0 + c_1 n_{12} + c_2 n_{22}$  be the linear cost function, where  $n_{12}$  and  $n_{22}$  are the second stage sample sizes from the two populations. The solutions of  $n_{12}$  and  $n_{22}$  that minimizes  $C$  subject to

$$z_{\alpha/4} \sqrt{\frac{\sigma_{1u}^2}{n_{12}} + \frac{\sigma_{2u}^2}{n_{22}}} \leq e$$

are

$$\begin{aligned} n_{12} &= \frac{z_{\alpha/4}^2}{e^2} \frac{\sqrt{c_1} \sigma_{1u} + \sqrt{c_2} \sigma_{2u}}{\sqrt{c_1}} \sigma_{1u} + 1, \\ n_{22} &= \frac{z_{\alpha/4}^2}{e^2} \frac{\sqrt{c_1} \sigma_{1u} + \sqrt{c_2} \sigma_{2u}}{\sqrt{c_2}} \sigma_{2u} + 1. \end{aligned} \quad (2.11)$$

- (iv) Let  $\bar{X}_2$  and  $\bar{Y}_2$  be the two sample means for the second stage samples. Then  $\bar{X}_2 - \bar{Y}_2$  as an estimator of  $\mu_1 - \mu_2$  satisfies (2.10).

## 2.1. Inference about the Variance of a Normal Distribution

Suppose that we are concerned with estimating the variance of a normal distribution  $\sigma^2$ , for which the mean,  $\mu$ , is unknown. If we use an unbiased estimate,  $S^2 = \frac{1}{n-1} \sum (x_i - \bar{x}_n)^2$ , then  $(n-1)s^2/\sigma^2$  could be considered as a pivot function for constructing a confidence interval or testing a statistical hypothesis concerning  $\sigma^2$ .

The absolute error requirement in (2.1) now may be changed to percentage error

$$Pr[|S^2/\sigma^2 - 1| \leq e | \sigma^2] \geq 1 - \alpha.$$

Writing  $X^2 = \frac{(n-1)S^2}{\sigma^2}$  and rearranging the above inequality gives the sample size as the solution to

$$Pr[(n-1)(1-e) \leq X^2 \leq (n-1)(1+e)] \geq 1 - \alpha, \quad (2.12)$$

where  $X^2$  is distributed as a chi-squared with  $n-1$  degrees of freedom, which may be denoted as  $\chi^2_{(n-1)}$ . Since the inequality (2.12) should be solved numerically, one may apply a rough approximation by using the Central Limit Theorem to obtain  $n$ . The distribution on  $(\chi^2_{(\nu)} - \nu)/\sqrt{2\nu}$  is asymptotically  $N(0, 1)$ . Now (2.12) may be approximated by

$$Pr[-e\sqrt{\frac{n-1}{2}} \leq Z \leq e\sqrt{\frac{n-1}{2}}] \geq 1 - \alpha,$$

where  $Z$  is the standard normal random variable. This can be easily solved for  $n$ , resulting in the following formula

$$n = \frac{2z_{\alpha/2}^2}{e^2} + 1. \quad (2.13)$$

**Remark:** If a study is designed to test a (commonly tested) hypothesis;  $H_0 : \sigma = \sigma_0$  vs  $H_1 : \sigma = \sigma_1 (> \sigma_0)$ , then for a power of  $1 - \beta$  when  $\sigma = \sigma_1 (> \sigma_0)$ ,  $n$  must be chosen so that

$$Pr\left\{S^2 > \frac{\sigma_0^2 \chi^2_{(n-1),\alpha}}{n-1} \mid \sigma^2 = \sigma_1^2\right\} = 1 - \beta,$$

where  $\chi^2_{(n-1),\alpha}$  denotes the upper  $100(1-\alpha)\%$  point of the chi-squared with  $n-1$  degrees of freedom. Following the power and size control rule strategy, this leads to the relation

$$\sigma_0^2 \chi^2_{(n-1),\alpha} = \sigma_1^2 \chi^2_{(n-1),\beta}, \quad (2.14)$$

which needs to be solved for  $n$ .

It should be noted that the  $\chi^2$  percentiles can be approximated by normal percentiles (see, for example, Desu and Raghavarao (1990), p.11). So (2.14) could be solved approximately to give

$$n = \frac{1}{2} \left( \frac{\left(\frac{\sigma_0^2}{\sigma_1^2}\right) z_\alpha + z_\beta}{\left(\frac{\sigma_0^2}{\sigma_1^2}\right) - 1} \right)^2 + \frac{3}{2}. \quad (2.15)$$

## 2.2. Multivariate Normal Distribution

The principles of sample size determination methods discussed at the beginning of this chapter can be extended to multivariate problems.

Consider a  $k$ -variate vector  $\mathbf{X}$  which is normally distributed with unknown vector mean  $\boldsymbol{\mu}$  and known variance-covariance matrix  $\boldsymbol{\Sigma}$ . The direct extension of the absolute error approach for the error vector  $\mathbf{e}$  may be written as

$$Pr\left[\bigcap_{i=1}^k (|\bar{X}_{in} - \mu_i| \leq e_i) \mid \mu_i, i = 1, 2, \dots, k\right] \geq 1 - \alpha, \quad (2.16)$$

where the  $\bar{X}_{in}$ 's are the elements of the sample mean vector  $\bar{\mathbf{X}}_n$ , and the  $\mu_i$ 's and  $e_i$ 's are, respectively, the elements of the mean vector  $\boldsymbol{\mu}$  and the error vector  $\mathbf{e}$ . The region  $R = \bigcap_{i=1}^k (|\bar{x}_{in} - \mu_i| \leq e_i)$  is known as a hyper-cubic region. The requirement (2.16) means that we must evaluate the general multivariate normal integral. As pointed out by Adcock (1997), an alternative is to consider elliptical regions, like

$$(\bar{\mathbf{X}}_n - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{X}}_n - \boldsymbol{\mu}) \leq e^2. \quad (2.17)$$

This leads to the rule

$$n \geq \chi_{(k),\alpha}^2 / e^2, \quad (2.18)$$

where, once again,  $\chi_{(k),\alpha}^2$  denotes the upper  $100(1 - \alpha)\%$  point of the chi-squared distribution with  $k$  degrees of freedom.

For elliptical regions and the case when  $\boldsymbol{\Sigma}$  is known it is straightforward to extend the power and size control methodology. However, if  $\boldsymbol{\Sigma}$  is unknown the problem of sample size determination becomes more complicated. In the absolute error procedure, one possibility may be to employ a multivariate version of Stein's two-stage method. Another sequential scheme may be found in Liu (1997).



## 2.3. Binomial Distribution

The determination of sample size for the binomial distribution is one of the oldest cases for which there is a formal procedure (see, for example, Cochran (1963)). The normal approximation to the binomial distribution may be used. If  $X \sim b(n, p)$  then the sample proportion  $\hat{p} = \frac{X}{n}$  is approximately normally distributed with mean  $p$  and variance  $\frac{p(1-p)}{n}$ . The absolute error for finding the sample size may now be written as

$$Pr[|\hat{p} - p| \leq e|p] \geq 1 - \alpha. \quad (2.19)$$

Since  $\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$  has, approximately, a standard normal distribution, one can conclude  $[\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}]^2$  has, approximately, a  $\chi^2_{(1)}$  distribution. Solving for  $n$  in (2.19) gives

$$n \geq \frac{p(1-p)\chi^2_{(1),\alpha}}{e^2}. \quad (2.20)$$

As noted by Adcock (1987), the rule at (2.20) has two limitations. First, some problems are multinomial, that is the number of classes is greater than two. Secondly, the population parameter  $p$  is unknown. The second limitation may be avoided by using the maximum value of  $p(1-p)$  which gives the rule

$$n \geq \frac{\chi^2_{(1),\alpha}}{4e^2},$$

which is conservative in the sense that the probability associated with it is at least  $1 - \alpha$ . Another alternative is to use an estimate of  $p$  which might have come from earlier related studies.

The use of absolute error is questionable. A given value  $\pm e$  when the true value of  $p$  is close to  $1/2$  is a more stringent requirement than when  $p$  is close to 0. If we elect to use the percentage error and require that

$$Pr[|\hat{p} - p| \leq ep|p] \geq 1 - \alpha \quad (2.21)$$

then assuming the normal approximation, we obtain

$$n \geq \left(\frac{1}{p} - 1\right)\chi^2_{(1),\alpha}/e^2. \quad (2.22)$$

As pointed out by Tortora (1978), this does not seem completely sensible either. As  $p \rightarrow 1$  the sample size tends to 0 and, as  $p \rightarrow 0$ , the sample size is

unbounded. Furthermore, if the true value of  $p$ , for example, is 0.075, then estimates of 0.05 and 0.1 give sample sizes which differ considerably for all values of  $\alpha$  and  $e$ . So the ability to take the uncertainty of the estimate of  $p$  into account becomes important.

**Remark:** A simple criterion that avoids the problems that arise when  $p$  is close to its maximum or minimum is to use the approximate pivot function  $Z = n^{1/2}(\hat{p} - p)/\{p(1 - p)\}^{1/2}$  in another way. Thus if we require that

$$Pr[|\hat{p} - p| \leq e\{p(1 - p)\}^{1/2}|p] \geq 1 - \alpha, \quad (2.23)$$

then using the normal approximation we obtain the rule

$$n \geq \chi_{(1),\alpha}^2/e^2. \quad (2.24)$$

The following remarks demonstrate how one may apply the power and size control method to obtain the sample size in binomial sampling.

**Remark:** The required sample size for testing the hypothesis  $H_0 : p = p_0$  vs  $H_1 : p \neq p_0$ , which is known as the one-sample problem in binomial sampling, may be calculated from

$$n = \left\{ z_{\alpha/2} \sqrt{p_0(1 - p_0)} + z_{\beta} \sqrt{p(1 - p)} \right\}^2 / e^2, \quad (2.25)$$

where  $e = |p - p_0|$ .

**Remark:** For testing the hypothesis concerning two population proportions i.e.,  $H_0 : p_1 = p_2$  vs  $H_1 : p_1 \neq p_2$ , which is known as the two-sample problem in binomial sampling, the required sample size may be calculated from

$$n = \left\{ z_{\alpha/2} \sqrt{2\bar{p}(1 - \bar{p})} + z_{\beta} \sqrt{p_1(1 - p_2) + p_2(1 - p_1)} \right\}^2 / e^2, \quad (2.26)$$

where  $\bar{p} = \frac{p_1 + p_2}{2}$  and  $e = |p_2 - p_1|$ .

It should be noted that in (2.26) we assume that the common value of  $p_1$  and  $p_2$  is denoted by  $p$  and  $n_1 = n_2 = n$ . So  $V(\hat{p}_1 - \hat{p}_2) = 2[p(1 - p)/n]$ , which depends on  $p$ . This may be estimated as the average of the two-sample proportions  $\bar{p}$  from the pilot or previous studies.

Applying the same analysis as in (2.4) we may arrive at the following equations

$$\frac{z_{\alpha/2} \sqrt{\frac{2\bar{p}(1-\bar{p})}{n}} - (p_2 - p_1)}{\sqrt{p_1(1-p_1)/n + p_2(1-p_2)/n}} = -z_{\beta}$$

$$\frac{-z_{\alpha/2} \sqrt{\frac{2\bar{p}(1-\bar{p})}{n}} - (p_2 - p_1)}{\sqrt{p_1(1-p_1)/n + p_2(1-p_2)/n}} = z_{\beta}$$

Putting  $e = |p_2 - p_1|$  and solving for  $n$  yields (2.26).

□

Traditional methods of determining sample sizes for testing the equality of proportions with a specified type *I* and *II* error probabilities are based on approximations to the binomial distribution. When proportions involved are small, the exact distribution of the test statistic may not follow the assumed distribution (e.g., normal or chi-square). Consequently, use of the test statistic with the sample size determined in the traditional way, as in (2.26), may result in a test with type *I* and type *II* error probabilities that differ from the values which have been specified.

A method for evaluating the exact type *I* and type *II* error probabilities of test under these circumstances was described by Lemeshow *et al* (1981). Their method has been applied to the case of a one-sided alternative hypothesis. They described a procedure that adjusts the sample size in an iterative manner until the value  $\beta$  for the test falls within a specified percent of the derived value. They showed that the *arcsin* approximation appears to give the closest approximation to the necessary sample sizes in the cases where the probability of occurrence of the event is small. This leads to the following commonly used formula (see, for example, Cochran and Cox (1957)) for a size  $\alpha$ , one-sample problem of testing  $H_0 : p = p_1$  versus  $H_1 : p = p_2 (> p_1)$

$$n = \frac{(z_{\alpha} + z_{\beta})^2}{2(\arcsin\sqrt{p_2} - \arcsin\sqrt{p_1})^2}. \quad (2.27)$$

## 2.4. Multinomial Distribution

The analysis for sample size determination in binomial sampling can be extended to multinomial sampling. Suppose that the random vector  $\mathbf{X} =$

$(X_1, X_2, \dots, X_k)$  has a density function of the form

$$f(x_1, x_2, \dots, x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}, \quad x_i = 0, 1, \dots, n, \quad i = 1, 2, \dots, k,$$

where  $p_i \geq 0$  ( $i = 1, 2, \dots, k$ ),  $\sum_{i=1}^k p_i = 1$ ,  $\sum_{i=1}^k x_i = n$ .

If  $\hat{p}_i = \frac{x_i}{n}$ , ( $i = 1, 2, \dots, k$ ) are the observed proportions, then an extension to the absolute error criterion is to find  $n$  so that

$$Pr\left[\bigcap_{i=1}^k (|\hat{p}_i - p_i| \leq e_i) \mid p_i, i = 1, 2, \dots, k\right] \geq 1 - \alpha. \quad (2.28)$$

When  $e_i = e$ ,  $\forall i$ , the Bonferroni inequality leads to the rule

$$n \geq \max_i \{p_i(1 - p_i) \chi_{(1), \alpha/k}^2 / e^2\},$$

which leads to the conservative rule

$$n \geq \frac{\chi_{(1), \alpha/k}^2}{4e^2}.$$

If  $e_i$  in inequality (2.28) is replaced by  $e\{p_i(1 - p_i)\}^{1/2}$  we obtain a simpler rule

$$n \geq \frac{\chi_{(1), \alpha/k}^2}{e^2}.$$

However, as noted by Thompson (1987), both these rules lead to a sample size which increases when the number of classes increases. He resolved this problem by showing that

$$n \geq \max_{k \geq 1} \left\{ \frac{1/k(1 - 1/k) \chi_{(1), \alpha/k}^2}{e^2} \right\} \quad (2.29)$$

satisfies (2.28), when  $e_i = e$ ,  $\forall i$ . He provided a look-up table which is valid for any number of classes  $k$ . The maximum of equation (2.29) occurs at low values of  $k$  and so the sample size does not increase indefinitely with the number of classes. A discussion about the limitations of Thompson's procedure, along with another way of satisfying (2.28), may be found in Sison and Glaz (1995).

Other contributions to frequentist sample size determination have been made by Feigel (1978), Haseman (1978), Whittemore (1981), Thigpen (1987), Miller (1990), Signorini (1991), and Singer (1997).

Before closing this section, let us consider the following example which demonstrates the use of confidence intervals in determining sample sizes for inference about the difference between the success rates of two binomial distributions.

## 2.5. An Example

Consider the design of a randomized clinical trial to compare the (success) rates of two different treatments. What sample size is needed to provide sufficient information to specify the difference between the two rates within a confidence interval of width  $e$ ?

The most common solution to this question for a  $100(1 - \alpha)\%$  confidence interval (CI) is to set

$$n = \frac{4z_{\alpha/2}^2[p_1(1 - p_1) + p_2(1 - p_2)]}{e^2}, \quad (2.30)$$

where  $p_1$  and  $p_2$  are the success rates of the two treatments. This conservative estimate, however, may be much larger than that truly required.

Obviously the sample size will be maximized when  $p_1 = p_2 = 1/2$ . For example if  $e = 0.01$  and  $\alpha = 0.1$ , the maximum sample size will be  $n = 54121$ . Suppose that there are two previous studies; one study of the first treatment found 4 out of 15 ( $\hat{p}_1 = 0.266$ , 90% CI= (0.078, 0.454)) patients cured while another study for the second treatment found 10 out of 65 ( $\hat{p}_2 = 0.154$ , 90% CI= (0.080, 0.228)) patients cured. Using  $\hat{p}_1$  and  $\hat{p}_2$  in (2.30) gives  $n = 35236$ . Putting the lower limits and the upper limits of the confidence intervals for  $p_1$  and  $p_2$  into formula (2.30) gives  $n = 15751$  and  $n = 45884$ , respectively. What value should we actually choose? We might, for example, use the average of these numbers.

The reason for this arbitrariness is that formulae such as (2.30) require point estimates of  $p_1$  and  $p_2$  while a better summary of the available information is a distribution over a range of values. Furthermore, the lengths of the intervals calculated at the end of a trial depend on the data collected, which is of course unknown at the planning stage.

In the next section some of the Bayesian methods for the sample size determination problem will be reviewed.

### 3. Bayesian methods of sample size determination; a review

As pointed out by Pham–Gia and Turkkan (1992), in contrast to the classical approach, in the Bayesian context sample size questions are less frequently discussed. There are some reasons for this. First, it is sometimes argued that a Bayesian statistician is interested more in a sequential procedure, using the posterior distribution at one step as the prior for the following step, until a conclusion is reached. So it seems that there is no need for a fixed sample size. In real life it is not always possible to use a sequential approach, and therefore it is important to find a suitable sample size prior to sampling. Secondly, precision conditions are not as widely used as in classical statistics, mainly because of computational problems. For example, the computation of exact credible intervals (or posterior confidence intervals) using the highest posterior density (HPD) region may require the use of special tables such as those constructed by Isaacs *et al* (1974) or special software. The associated question of the sample size required so that these intervals are shorter than a given quantity then becomes a complicated iterative process.

A number of researchers have considered inferential Bayesian approaches to the sample size determination problem. For the mean of a normal distribution, Adcock (1988) developed closed formulae for both the known and unknown variance cases by averaging the coverage of fixed length posterior credible sets over the predictive distribution of the data. Joseph and Bélisle (1997) used similar techniques to derive closed form formulae for the case of average lengths of fixed coverage posterior credible sets, as well as *worst case* criteria.

Bayesian sample size determination for estimating the success probability in binomial sampling has received considerable attention (see, for example, Adcock (1987), (1992), (1995), Pham–Gia and Turkkan (1992), Joseph *et al* (1995)). The case of the difference between two binomial parameters has been considered by Joseph *et al* (1997). In the following sections the outlines of some of the Bayesian methods for the sample size determination problem are reviewed.

### 3.1. Bayesian and Mixed Bayesian/Likelihood Criteria

Let  $\theta \in \Theta$  denote the parameter of interest and  $\pi(\theta)$  the prior density of  $\theta$ . Suppose that the experimental data are  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ , where  $n$  is the sample size and the components of  $\mathbf{X}$  are exchangeable and belong to the data space  $\mathcal{X}$ . The predictive density of  $\mathbf{X}$ , which we may also describe as the pre-posterior marginal density of the data, is

$$f(\mathbf{x}) = \int_{\Theta} \ell(\mathbf{x} | \theta) \pi(\theta) d\theta, \quad (2.31)$$

and the posterior density of  $\theta$  given  $\mathbf{x}$  is

$$\pi^n(\theta | \mathbf{x}) = \frac{\ell(\mathbf{x} | \theta) \pi(\theta)}{\int_{\Theta} \ell(\mathbf{x} | \theta) \pi(\theta) d\theta}, \quad (2.32)$$

where  $\ell(\mathbf{x} | \theta)$  is the likelihood of the data. If the prior density  $\pi(\theta)$  is a (possibly improper) uniform density over  $\Theta$ , then (2.32) is the normalized likelihood.

In the terminology of Joseph *et al* (1997), if  $\pi(\theta)$  represents the true prior information in both (2.31) and (2.32) then the methods of sample size determination, described in sections 3.1.1. to 3.1.3., are called Bayesian<sup>1</sup>, and if  $\pi(\theta)$  is the true prior information in (2.31) but a uniform density for  $\pi(\theta)$  is substituted in (2.32), then they are called mixed Bayesian/Likelihood criteria for sample size determination.

#### 3.1.1. Tolerance Intervals or Average Coverage Criterion (ACC)

Following the alpha-expectation tolerance region approach of Fraser and Guttman (1956), which Raiffa and Schlaifer (1961) called pre-posterior analysis, Adcock (1987) introduced the average coverage criterion. The essence of this approach is to hold the HPD length,  $l$ , fixed and allow the coverage probability  $1 - \alpha$  to vary with  $\mathbf{x}$ . In other words, one requires that, conditional on  $\mathbf{x}$ ,  $\theta$  is contained in a specified interval of length  $l$ , with probability  $1 - \alpha$ , that is

$$\int_{a(\mathbf{x},n)}^{a(\mathbf{x},n)+l} \pi^n(\theta | \mathbf{x}) d\theta = 1 - \alpha, \quad (2.33)$$

---

<sup>1</sup>Joseph *et al* (1997) call this the *fully Bayesian* approach to the sample size problem but, as already mentioned, we shall restrict this term to cases where a utility or a loss function is employed and the problem is treated as a decision problem.

where  $a(\mathbf{x}, n)$  is the lower limit of the HPD interval of length  $l$  for the posterior density  $\pi^n(\theta | \mathbf{x})$  and in general depends upon both  $\mathbf{x}$  and  $n$ .

This solution varies with  $\mathbf{x}$  and so before sampling one requires that (2.33) should hold on average over all possible samples. Thus we choose  $n$  so that

$$\int_{\mathcal{X}} \left\{ \int_{a(\mathbf{x}, n)}^{a(\mathbf{x}, n)+l} \pi^n(\theta | \mathbf{x}) d\theta \right\} f(\mathbf{x}) d\mathbf{x} = 1 - \alpha. \quad (2.34)$$

### 3.1.2. Average Length Criterion (ALC)

The essence of the ALC is that the coverage probability  $1 - \alpha$  is fixed and the HPD interval length varies depending on the data. In this method, which was first introduced by Joseph *et al* (1995), for each  $\mathbf{x} \in \mathcal{X}$  one must first find the HPD length  $l'(\mathbf{x}, n)$  such that

$$\int_{a(\mathbf{x}, n)}^{a(\mathbf{x}, n)+l'(\mathbf{x}, n)} \pi^n(\theta | \mathbf{x}) d\theta = 1 - \alpha, \quad (2.35)$$

then the sample size is the minimum integer  $n$  that satisfies

$$\int_{\mathcal{X}} l'(\mathbf{x}, n) f(\mathbf{x}) d\mathbf{x} \leq l, \quad (2.36)$$

where  $l$  is the prespecified average length. The left hand side of (2.36) averages the length of fixed coverage HPD intervals, weighted by the predictive density  $f(\mathbf{x})$ .

### 3.1.3. Worst Outcome Criterion (WOC)

The ACC and ALC are based on averages over all possible samples, but since inferences are conditional on the observed sample, they lead to coverages and lengths which are undesirably small or large, respectively, for some samples. A criterion that recognizes this, was described by Pham-Gia and Turkkan (1992) and developed by Joseph *et al* (1995). A conservative approach is to ensure a maximum length  $l$  and a minimum coverage probability  $1 - \alpha$ , regardless of the data  $\mathbf{X}$ . Thus the sample size is the minimum value of  $n$



such that

$$\inf_{\mathbf{x} \in \mathcal{X}} \int_{a(\mathbf{x},n)}^{a(\mathbf{x},n)+l} \pi^n(\theta | \mathbf{x}) d\theta \geq 1 - \alpha. \quad (2.37)$$

This rule is known as the worst outcome criterion. (Lindley (1997) opines that it is not a Bayes rule because it eschews the expected value operation over the marginal distribution of  $\mathbf{x}$ .)

### 3.2. Bayes Factor or Weiss's Method

The standard Bayesian mechanism for testing two hypotheses is to compute the posterior probability of either hypothesis given the data, the model, the study design, and the prior probabilities. However, sometimes prior probabilities for the two hypotheses are not readily available. In this case inference may be based on the Bayes factor  $B(0 : 1 | \mathbf{x})$  in favour of  $H_0$  against  $H_1$ :

$$B(0 : 1 | \mathbf{x}) = \frac{f(\mathbf{x} | H_0)}{f(\mathbf{x} | H_1)},$$

where, using Weiss's (1997) notation,

$$f(\mathbf{x} | H_j) = \int f(\mathbf{x} | \theta_j) \pi(\theta_j | H_j) d\theta_j,$$

in which  $\theta_j$  are the parameters under hypothesis  $H_j$ ,  $j = 0, 1$ ,  $f(\mathbf{x} | \theta_j)$  is the sampling distribution of  $\mathbf{X}$  under  $H_j$  and  $\pi(\theta_j | H_j)$  is the prior under the hypothesis  $H_j$ . Since  $B(0 : 1 | \mathbf{x})$  is non-negative, it is convenient to take logarithms and to define

$$b_{01} = b(0 : 1 | \mathbf{x}) = \log B(0 : 1 | \mathbf{x}).$$

A positive value of  $b_{01}$  means that the data support  $H_0$ , and a negative value means that the data support  $H_1$ . Kass and Raftery (1995) suggest interpreting  $b_{01}$  between 3 and 5 as strong evidence in favour of  $H_0$  and  $b_{01} = 5$  as very strong evidence. Negative values of  $b_{01}$  favour  $H_1$  in a symmetrical way.

Bayesian sample size selection might aim for a reasonable probability that  $b_{01}$  is greater than  $k$  or less than  $-k$  for a prescribed non-negative  $k$ , under  $H_0$  and  $H_1$ , respectively. The type *I* error may be defined as  $Pr(b_{01} < k | H_0)$  and the type *II* error may be  $Pr(b_{01} \geq -k | H_1)$ . The sample size may be

chosen to meet constraints on these two errors.

Weiss (1997) suggests another approach and gives a detailed presentation of Bayes factor use for sample size selection for a normal distribution with known variance. The basis of his procedure is to find the sample size by specifying  $\alpha$  and  $\beta$  and solving the pair of equations

$$\begin{cases} Pr\{b_{01} > b_{cut} \mid H_0\} = 1 - \alpha, \\ Pr\{b_{01} \leq b_{cut} \mid H_1\} = 1 - \beta, \end{cases} \quad (2.38)$$

for given  $H_0$  and  $H_1$ , to determine both the sample size  $n$  and the cut-off point  $b_{cut}$ .

### 3.3. Methods of moments

The methods addressed above are all well defined in principle but may be difficult to implement in practice. An approach that may avoid the implementation difficulties is to use the moments of the posterior distribution of  $\theta$  given  $\mathbf{x}$  and then average over the predictive distribution. Since the posterior variance provides information on the dispersion of the posterior distribution, some authors address the sample size problem by setting an upper bound on the posterior variance (Pham-Gia and Turkkan (1992)). For example De Groot (1985, p.323) has noticed that in binomial sampling, if the prior density is uniform (a special case of beta distribution), then it is sufficient to take  $n = 22$  samples to ensure that the variance of the posterior distribution is less than or equal to 0.01.

If  $V_{post}$  denotes the variance of the posterior distribution then a general criterion for sample size determination using the posterior variance is to require that

$$E_{\mathbf{x}}\{V_{post}\} \leq \epsilon, \quad (2.39)$$

where  $\epsilon (> 0)$  is a prespecified (known) constant.

#### 3.3.1. Expected Value of Sample Information

In some applications of Bayesian decision theory and quality management, the value of information (in money or an equivalent) may be considered.

When quantified, population or sample information can be related to the mean or variance of the related distribution.

**Definition 1:** If  $\Theta \neq \emptyset$  is the parameter space and  $\mathcal{A}$  is a non-empty set of actions (or decisions) available to the statistician then a real valued function  $L(\theta, a)$  defined as  $L : \Theta \times \mathcal{A} \rightarrow [0, \infty)$  is called the loss due to taking action  $a \in \mathcal{A}$  when parameter has the value  $\theta$ .

In Bayesian estimation theory if  $a$  is the estimate of  $\theta$  then the loss function  $L(\theta, a)$  reflects the discrepancy between the value of  $\theta$  and its estimate. Normally  $L(\theta, a) \geq 0$  and  $L(\theta, \theta) = 0$ .

**Definition 2:** The expected value of perfect information (*EVPI*) is the expected loss of the action that is optimal under the current state of information, prior to sampling.

As pointed out by Claxton and Posnett (1996), there is a probability that a decision based on the prior information will be wrong and opportunity losses will be incurred. The expected loss is the expected cost of the uncertainty surrounding the decision problem: this is the expected value of perfect information.

**Definition 3:** The expected value of sample information (*EVSI*( $n$ )), as a function of sample size  $n$ , is the amount that may be paid for the information provided by the sample without being worse off in terms of expected loss than without taking the sample.

The expected benefit of sample information is measured by the reduction in expected loss.

**Definition 4:** The expected net gain of sampling (*ENG**S*( $n$ )) is the difference between *EVSI*( $n$ ) and the sampling cost  $cn$ . Thus

$$ENG S(n) = EV SI(n) - cn. \quad (2.40)$$

Maximizing *ENG**S*( $n$ ), one can find the optimal sample size.

**Remark:** Under a squared error loss function  $L(\theta, a) = k(\theta - a)^2$ , where  $k$  is a positive constant

$$EV PI = kV_{prior},$$

and

$$EV SI(n) = k(V_{prior} - E(V_{post})),$$

in which  $V_{prior}$  denotes the variance of the prior distribution.

Following Morris (1968), Pham-Gia and Turkkan (1992) showed that if in a binomial sampling the parameter under investigation,  $p$ , has a beta prior distribution with parameters  $\alpha$  and  $\beta$ , then the optimal sample size is

$$n^* = \max\{0, [\frac{k(\alpha + \beta)V_{prior}}{c}]^{1/2} - (\alpha + \beta)\}.$$

As long as  $ENG S(n)$  is positive, gain outweighs expenses, but the net gain is maximized at  $n^*$ .

**Remark:** The average loss (with respect to the posterior distribution of  $\theta$  conditional on  $\mathbf{x}$ ) is called posterior expected loss. If  $a$  is an unbiased estimator for  $\theta$  and  $L(\theta, a) = k(\theta - a)^2$ , then the posterior expected loss is equal to  $kV_{post}$ . Its average with respect to the predictive distribution of  $\mathbf{X}$  is called integrated risk and denoted by  $r(\theta, a)$ . To find the optimal sample size for the point estimation of  $\theta$ , sometimes the inequality  $r(\theta, a) \leq \epsilon$  is used, which is another representation of (2.39).

### 3.4. Fully Bayesian Approach (Decision Theory)

The first fully Bayesian treatment of the sample size problem was set out in detail by Raiffa and Schlaifer (1961). Their method specifies a utility function  $u(n, \mathbf{x}, a, \theta)$  expressing the benefit of choosing  $n$ , obtaining the result  $\mathbf{x}$  and taking the action  $a$ , when the parameter has the value  $\theta$ .

Assuming a (prior) density for  $\theta$ ,  $\pi(\theta)$ , their resolution of the problem (as recently described by Lindley (1997)) is to proceed in reverse time order,  $\theta$ ,  $\alpha$ ,  $\mathbf{x}$ ,  $n$ , taking expectations of utility over the random quantities  $\theta$  and  $\mathbf{x}$ , and maximizing over the deterministic quantities  $a$  and  $n$ . The expected utility may be written as

$$\max_n \left[ \int_{\mathbf{x}} \max_a \left\{ \int_{\theta} u(n, x, a, \theta) \pi^n(\theta | a, \mathbf{x}) d\theta \right\} f(\mathbf{x} | n) d\mathbf{x} \right]. \quad (2.41)$$

They assume that the utility does not depend on  $\mathbf{x}$  and is additive and linear in  $n$ , so that, for some  $c$ ,

$$u(n, \mathbf{x}, a, \theta) = u(a, \theta) - cn.$$

So (2.41) may be written as

$$\max_n \left[ \int_{\mathbf{x}} \max_a \left\{ \int_{\theta} u(a, \theta) \pi^n(\theta | \mathbf{x}) d\theta \right\} f(\mathbf{x} | n) d\mathbf{x} - cn \right]. \quad (2.42)$$

Note that the choice of  $a$  does not influence the opinion of  $\theta$ , given  $\mathbf{x}$  and  $n$ , so in (2.42) we have written  $\pi^n(\theta | \mathbf{x})$  rather than  $\pi^n(\theta | a, \mathbf{x})$ .

This method of sample size determination is called maximization of expected utility (MEU).

As pointed out by Lindley (1997), on theoretical grounds the use of MEU in the determination of sample size has everything in its favour. First, it is the only method that is guaranteed to be coherent: a user of MEU can never be exposed as making absurd comparisons (technically, comparisons that violate the axioms). Second, it is accompanied by a well-defined algorithm, set out in expression (2.42), for its solution. This consists of a series of expectations and maximizations that can be performed on a computer. Third it is very flexible, since a wide range of utility functions can be accommodated.

Note that the methods based on the width of the posterior credible interval are rather different from those based on hypothesis testing — which is the basic idea in the classic or the frequentist approach. All methods are rather arbitrary in their choice of criteria unless they are fully Bayesian.

### 3.5. An Example

In this section we apply the Bayesian rules for sample size determination, noted above, to a normal distribution, and show how they work for inference about the mean for both the known and unknown variance cases.

#### 3.5.1. Sample Size Determination for Normal Mean with Known Variance

Suppose that conditional on  $\theta$  the observations are independently and identically distributed as  $N(\theta, \sigma^2)$  where  $\sigma^2$  is known and assume  $\theta$  has a prior

distribution which is  $N(\mu, \sigma^2/n_0)$ ,  $n_0 > 0$ . The posterior distribution of  $\theta$  given the sample mean  $\bar{x}_n$  and the predictive (or the unconditional) distribution of  $\bar{X}_n$  are, respectively,

$$\theta \mid \bar{x}_n \sim N\left(\frac{n\bar{x}_n + n_0\mu}{n + n_0}, \frac{\sigma^2}{n + n_0}\right), \quad \bar{X}_n \sim N\left(\mu, \sigma^2\left[\frac{1}{n} + \frac{1}{n_0}\right]\right).$$

As Joseph and Bélisle (1997) point out, the ACC, ALC and WOC approaches all give the same result, for this special case. Let  $\mu'$  denote the posterior mean of  $\theta$  and specify the HPD interval by  $|\theta - \mu'| \leq e$ . The ACC for this HPD of length  $2e$ , may be written as

$$\int_{\bar{x}_n} Pr[|\theta - \mu'| \leq e \mid \bar{x}_n] f(\bar{x}_n) d\bar{x}_n \geq 1 - \alpha.$$

Note that  $Pr[|\theta - \mu'| \leq e \mid \bar{x}_n]$  does not depend on  $\bar{x}_n$ , even though  $\mu'$  does. It follows that the sample size must be such that

$$\Phi[\sqrt{(n + n_0)e/\sigma}] \geq 1 - \alpha/2, \quad (2.43)$$

where  $\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$ . Solving for  $n + n_0$  in (2.43) leads to

$$n + n_0 \geq \sigma^2 z_{\alpha/2}^2 / e^2. \quad (2.44)$$

This demonstrates clearly the effect of the prior distribution of  $\theta$ . If we assume that the prior information is so uninformative that  $n_0$  is effectively 0, then we obtain the classical result (2.2). When  $n_0 \geq \sigma^2 z_{\alpha/2}^2 / e^2$  then strictly speaking it is unnecessary to take any sample.

**Remark:** The ACC may be extended to give a rule which is analogous to the frequentist hypothesis testing rule (2.5). If the two hypotheses are

$$H_0 : \mu' = \mu'_0$$

$$H_1 : \mu' = \mu'_0 + e,$$

then the following equations

$$\int_{\bar{x}_n} \left\{ Pr[|\theta - \mu'| \leq e \mid \bar{x}_n \text{ and } H_0] \right\} f(\bar{x}_n) d\bar{x}_n = 1 - \alpha$$

$$\int_{\bar{x}_n} \left\{ Pr[|\theta - \mu'| \geq e \mid \bar{x}_n \text{ and } H_1] \right\} f(\bar{x}_n) d\bar{x}_n = 1 - \beta$$

give

$$n + n_0 = \sigma^2(z_{\alpha/2} + z_{\beta})^2/e^2. \quad (2.45)$$

Following Weiss (1997), let us consider a novel Bayesian version of the hypothesis test

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \sim N(\theta_0 + e, \sigma^2/n_0).$$

As already mentioned, the basis of Weiss's method is to compute  $b_{01}$ , the logarithm of the Bayes factor. The sample size  $n$  and cut-off point are then found by solving the pair of equations in (2.38). For finite  $n_0$  the solution requires iterative computations. If  $n_0 \rightarrow \infty$ , for which the alternative hypothesis is called precise, then the sample size is

$$n \geq \sigma^2(z_{\alpha} + z_{\beta})^2/e^2 \quad (2.46)$$

which is very similar to (2.5). If the alternative is vague in the sense that  $n_0 \rightarrow 0$ , then the power approaches 1 at all sample sizes. Care needs to be exercised if  $n_0$  is set judgementally.

Following Lindley (1997), let us consider a utility function of the form

$$u(n, \mathbf{x}, a, \theta) = \xi - cn, \quad (2.47)$$

where  $\xi = 1$  if, conditional on the sample,  $\theta \mid \bar{x}_n \in [a, a + e]$  and  $\xi = 0$  otherwise. The unit cost of sampling  $c$ , and the length of interval  $e$  are given. The action  $a$  which is the starting point of the interval, is found as part of the expected utility maximization

$$\max_n \left[ \int_{\mathbf{x}} \max_a \left\{ \int_{\theta} \xi \pi^n(\theta \mid \mathbf{x}) d\theta \right\} f(\mathbf{x} \mid n) d\mathbf{x} - cn \right]. \quad (2.48)$$

It is obvious that

$$\int \xi \pi^n(\theta \mid \mathbf{x}) d\theta = \int_a^{a+e} \pi^n(\theta \mid \mathbf{x}) d\theta = \Phi\left(\frac{a+e-\mu'}{\sqrt{\frac{\sigma^2}{n+n_0}}}\right) - \Phi\left(\frac{a-\mu'}{\sqrt{\frac{\sigma^2}{n+n_0}}}\right).$$

To find the maximum of the above expression we use  $\frac{d}{da} \int_a^{a+e} \pi^n(\theta \mid \mathbf{x}) d\theta = 0$ . It turns out that  $a = \mu' - e/2$  is the optimum act (decision) and

$$\max_a \left\{ \int_{\theta} \xi \pi^n(\theta \mid \mathbf{x}) d\theta \right\} = 2\Phi\left(\frac{e\sqrt{n+n_0}}{2\sigma}\right) - 1. \quad (2.49)$$

Putting (2.49) in (2.48) and noting that (2.47) does not depend on  $\mathbf{x}$ , the optimal sample size is the solution to

$$\max_n [2 \Phi(\frac{e/2\sqrt{n+n_0}}{\sigma}) - 1 - cn]. \quad (2.50)$$

To find the maximum of the above expression we use differentiation with respect to  $n$ , and set it equal to zero

$$2 [\Phi'(\frac{e/2\sqrt{n+n_0}}{\sigma})] \frac{e}{4\sigma} (n+n_0)^{-1/2} - c = 0.$$

Hence, the optimal sample size is the solution to

$$n + n_0 = \left[ \frac{e}{2\sigma c} \phi\left\{ \frac{e}{2\sigma} \sqrt{n+n_0} \right\} \right]^2, \quad (2.51)$$

where  $\phi(\cdot)$  denotes the standard normal density function. For sample size determination in this case of a normal distribution with known variance, the key point is that the ACC, ALC, and WOC lead to the same result. Lindley's method is very different mathematically but as pointed out by Adcock (1997) it may be made equivalent to the ACC by a suitable choice of the unit cost  $c$ . It is shown in the following table.

**Table 1: Example of sample sizes for normal distribution with known variance**

<i>Rule</i>	<i>Sample Size</i>
Frequentist (formula (2.5))	71
ACC, ALC, WOC (formula (2.44))	33
Fully Bayesian (formula (2.51))	
$c = 0.001$	126
$c = 0.003$	68
$c = \mathbf{0.007}$	<b>33</b>
$c = 0.01$	22

$\alpha = 0.05$ ,  $\beta = 0.2$ ,  $n_0 = 2$ ,  $e = 1$ , and  $\sigma = 3$ .



### 3.5.2. Sample Size Determination for Normal Mean with Unknown Variance

When the variance is known, the results in the previous section show that there is a close correspondence between sample size determination from the frequentist, and the Bayesian, standpoints. When the variance is unknown the differences are greater. In the frequentist approach, it may be necessary to sample twice, or otherwise to have access to an independent estimate of  $\sigma^2$ . In the Bayesian approach the key idea is to consider a prior density for  $\sigma^2$ . Joseph and Bélisle (1997) give a comprehensive account of sample size selection for inference about the mean, and for inference about the difference between the means of two normal distributions. Following Adcock (1988), let us assume that  $\sigma^2$  is unknown but the quantity  $w\sigma_0^2/\sigma^2$  is distributed as  $\mathcal{X}_{(w)}^2$ ,  $w > 0$ ; this being the standard form of prior distribution for variance. As above, the tolerance interval is  $|\theta - \mu'| \leq e$ , which is also the HPD interval. The ACC sample size satisfies

$$n + n_0 \geq \sigma_0^2 t_{\alpha/2}(w)/e^2. \quad (2.52)$$

This is very similar to Stein's (1945) rule with  $n_0 - 1 = w$  and  $\sigma_0^2$  replaced by  $S^2$ . Joseph and Bélisle (1997) describe analogous ALC and WOC rules. For the utility defined by equation (2.47) we obtain

$$n + n_0 = \left[ e/2 f_w \left\{ \frac{e}{2\sigma_0} \sqrt{n + n_0} \right\} / \sigma_0 c \right]^2, \quad (2.53)$$

where  $f_w(\cdot)$  is the density of the Student  $t$ -distribution with  $w$  degrees of freedom, and  $c$  is the unit cost of sampling. The methods of Bayes factors may also be used but is more complicated to implement because of the form of the marginal distributions  $f(\bar{x}_n | H_0)$  and  $f(\bar{x}_n | H_1)$ . □

Other contributions related to Bayesian sample size determination include Goldstein (1981), Berger (1985), Gould (1993), Hutton and Owens (1993), Stallard (1998), Pezeshk and Gittins (1999) and Gittins and Pezeshk (2000).

# Chapter 3

## A Simplified Introductory Model

### 1. Introduction and An Executive Summary

In this chapter we discuss optimal sample size for the case where the number of subsequent users of the new treatment is a linear function of the posterior expectation of the difference  $\delta$  between the performance of the two treatments. We return to the fully Bayesian analysis of Raiffa and Schlaifer (1961), with the modification that instead of a utility-maximising terminal decision a plausible model is assumed for the way patients and their medical advisers respond to the evidence from a trial. Our assumptions are somewhat oversimplified, the aims being to show that the approach is computationally tractable and gives sensible answers.

We suppose that the cost of a trial is a known function of its size, and the benefit per patient of using the new treatment is a known function of  $\delta$ . This is assumed to be expressed in monetary terms so that the cost of trial may be directly offset against it. The optimal sample size is obtained by maximizing expected net benefit. This is the total benefit from the resulting change in the number of patients using the new treatment minus the cost of trial. Not surprisingly it turns out that the optimal sample size increases when the prior uncertainty about  $\delta$  increases, and decreases when the cost/benefit ratio of using the new treatment increases.

We give a Bayesian solution to a simplified model for sample size determination in a clinical trial, where the data are from a normal distribution with unknown mean  $\theta$  (this might be the performance parameter of the new

treatment), and known variance  $\sigma^2$ . A normal prior distribution is assumed to describe our prior knowledge about  $\theta$ . The objective function is based on the number of subsequent users of the new treatment, the expected benefit of using the new treatment, and the cost of conducting the trial.

In section 2 we state our sample size problem and introduce the notation. As a check on the consistency of our calculations, we solve the problem for the trivial case for which the number of subsequent users does not depend on the outcome of the trial.

In section 3 we provide an exact solution for the problem in the case where the number of subsequent users is proportional to the number of standard deviations by which the posterior mean of  $\theta$  exceeds the performance of the current treatment. An explicit expression is obtained for the expected net benefit of the trial, and for the maximizing sample size. The limitations of the model, such as the problem caused by assuming an unbounded linear increasing function representing the number of subsequent users of the new treatment, are also discussed in this section. It should be noted that, because of the above assumption, the model introduced in this chapter is unrealistic but it is discussed because it leads to an analytical solution.

## 2. Sample Size Problem

Suppose  $X_1, X_2, \dots$  are *i.i.d.* observations from  $N(\theta, \sigma^2)$  where  $\sigma^2$  is known from previous experience, or else can be estimated from the results themselves with sufficient precision so that errors of estimation of  $\sigma^2$  can be ignored, and  $\theta$  is unknown. We can think of  $\theta$  as the performance parameter of a new treatment and  $\theta_0$  (known) as the performance parameter of the treatment in current use. We assume that for every patient treated by the new treatment, there is a benefit  $b(\theta - \theta_0)$ , where  $b (\geq 0)$  is known. Let  $C(n) = cn$  be the cost of doing an experiment on a sample of size  $n$ . To formulate our prior knowledge about  $\theta$  let us suppose a prior density function for  $\delta = \theta - \theta_0$  which is  $N(\mu, \tau^2)$ . If  $\mu'$  is the expectation of the posterior density of  $\delta$  after doing the experiment, then the function,  $m(\mu')$  is the number of subsequent users of the new treatment. The question is what sample size maximizes the expected net benefit?

The objective function is the total benefit from the resulting change in the number of patients using the treatment minus the cost of the trial. Putting

$Z_i = X_i - \theta_0$ ,  $i = 1, 2, \dots$ , the objective (or the expected net benefit) function can be written as

$$r(n) = \int_{\bar{z}_n} \int_{\delta} m(\mu'(\bar{z}_n))(b\delta) d\Pi^n(\delta|\bar{z}_n) f(\bar{z}_n) d\bar{z}_n - cn \quad (3.1)$$

where  $\Pi^n(\delta|\bar{z}_n)$  is the posterior distribution of  $\delta$ , given the sufficient statistic  $\bar{z}_n$ . We are looking for  $n^*$  which maximizes  $r(n)$ .

It is well known that if  $z_i|\delta \sim N(\delta, \sigma^2)$  and  $\delta \sim N(\mu, \tau^2)$  then  $\delta|z_i \sim N(\delta_1, \psi_1^2)$  where  $\psi_1^2 = \frac{1}{\tau^{-2} + \sigma^{-2}}$  and  $\delta_1 = \psi_1^2(\frac{\mu}{\tau^2} + \frac{z_i}{\sigma^2})$ . So if  $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)$  is the vector of sample observations then  $\delta|\bar{z}_n \sim N(\mu'(\bar{z}_n), \tau'^2)$  where

$$\mu'(\bar{z}_n) = \frac{\sigma^2\mu + n\tau^2\bar{z}_n}{\sigma^2 + n\tau^2}, \quad \tau'^2 = \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}.$$

By calculating the inner integral in (3.1) we arrive at the following representation of  $r(n)$ :

$$r(n) = \int_{\bar{z}_n} m(\mu'(\bar{z}_n))(b\mu'(\bar{z}_n))f(\bar{z}_n)d\bar{z}_n - cn. \quad (3.2)$$

It is straightforward to show that  $\bar{Z}_n \sim N(\mu, \frac{n\tau^2 + \sigma^2}{n})$ . It thus follows from (3.2) that

$$r(n) = \int_{\bar{z}_n} m(\mu'(\bar{z}_n))b\mu'(\bar{z}_n) \phi\left(\frac{\bar{z}_n - \mu}{\sqrt{(n\tau^2 + \sigma^2)/n}}\right) \sqrt{\frac{n}{n\tau^2 + \sigma^2}} d\bar{z}_n - cn. \quad (3.3)$$

To check the consistency of the above result with a trivial case for which the solution is known, suppose the number of subsequent users of the new treatment does not depend on the outcome of the trial, so we may write  $m(\mu'(\bar{z}_n)) = m'$ . Equation (3.3) now becomes:

$$r(n) = \int_{\bar{z}_n} m'b\mu'(\bar{z}_n)\phi\left(\frac{\bar{z}_n - \mu}{\sqrt{(n\tau^2 + \sigma^2)/n}}\right) \sqrt{\frac{n}{n\tau^2 + \sigma^2}} d\bar{z}_n - cn,$$

or

$$r(n) = m'bE(\mu'(\bar{Z}_n)) - cn,$$

or

$$r(n) = m'b\mu - cn.$$

So the expected net benefit is a decreasing function of  $n$  and the optimal value of  $n$  is zero, as we might have expected. In general  $m(\mu')$  is not constant. In section 3 we consider it as a linear increasing function of  $\mu'$ .

### 3. A Bayesian Solution

As noted in section 3 of the previous chapter, in the fully Bayesian framework the optimal sample size may be found by maximizing the expected total benefit of conducting the trial minus its cost.

#### 3.1. Maximizing the Expected Net Benefit

Throughout we assume that the number of subsequent users of the new treatment depends upon the posterior distribution of  $\delta$ . For example suppose that

$$m(\mu'(\bar{z}_n)) = k_1\left(\frac{\mu'(\bar{z}_n)}{\tau'}\right) + k_2, \quad (3.4)$$

where  $k_1(> 0)$  and  $k_2(\geq 0)$  are known.

A high value of  $k_1$  corresponds to a large number of potential users. The constant  $k_2$  has no influence on the size of the trial, as it represents a fixed benefit, independent of the outcome. We shall see that the optimum sample size depends only upon  $\frac{c}{k_1 b}$  (a kind of cost/benefit ratio),  $\frac{\tau}{\sigma}$  (a measure of prior uncertainty), and  $\frac{\mu}{\sigma}$  (a location parameter), which we denote by  $C$ ,  $T$ , and  $D$ .

**Theorem:** If the number of subsequent users of the new treatment is given by (3.4) then the optimal sample size for the problem of section 2 with the scaled variable  $\frac{b}{\sigma}$  in place of  $b$ , is

$$n^* = \left[ \left\{ C + \sqrt{\frac{(D^2 + T^2)^3}{27} + C^2} \right\}^{1/3} + \left\{ C - \sqrt{\frac{(D^2 + T^2)^3}{27} + C^2} \right\}^{1/3} \right]^{-2} - T^{-2}. \quad (3.5)$$

**Proof:** The objective function may be written as

$$r(n) = \int_{\bar{z}_n} [k_1\left(\frac{\mu'(\bar{z}_n)}{\tau'}\right) + k_2] b\left(\frac{\mu'(\bar{z}_n)}{\sigma}\right) f(\bar{z}_n) d\bar{z}_n - cn, \quad (3.6)$$

or

$$\begin{aligned} r(n) &= \int_{\bar{z}_n} \frac{k_1 b}{\sigma} \left( \frac{\frac{\sigma^2 \mu + n \tau^2 \bar{z}_n}{\sigma^2 + n \tau^2}}{\sqrt{\frac{\sigma^2 \tau^2}{\sigma^2 + n \tau^2}}} \right) \left( \frac{\sigma^2 \mu + n \tau^2 \bar{z}_n}{\sigma^2 + n \tau^2} \right) f(\bar{z}_n) d\bar{z}_n \\ &\quad + \int_{\bar{z}_n} \frac{k_2 b}{\sigma} \left( \frac{\sigma^2 \mu + n \tau^2 \bar{z}_n}{\sigma^2 + n \tau^2} \right) f(\bar{z}_n) d\bar{z}_n - cn. \end{aligned}$$

It can be shown that the above expression becomes

$$r(n) = \frac{k_1 b}{\sigma \sqrt{\frac{\sigma^2 \tau^2}{\sigma^2 + n \tau^2}}} \int_{\bar{z}_n} [\mu + \frac{\tau^2}{\frac{\sigma^2}{n} + \tau^2} (\bar{z}_n - \mu)]^2 f(\bar{z}_n) d\bar{z}_n + \frac{k_2 b}{\sigma} \mu - cn.$$

After some calculations we arrive at the following expression:

$$r(n) = \frac{k_1 b}{\sigma \sqrt{\frac{\sigma^2 \tau^2}{\sigma^2 + n \tau^2}}} [\mu^2 + (\frac{\tau^2}{\frac{\sigma^2}{n} + \tau^2})^2 (\frac{\sigma^2}{n} + \tau^2)] + \frac{k_2 b}{\sigma} \mu - cn.$$

This can be written as

$$r(n) = k_1 b (T^{-2} + n)^{1/2} [D^2 + n T^2 (T^{-2} + n)^{-1}] + k_2 b D - cn. \quad (3.7)$$

To find the maximum of  $r(n)$  we suppose  $n \geq 0$  is a continuous variable and set  $\frac{dr(n)}{dn} = 0$ . Thus we have

$$\begin{aligned} \frac{dr(n)}{dn} = k_1 b [1/2 (T^{-2} + n)^{-1/2} (D^2 + n T^2 (T^{-2} + n)^{-1}) \\ + (T^{-2} + n)^{1/2} (T^2 (T^{-2} + n)^{-1} - n T^2 (T^{-2} + n)^{-2}) - C] = 0. \end{aligned}$$

Writing  $H = (T^{-2} + n)^{-1/2}$  it follows that

$$\frac{dr(n)}{dn} = 0 \iff H^3 + (D^2 + T^2)H - 2C = 0. \quad (3.8)$$

To solve this equation we use the standard method for solving the cubic equation  $x^3 + bx^2 + cx + d = 0$ . Define

$$\begin{aligned} f = \frac{b}{3} = 0, \quad g = \frac{c}{3} - f^2 = \frac{(D^2 + T^2)}{3}, \quad h = \frac{d}{2} + f^3 - \frac{fc}{2} = -C \\ j = g^3 + h^2 = \frac{(D^2 + T^2)^3}{27} + (-C)^2 \end{aligned}$$

Since  $j > 0$  the equation has one real root. The real root  $H^*$  of the equation can be found as follows

Let  $k = \sqrt{j}$ ,  $l_1 = (-h + k)^{1/3}$ , and  $l_2 = (-h - k)^{1/3}$ . Then the real root of the equation is equal to  $l_1 + l_2$ . Thus

$$H^* = \left\{ C + \sqrt{\frac{(D^2 + T^2)^3}{27} + (-C)^2} \right\}^{1/3} + \left\{ C - \sqrt{\frac{(D^2 + T^2)^3}{27} + (-C)^2} \right\}^{1/3}.$$

Since

$$\begin{aligned} \frac{d^2 r(n)}{dn^2} &= (3H^2 \frac{dH}{dn} + (D^2 + T^2) \frac{dH}{dn}) \left( \frac{k_1 b}{2} \right) \\ &= \left\{ 3H^2 \left( -1/2(T^{-2} + n)^{-3/2} \right) + (D^2 + T^2) \left( -1/2(T^{-2} + n)^{-3/2} \right) \right\} \left( \frac{k_1 b}{2} \right) \\ &< 0, \end{aligned}$$

this turning point gives the maximum value of  $r(n)$ . We now have  $n^* = H^{*-2} - T^{-2}$ . It should be noted here that the optimal sample size,  $n^*$ , is positive if  $\frac{dr(n)}{dn} \big|_{n=0} > 0$ . Putting  $n = 0$  in (3.8) it follows that  $C < T(D^2/2 + T^2)$ . So if the cost of sampling is greater than  $T(D^2/2 + T^2)$  it is not worth doing any trial and  $n^* = 0$ .

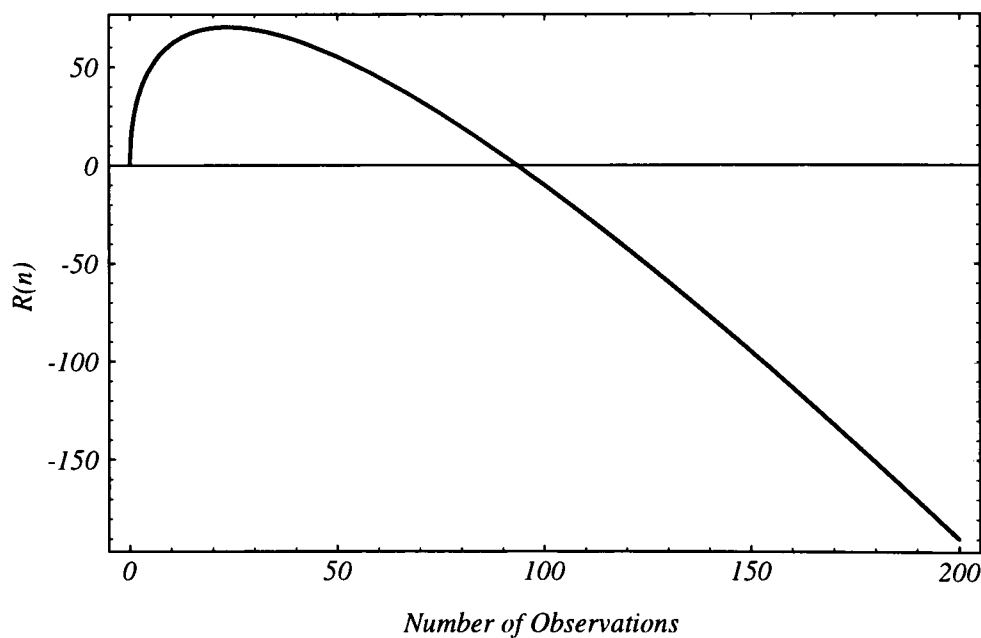


Figure 3.1: Expected net benefit when  $D = 2$ ,  $T = 5$ , and  $C = 3$ .

Figures 3.1 and 3.2 illustrate the variation of the scaled objective function  $R(n) = \frac{r(n)}{k_1 b} - \frac{k_2}{k_1} D$ . The optimal size of trial  $n^*$  for the case shown by figure 3.1 is 23 and for that shown by figure 3.2 is 211.

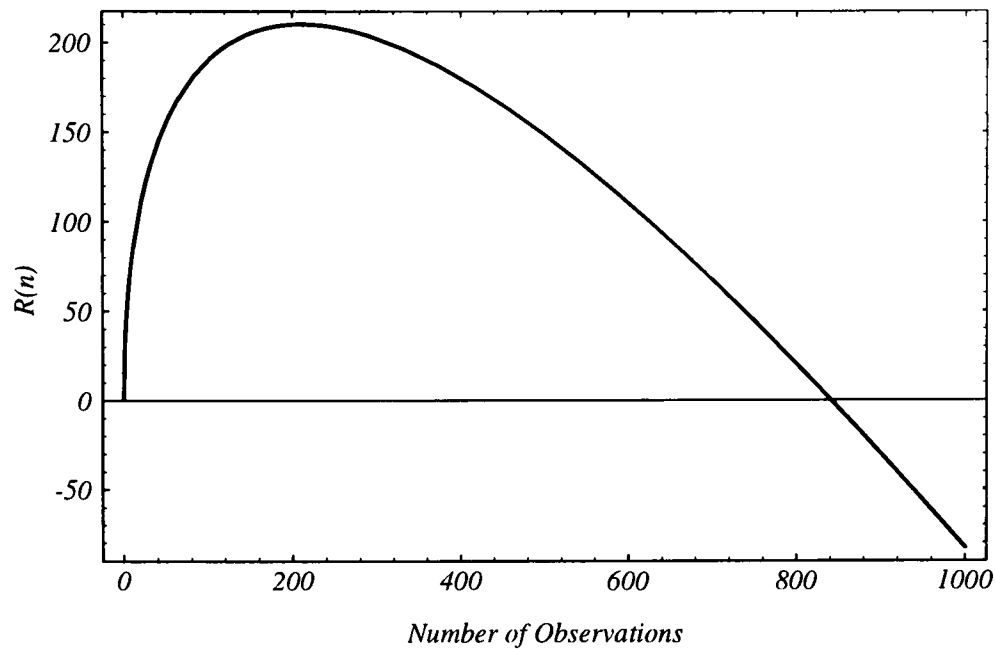


Figure 3.2: Expected net benefit when  $D = 2$ ,  $T = 5$ , and  $C = 1$ .

Figure 3.3 illustrates the variation of the logarithm of the optimal sample size  $n^*$  as a function of  $T$  and  $C$ . The optimal sample size increases when prior uncertainty of  $\delta$  increases and decreases when the cost/benefit ratio of using the new treatment increases. In this example we assume that the benefit of using the new treatment is  $b\mu'$  and  $D$  the measure of our prior expectation of the superiority of the new treatment is equal to two.



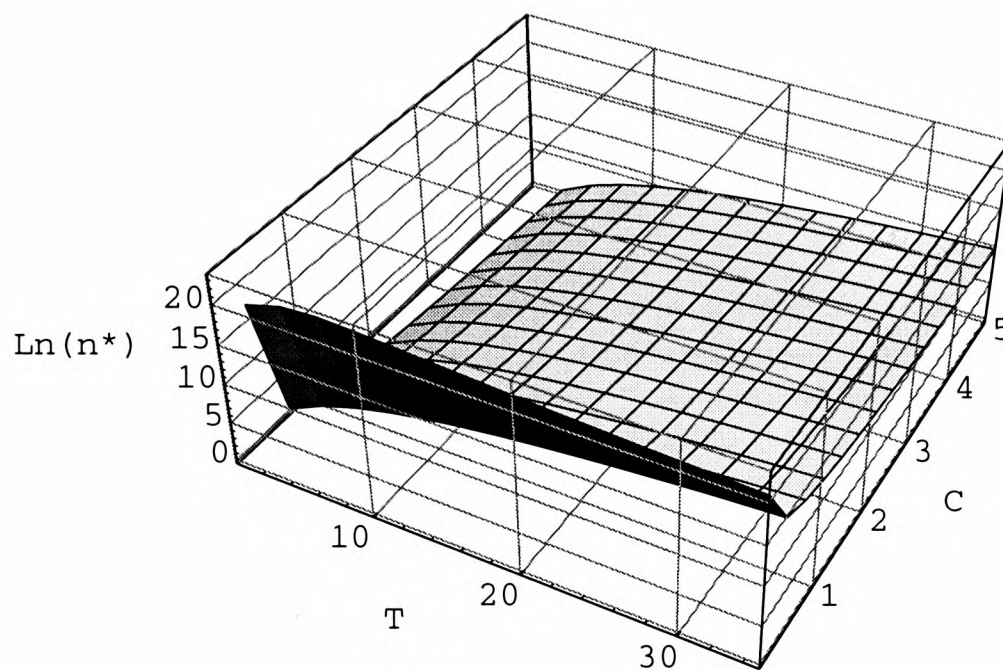


Figure 3.3: Logarithm of the optimal sample size  $n^*$ , as a function of cost/benefit ratio  $C$ , and the parameter of uncertainty  $T$ .

### 3.1.1. A Note on the Benefit Structure

If the benefit of using the new treatment is written from the standpoint of the pharmaceutical company selling the new treatment, it may be appropriate to assume that the total benefit would be equal to the number of patients using the new treatment times a fixed benefit  $b$  per each patient. For this case we may have the following lemma.

**Lemma:** If the benefit of using the new treatment is  $b$  and if the number of subsequent users of the new treatment is given by (3.4) then the optimal sample size  $n^*$  for the problem of section 2 is

$$n^* = \left(\frac{D}{2C}\right)^2 - T^{-2}$$

**Proof:** The objective function may be written as

$$r(n) = \int_{\bar{z}_n} [k_1(\frac{\mu'(\bar{z}_n)}{\tau'}) + k_2] b f(\bar{z}_n) d\bar{z}_n - cn, \quad (3.9)$$

or

$$r(n) = \int_{\bar{z}_n} k_1 b \left( \frac{\frac{\sigma^2 \mu + n \tau^2 \bar{z}_n}{\sigma^2 + n \tau^2}}{\sqrt{\frac{\sigma^2 \tau^2}{\sigma^2 + n \tau^2}}} \right) f(\bar{z}_n) d\bar{z}_n + \int_{\bar{z}_n} k_2 b f(\bar{z}_n) d\bar{z}_n - cn.$$

After some calculations it turns out that

$$\frac{r(n)}{k_1 b} - \frac{k_2}{k_1 b} \left( \frac{\mu}{\sigma} \right) = \left\{ \left( \frac{\tau}{\sigma} \right)^{-2} + n \right\}^{-1/2} \frac{\mu}{\sigma} - \frac{c}{k_1 b} n.$$

Putting, once again,  $R(n) = \frac{r(n)}{k_1 b} - \frac{k_2}{k_1 b} \left( \frac{\mu}{\sigma} \right)$ , and using the same notation as in (3.5), this gives

$$R(n) = (T^{-2} + n)^{1/2} D - Cn. \quad (3.10)$$

It is straightforward to show that the value of  $n$  which maximizes the above objective function is given by

$$n^* = \left( \frac{D}{2C} \right)^2 - T^{-2} \quad (3.11)$$

It should be noted that if  $C > \frac{1}{2}DT$ , it is not worth doing any trial and  $n^* = 0$ .

Figure 3.4 illustrates the variation of the scaled objective function for the case where the benefit of using the new treatment is written from the standpoint of the pharmaceutical company conducting the trial. The optimal sample size for this example is  $n^* = 99$ .

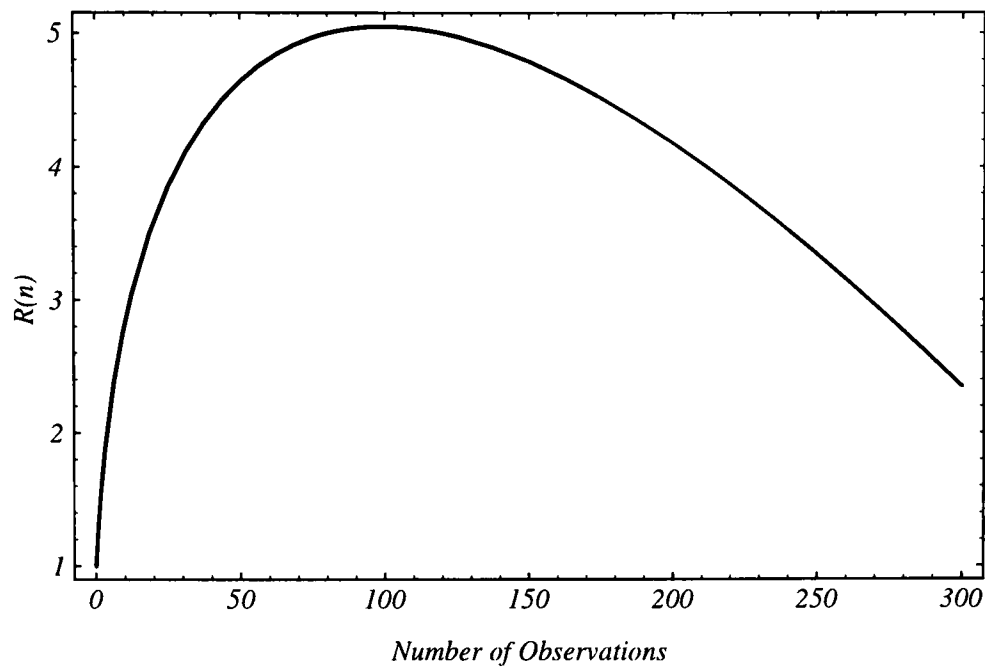


Figure 3.4: Expected net benefit when  $D = 1$ ,  $T = 1$ , and  $C = 0.05$ .

### 3.2. The Parameters Determining Optimal Sample Size

From the equations (3.5) and (3.11) it is clear that the optimal value  $n^*$  of  $n$  depends only on the parameters  $C$ ,  $T$ ,  $D$ . These, we recall, have the following interpretations;

$C$  is a cost/benefit ratio,

$T$  is a measure of our prior ignorance of  $\mu$ ,

$D$  is a measure of our prior expectation of the superiority of the new treatment.

Calculations show that  $n^*$  is a decreasing function of  $C$  and an increasing function of both  $T$  and  $D$ .

Figures 3.5 and 3.6 refer to the solution (3.5). The optimal sample size,  $n^*$ , is a decreasing function of  $C$  (figure 3.5) and an increasing function of  $T$  (figure 3.6).

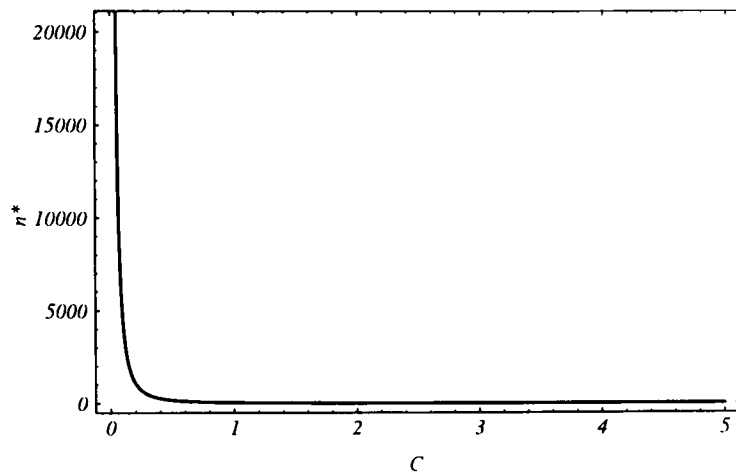


Figure 3.5: Optimal sample size,  $n^*$ , as a function of cost/benefit ratio,  $C$ : clearly  $n^*$  is a decreasing function of  $C$ ;  $T = 3$ .

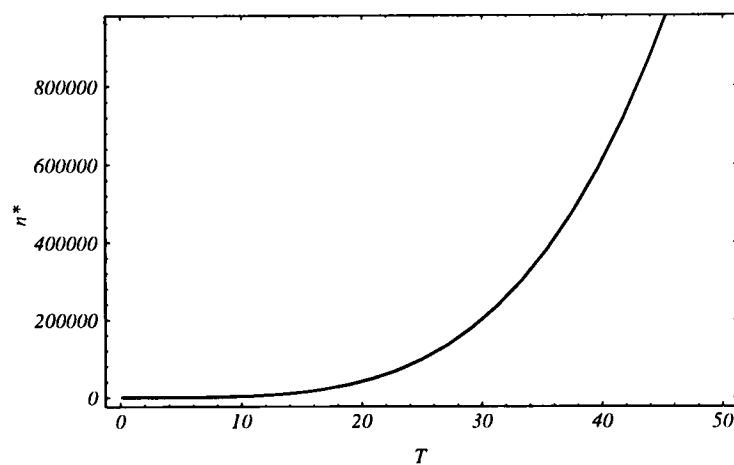


Figure 3.6: Optimal sample size,  $n^*$ , as a function of the parameter of uncertainty  $T$ ;  $C = 1$ .

That  $n^*$  should decrease as the cost/benefit ratio increases (figure 3.5), and that it should increase as the prior ignorance increases (figure 3.6), is just what common sense suggests. That it should also be large when we expect the new treatment to be much better than the current one at first looks odd. In that case why do we need a trial at all? The reason for this is the factor  $\frac{\mu'(\bar{z}_n)}{\tau'}$  in the objective function (3.6). For instance, if it turns out that  $\frac{\mu}{\tau} = 6$  and we choose the sample size big enough to ensure that  $\tau' = \frac{1}{2}\tau$ , then  $E(\frac{\mu'(\bar{z}_n)}{\tau'})$  becomes approximately 12, with a consequent increase in the

objective function,  $r(n)$ . On the other hand a smaller value of  $\frac{\mu}{\tau}$  leads to a smaller increase in  $r(n)$ , and consequently a smaller value of  $n^*$ .

## Chapter 4

# A Behavioural Bayes Approach to the Sample Size Question; Normal Distribution, Known Variance

In this chapter the determination of the optimal sample size of a clinical trial is considered when the number of subsequent users of the new treatment is a function both of the magnitude of the apparent difference between the performance of the new treatment and that of the treatment in current use and of the statistical significance of the difference. A general model is introduced with different features which allow for various contingencies.

### 1. Introduction

The determination of the optimal sample size of a clinical trial is considered when the number of subsequent users of a new treatment is a function of both the magnitude of the apparent difference between the performance of the new treatment and that of the treatment in current use and of the statistical significance of the difference. An objective function is proposed consisting of the total benefit from the resulting change in the number of patients using the new treatment minus the cost of the trial. From this the optimal sample size may be calculated. The model has features which allow for the following contingencies: a cost difference between the two treatments; a payoff function defined either from the public health or from a drug com-

pany standpoint; the performance of the existing treatment may be either known or unknown; varying degrees of severity of the condition to be treated; and a set-up cost for conducting the trial.

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, depend on the strength of the evidence provided by the trial.

A detailed description of the function representing the number of subsequent users of the new treatment is given in section 2.

In section 3 we state the sample size problem and introduce a general model which incorporates the above mentioned features. Both public health and drug company oriented benefit functions are used.

Section 4 defines the function representing the number of subsequent users of the new treatment. Various cases which may be handled within the general model framework are discussed in section 5.

Sections 6 and 7 present a detailed discussion for the problem when there is a regulator whose approval of a licence for the new treatment is necessary.

In section 8 the general model is applied to two recent trials which were carried out by a pharmaceutical company. These are trials 1 and 2 of those 6 trials described in chapter 1.

We assume throughout this chapter that the observed effect is normally distributed with a known variance. Although the known variance assumption is often a realistic assumption but it is relaxed in the next chapter.

## **2. The Number of Subsequent Users of the New Treatment**

The number of subsequent users of the new treatment will not be high, unless it is, in the statistical sense, significantly better than the existing one. There must also be a reasonable expectation that the new treatment achieves a sufficiently large improvement to justify a switch.

For the moment we ignore the question of statistical significance and just deal with the difference  $\delta$  between the performances, on some appropriate scale, of the new treatment and of the existing treatment. Let  $\Delta C = \text{expected cost of new treatment} - \text{expected cost of old treatment}$ . First suppose  $\Delta C = 0$ . It is reasonable to suppose that  $m$ , the number of subsequent users, depends on  $\delta$  as shown below.

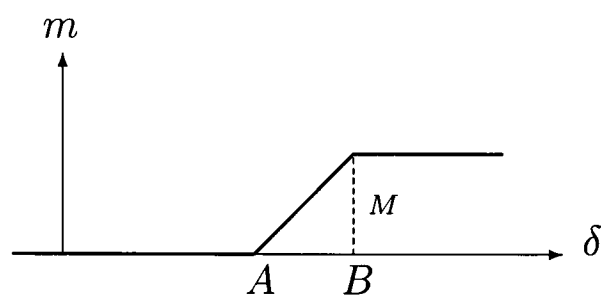


Figure 4.1: Number of subsequent users of the new treatment when  $\Delta C = 0$ .

Here  $M$  is the expected total number of users when  $\Delta C = 0$ , given a substantial improvement in performance. The above function implies that if the amount of improvement is less than a certain value  $A$ , then nobody switches to the new treatment; if it is between  $A$  and another higher value  $B$  then some patients switch to the new treatment; and if it is higher than  $B$  then all the patients switch to the new treatment.

Now suppose  $\Delta C > 0$ . The function representing the number of subsequent users will be of the form shown below.

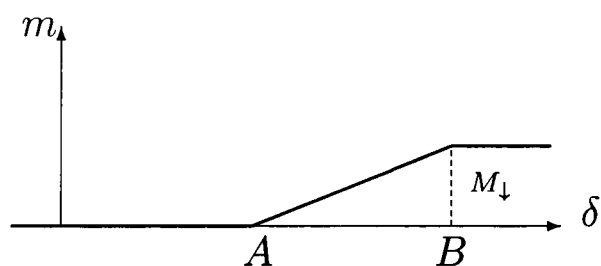


Figure 4.2: Number of subsequent users of the new treatment when  $\Delta C > 0$ .

For some people the cost makes no difference to usage, so  $A$  is unchanged. For some others the extra cost is prohibitive, so the maximum expected number of subsequent users  $M$  is decreased. For others the extra cost means that a



higher apparent improvement in performance is needed than when  $\Delta C = 0$ , so  $B$  is increased.

Finally, suppose  $\Delta C < 0$ . The number of subsequent users will be a function of the form shown below.

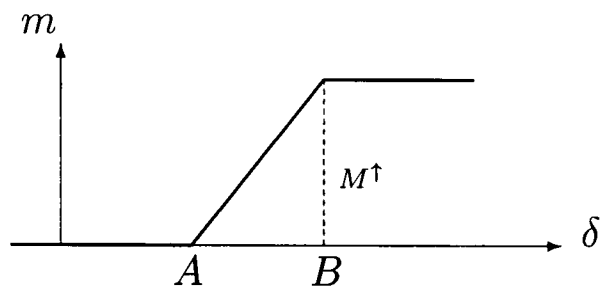


Figure 4.3: Number of subsequent users of the new treatment when  $\Delta C < 0$ .

Compared to the case where  $\Delta C = 0$ , now  $B$  is unchanged,  $M$  is increased and, because some additional patients can now afford to use the new treatment,  $A$  is decreased ( $A$  could be negative).

The above discussion assumes that the actual amount of improvement in performance is known. It may be extended to the case where there is uncertainty about the amount of improvement. It is sufficient to shift the above mentioned functions  $1\frac{1}{2}\tau'$  or  $2\tau'$  to the right, where  $\tau'$  is the standard deviation of the posterior distribution of the unknown parameter  $\delta$ . This corresponds to assuming that individuals are prepared to switch to the new treatment provided that the probability of the difference between the two treatments exceeding their personal threshold difference is at least some minimum value. Figure 4.4 sets out in graphical terms, where  $\mu'$  is the posterior mean of  $\delta$ .

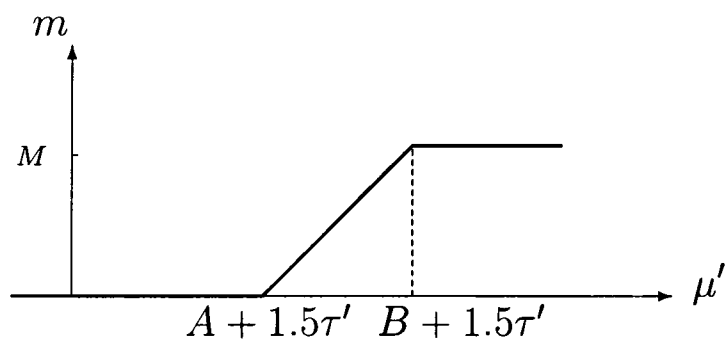


Figure 4.4: Number of subsequent users,

Thus

$$m = \begin{cases} 0 & \mu' < A + 1.5\tau' \\ \frac{M}{B-A}[\mu' - A - 1.5\tau'] & A + 1.5\tau' < \mu' < B + 1.5\tau' \\ M & B + 1.5\tau' < \mu'. \end{cases} \quad (4.1)$$

We may also note how  $A$  and  $B$  may be expected to depend on the severity of the condition being treated. For an increasingly severe condition  $A$  and  $B$  both decrease and approach zero, for  $\Delta C = 0$ . Also the dependence of  $m$  on  $\Delta C$  is less marked for a severe condition than for a more mild condition. We assume that some of the patients are prepared to switch to the new treatment if the actual amount of improvement  $\delta$  is greater than  $0.8\mu$  and if  $\delta > 1.2\mu$  then all of them switch to the new treatment.

### 3. A General Model

Our discussion starts by considering paired comparison of normally distributed data. This is simply for convenience. As we shall see in section 5 the resulting methodology applies without these restrictions.

Suppose that the  $X_i$ 's are the clinical outcomes on some appropriate scale for patients using the new treatment and the  $Y_i$ 's are those for patients using another (current) treatment. Assume that  $Z_i = X_i - Y_i$  ( $i = 1, 2, \dots$ ) has a normal density which is  $N(\delta, \sigma^2)$ , and let us formulate our previous knowledge about  $\delta$  by assuming a prior density for it which is  $N(\mu, \tau^2)$ . The random variable  $n^{-1} \sum_{i=1}^n Z_i$  will be written as  $\bar{Z}_n$ , and  $\bar{z}_n$  denotes the value which it takes. Under our assumptions of normality with known variance,  $\bar{z}_n$  is a sufficient statistic for the parameter  $\delta$ . Let  $m$  be the number of subsequent users of the new treatment.

For every patient who goes on to use the new treatment as a result of the trial there is a benefit. The objective (or expected net benefit) function  $r(n)$ , for a trial with  $n$  patients on each treatment, is the total expected benefit from the resulting change in the number of patients using the new treatment minus the cost of trial. The BeBay program is based on a benefit per patient of  $b_1\delta + b_2$  ( $b_1 \geq 0$ ). The discussion throughout this dissertation is based on the two extreme cases  $b_1 = 0$  and  $b_2 = 0$ . The expected benefit is in general the sum of these two components. The case  $b_2 = 0$  might correspond in aggregate to a public health benefit, while the second case  $b_1 = 0$  might be more appropriate for a drug company, the total benefit in this case

depending on the number of patients using the new treatment but not on its performance. It will be convenient to refer to the two extreme cases as the public health and commercial cases, oversimplified as this undoubtedly is.

Let the cost of carrying out a trial with  $n$  pairs of patients be  $cn + d$  if  $n > 0$ , and 0 if  $n = 0$ . Thus  $d$  is a set-up cost. The question is how large should the trial be? We can proceed to calculate the optimal trial size  $n^*$  as though  $d = 0$ , and then compare the values of  $r(n)$  and  $r(0)$  to see whether a trial should be done. To simplify our discussion we shall assume from now on that  $d = 0$ .

The model, which will be described in 3.1 and 3.2, is stated for normally distributed data but it is much more widely applicable as we shall see in section 5.

We shall now proceed to write our objective function for these two extreme cases of benefit function in turn, and give some numerical examples.

### 3.1. Public Health Benefit Function

The benefit per patient of using the new treatment is assumed to be  $b_1\delta$  which is measured in the cost units. For the sake of notation let us assume  $b_1 = \frac{b}{\sigma}$  ( $b \geq 0$ ). The objective (ie expected net benefit) function, which is proposed, consisting of the total benefit from the resulting change in the number of patients using the new treatment minus the cost, can be written as

$$r(n) = \int_{\bar{z}_n} \int_{\delta} m\left(\frac{b\delta}{\sigma}\right) d\Pi^n(\delta|\bar{z}_n) f(\bar{z}_n) d\bar{z}_n - cn. \quad (4.2)$$

Here  $\Pi^n(\delta|\bar{z}_n)$  is the posterior distribution of the unknown parameter  $\delta$  given  $\bar{z}_n$ . This is  $N(\mu'(\bar{z}_n), \tau'^2)$ , where

$$\mu'(\bar{z}_n) = \frac{\sigma^2\mu + n\tau^2\bar{z}_n}{\sigma^2 + n\tau^2}, \quad \tau'^2 = \frac{\sigma^2\tau^2}{\sigma^2 + n\tau^2}.$$

Thus (4.2) can be written as

$$r(n) = \int_{\bar{z}_n} m \frac{b\mu'(\bar{z}_n)}{\sigma} f(\bar{z}_n) d\bar{z}_n - cn, \quad (4.3)$$

where  $f(\bar{z}_n)$  is the predictive density of  $\bar{Z}_n$ . Since  $\delta \sim N(\mu, \tau^2)$ ,  $\bar{Z}_n - \delta \sim N(0, \sigma^2/n)$ , and  $\delta$  and  $\bar{Z}_n - \delta$  are independent, it follows that  $\bar{Z}_n \sim N(\mu, \tau^2 + \sigma^2/n)$ .

It is convenient to express the problem in terms of the scaled variables  $R(n) = \frac{r(n)}{Mb}$ ,  $C = \frac{c}{Mb}$ ,  $D = \frac{\mu}{\sigma}$ , and  $T = \frac{\tau}{\sigma}$ . Now we show that  $R(n)$  depends on  $c$ ,  $M$ ,  $b$ ,  $\mu$ ,  $\tau$ , and  $\sigma$  as a function of the three ratios  $C$ ,  $D$ , and  $T$ .

Let the number of subsequent users of the new treatment be of the form shown in figure 4.4. It is convenient to work with  $A\sigma$  and  $B\sigma$  in places of  $A$  and  $B$  in (4.1). This change of variable makes a considerable reduction in number of parameters appeared in the objective function and from now on through this chapter we work with them. Thus equation (4.3) may be written as

$$\begin{aligned} R(n) = 0 + \int_{H_1(A,n)}^{H_2(B,n)} \frac{1}{B-A} \left[ \frac{\mu'(\bar{z}_n)}{\sigma} - A - 1.5 \frac{\tau'}{\sigma} \right] \\ \times \left( \frac{\mu'(\bar{z}_n)}{\sigma} \right) f(\bar{z}_n) d\bar{z}_n \\ + \int_{H_2(B,n)}^{\infty} \left( \frac{\mu'(\bar{z}_n)}{\sigma} \right) f(\bar{z}_n) d\bar{z}_n - Cn, \end{aligned} \quad (4.4)$$

where  $H_1(A, n)$  and  $H_2(B, n)$  can be found from  $A + 1.5 \frac{\tau'}{\sigma} < \frac{\mu'(\bar{z}_n)}{\sigma} < B + 1.5 \frac{\tau'}{\sigma}$ ,

$$H_1(A, n) = \frac{\sigma[A + 1.5 \frac{\tau'}{\sigma}](\sigma^2 + n\tau^2) - \sigma^2\mu}{n\tau^2},$$

$$H_2(B, n) = \frac{\sigma[B + 1.5 \frac{\tau'}{\sigma}](\sigma^2 + n\tau^2) - \sigma^2\mu}{n\tau^2}.$$

Writing  $u = \frac{\bar{z}_n - \mu}{\sqrt{\frac{\sigma^2}{n} + \tau^2}}$  we find after some calculations that

$$\begin{aligned}
R(n) = & \frac{1}{B-A} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T^{-2} + n)^{-1/2} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \quad (4.5) \\
& \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{h_2(B,n)}^{\infty} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn,
\end{aligned}$$

where

$$\begin{aligned}
h_1(A, n) &= \frac{[A + 1.5(T^{-2} + n)^{-1/2} - D](T^{-2} + n)^{1/2}}{T\sqrt{n}}, \\
h_2(B, n) &= \frac{[B + 1.5(T^{-2} + n)^{-1/2} - D](T^{-2} + n)^{1/2}}{T\sqrt{n}}.
\end{aligned}$$

Numerical calculations show that in all cases  $R(n)$  has a unique maximum as a function of  $n$ . Figures 4.5 and 4.6 illustrate the variation of the scaled

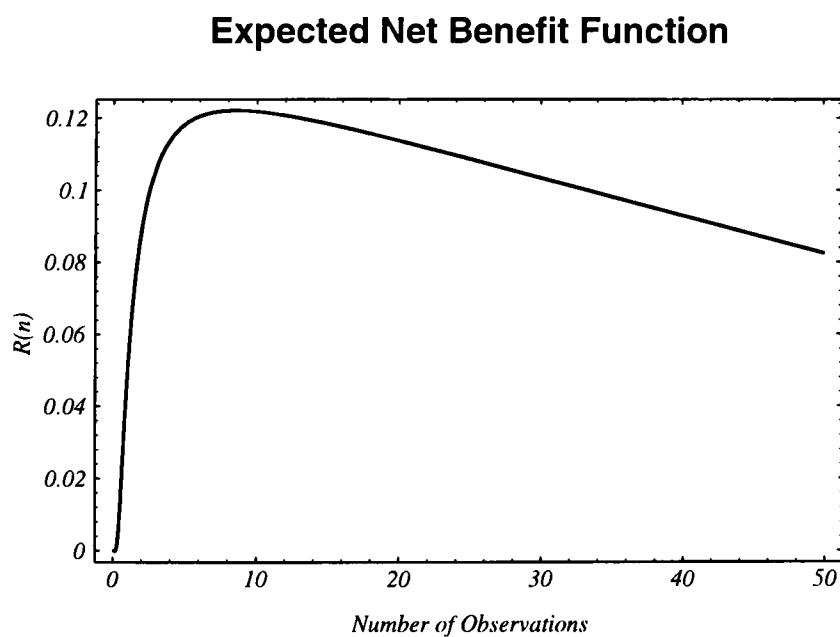


Figure 4.5: Expected benefit when  $D = 1$ ,  $T = 1$ ,  $C = 0.001$ , and  $(A, B) = (1, 1.5)$ ; public health benefit function.

objective function. The optimal size of trial  $n^*$  for the case shown by figure 4.5 is 9 and for that shown by figure 4.6 is 18.

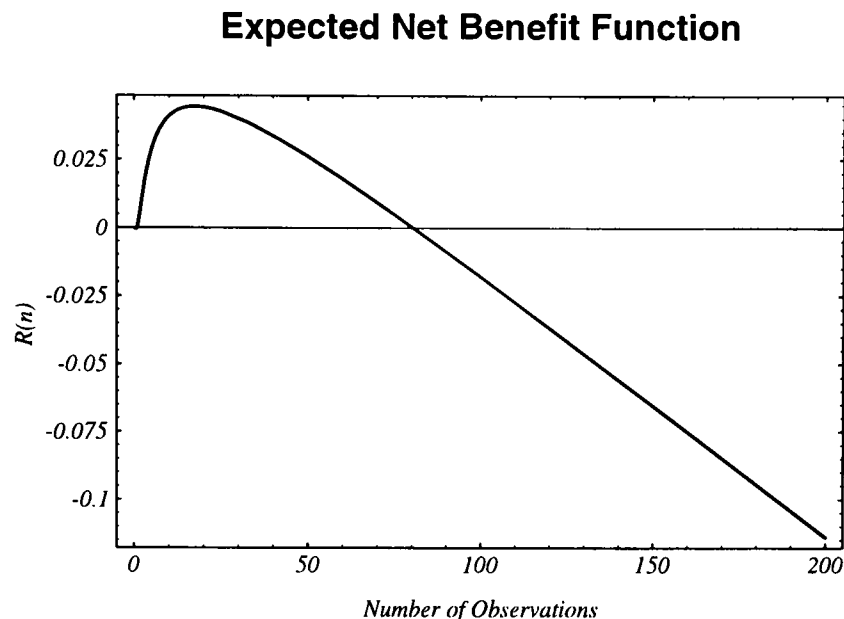


Figure 4.6: Expected net benefit when  $D = 1$ ,  $T = 1$ ,  $C = 0.001$ , and  $(A, B) = (2, 2.5)$ ; public health benefit function.

### 3.2. Commercial Benefit Function

For each patient using the new treatment we assume there is a benefit  $b_2 = b$  ( $b \geq 0$ ) independent of  $\delta$ . Thus

$$r(n) = \int_{\bar{z}_n} mbf(\bar{z}_n)d\bar{z}_n - cn. \quad (4.6)$$

As before scaled variables  $R(n)$ ,  $C$ ,  $D$ , and  $T$  may be defined. Again  $C = \frac{c}{Mb}$ , though the meaning of  $b$  is now different. We now show that  $R(n)$  depends on  $c$ ,  $M$ ,  $b$ ,  $\mu$ ,  $\tau$  and  $\sigma$  as a function of the three ratios  $C$ ,  $D$ , and  $T$ .

With  $m$  as in equation (4.1) we have

$$\begin{aligned} R(n) = 0 + & \int_{H_1(A,n)}^{H_2(B,n)} \frac{1}{B-A} \left[ \frac{\mu'(\bar{z}_n)}{\sigma} - A - 1.5 \frac{\tau'}{\sigma} \right] f(\bar{z}_n) d\bar{z}_n \\ & + \int_{H_2(B,n)}^{\infty} f(\bar{z}_n) d\bar{z}_n - Cn, \end{aligned}$$

where  $H_1(A, n)$  and  $H_2(B, n)$  are as in (4.4).

Proceeding as in (4.5) it may be shown that

$$\begin{aligned}
 R(n) = & \frac{1}{B-A} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & - \frac{A}{B-A} \int_{h_1(A,n)}^{h_2(B,n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & - \frac{1}{B-A} 1.5(T^{-2} + n)^{-1/2} \int_{h_1(A,n)}^{h_2(B,n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & + \int_{h_2(B,n)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn,
 \end{aligned} \tag{4.7}$$

where  $h_1(A, n)$  and  $h_2(B, n)$  are as in (4.5). Again,  $R(n)$  has a unique maximum as a function of  $n$ .

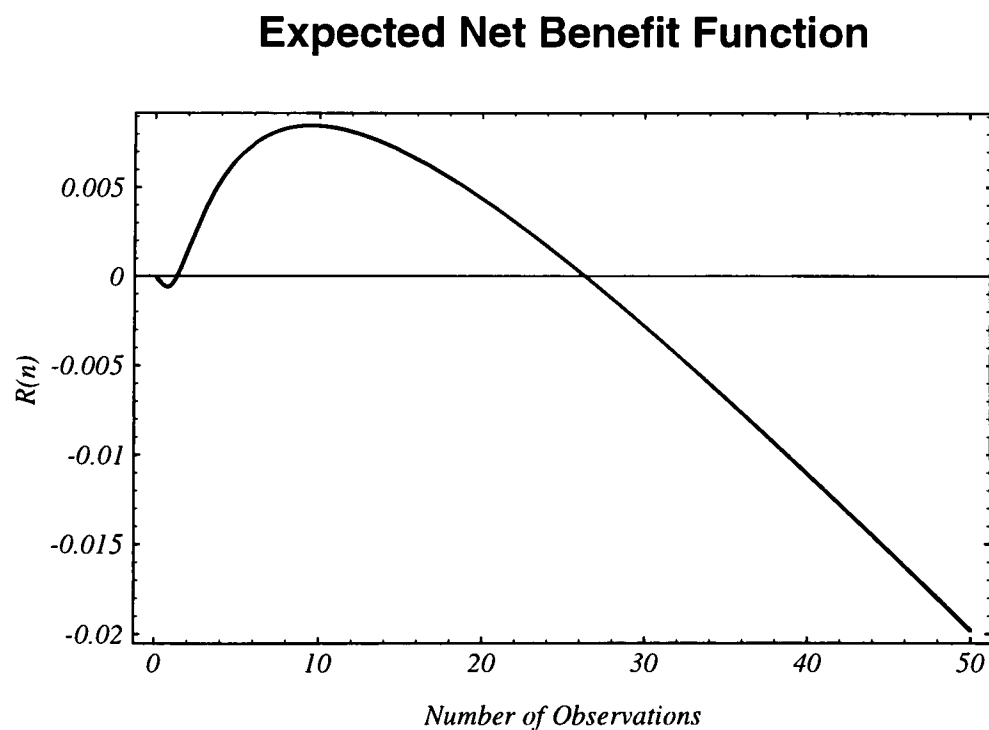


Figure 4.7: Expected net benefit when  $D = 1$ ,  $T = 1$ ,  $C = 0.001$ , and  $(A, B) = (2, 2.5)$ ; commercial benefit function.

Figure 4.7 illustrates the variation of the scaled objective function. The optimal size of trial  $n^*$  for the case shown by this figure is 10.

## 4. Distribution of the Number of Subsequent Users of the New Treatment

Before stating a lemma to obtain the distribution function of  $m$  consider the case where prior knowledge about  $\delta$ , the difference between the performance of the new treatment and that of the treatment in current use (or control treatment), is such that  $\mu < A$ . Thus if we decide not to conduct a trial the expected benefit, and hence the expected net benefit, is zero:

If the number of observations  $n$  is zero, then  $cn = 0$  and the posterior mean and variance for  $\delta$  are in fact the prior mean and prior variance. The function representing the number of subsequent users of the new treatment now is as in figure 4.8.

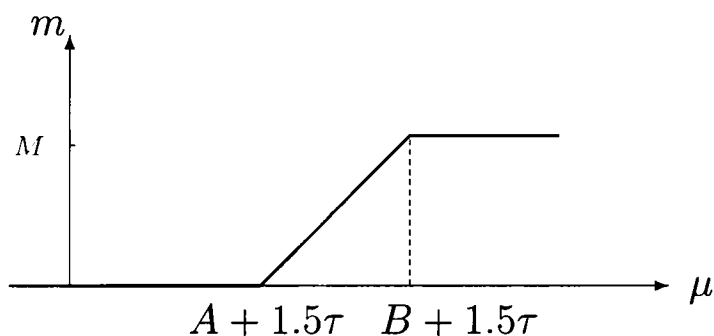


Figure 4.8: Number of subsequent users when  $n = 0$ .

Since  $\mu < A$  implies  $\mu < A + 1.5\tau$ , the number of subsequent users of the new treatment  $m$ , and hence the expected benefit from the trial, is zero. Thus the objective function, which is the benefit resulting from subsequent use of the new treatment minus the cost of the trial, is zero.

It should be noted that if  $\mu < A$  but the uncertainty parameter  $\tau$  is large it may well be worth conducting a trial.

In this section and in sections 6 and 7 our discussion is in terms of the scaled parameter defined in section 3.

**Lemma:** The prior distribution function of the number  $m$  of subsequent



users of the new treatment is

$$F_m(y) = \begin{cases} 0 & y < 0 \\ \Phi\left(\left[\frac{y}{M}(B - A) + A + 1.5(T^{-2} + n)^{-1/2} - D\right] \times \frac{(T^{-2} + n)^{1/2}}{T\sqrt{n}}\right) & 0 \leq y < M \\ 1 & M \leq y. \end{cases} \quad (4.8)$$

**Proof:** If  $0 \leq y < M$  then as shown by figure 4.9 we can write

$$\begin{aligned} F_m(y) &= P_r(m \leq y) \\ &= P_r\left\{\frac{\mu'(\bar{z}_n)}{\sigma} \leq \frac{y}{M}(B - A) + A + 1.5\tau'/\sigma\right\}. \end{aligned}$$

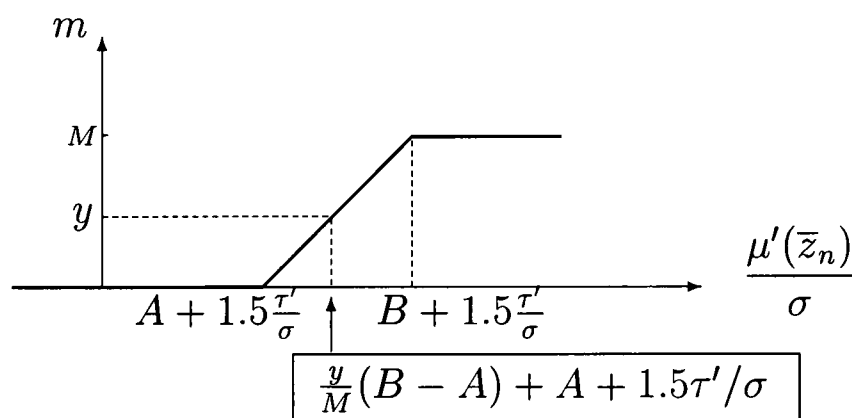


Figure 4.9: Number of subsequent users.

This may be written as

$$\begin{aligned} P_r\left\{\frac{\bar{Z}_n - \mu}{\sqrt{\frac{\sigma^2 + n\tau^2}{n}}}\right\} \\ \leq \frac{\left[\frac{y}{M}(B - A) + A + 1.5(T^{-2} + n)^{-1/2} - D\right](T^{-2} + n)^{1/2}}{T\sqrt{n}} \Bigg\}, \end{aligned}$$

or

$$F_m(y) = \Phi\left(\left[\frac{y}{M}(B - A) + A + 1.5(T^{-2} + n)^{-1/2} - D\right] \frac{(T^{-2} + n)^{1/2}}{T\sqrt{n}}\right),$$

$$0 \leq y < M.$$

Note that

$$\begin{aligned} P_r(m = 0) &= P_r\left\{\frac{\mu'(\bar{z}_n)}{\sigma} < A + 1.5\tau'/\sigma\right\} \\ &= P_r\left\{\frac{\bar{Z}_n - \mu}{\sqrt{\frac{\sigma^2 + n\tau^2}{n}}} < [A + 1.5(T^{-2} + n)^{-1/2} - D]\frac{(T^{-2} + n)^{1/2}}{T\sqrt{n}}\right\} \\ &= \Phi(h_1(A, n)), \end{aligned}$$

and similarly

$$P_r(m = M) = P_r\left\{\frac{\bar{Z}_n - \mu}{\sqrt{\frac{\sigma^2 + n\tau^2}{n}}} \geq h_2(B, n)\right\} = 1 - \Phi(h_2(B, n)).$$

□

**Remark:** The prior density function of  $m$  is of the form

$$f_m(y) = \begin{cases} P_r(m = y) = \Phi(h_1(A, n)) & y = 0 \\ \phi\left([ \frac{y}{M}(B - A) + A + 1.5(T^{-2} + n)^{-1/2} - D ] \right. \\ \quad \left. \times \frac{(T^{-2} + n)^{1/2}}{T\sqrt{n}} \right) \frac{B - A}{MT\sqrt{n}} (T^{-2} + n)^{1/2} & 0 < y < M \\ P_r(m = y) = 1 - \Phi(h_2(B, n)) & y = M. \end{cases} \quad (4.9)$$

Using the above density one can find the expected number  $E(m)$  of subsequent users of the new treatment.

Figure 4.10 illustrates the variation of  $E(m)$  as a function of  $n$ . For this example it is assumed that  $(A, B) = (0.836, 1.254)$ ,  $\mu = \frac{A+B}{2} = 1.045$ ,  $T = 0.5225$ , and the maximum number  $M$  of potential subsequent users of the new treatment is 10000.

### Expected Number of Subsequent Users (M=10000)

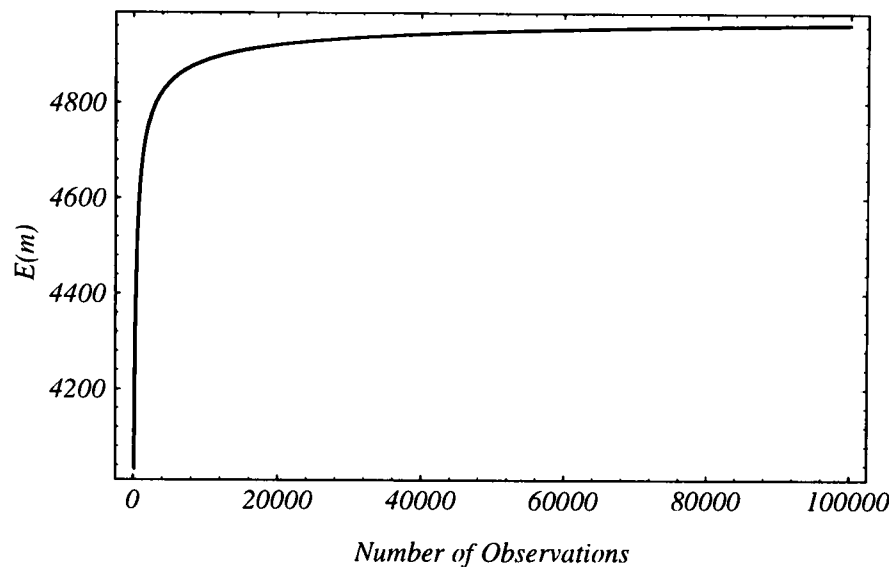


Figure 4.10: Expected number of subsequent users as a function of  $n$ .

As shown by the graph it is a bounded and increasing function with the upper bound of  $M/2 = 5000$ . For this example we have

$$\lim_{n \rightarrow \infty} E(m) = \frac{M}{2}.$$

## 5. Range of Application

Our general model was formulated on the assumption of normally distributed effects with a known variance. Here we show that, by using the central limit theorem (CLT) to justify the assumption of normality, the same analysis may be applied in a wide variety of other situations. In each case the prior and posterior density for  $\delta$ , and the predictive density for  $\bar{Z}_n$ , take the same forms as in section 3, as do the calculations leading to the optimal size of the trial.

For most trials there are control groups rather than paired comparisons. A control group may be an untreated group or a group receiving an active treatment or a placebo. The need to increase the number of patients who have received a new treatment prior to marketing approval, combined with the need to study these high numbers in a limited time, and the need to have

a reliable estimate of performance of the new treatment in various situations, for instance in different hospitals, means that many trials, in particular in phase III, are carried out at more than one centre. To avoid or control the bias in the results of a multi-centre trial there are usually control groups at each centre. We shall see how the general model deals with sample size determination for such trials in different situations.

It should be noted here that if the control group is decided not to be the same size as the actively treated group we can still apply our methodology for determining the optimal sizes of each of them by considering the different prior information and using our approach twice; once on the treatment group and once on the control group.

Our model can be used for a clinical trial comparing two treatments for a disease with a binary outcome, loosely termed *success* and *failure*.

Sometimes a patient response takes the form of a count: the number of times an event occurs during a specified period of time. For instance the number of angina attacks during the month following treatment. Usually the events are undesirable and the experimental treatment is intended to lower the event rate. Our model can deal with this situation assuming that the count has a Poisson distribution.

Note that in the following cases  $X_i$ 's and  $Y_i$ 's are, respectively, the clinical outcomes for patients using the new treatment and for the existing (control) treatment. It follows from the central limit theorem that strict normality is not required for the first three cases.

#### *Normal data, paired comparisons*

For this case, which was discussed in section 3, we can assume that  $X_i|\xi_i \sim N(\xi_i, \phi^2)$ ,  $Y_i|\xi_i \sim N(\xi_i - \delta, \phi^2)$ . Thus  $Z_i = X_i - Y_i \sim N(\delta, 2\phi^2)$ , and  $\bar{Z}_n \sim N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 2\phi^2$ .

#### *Normal data, treated group and (same size) control group, but unpaired*

For this case let  $X_i|\xi_i \sim N(\xi_i, \phi^2)$ ,  $\xi_i \sim N(\theta + \delta, \rho^2)$ ,  $\bar{X}_n = n^{-1} \sum X_i$ ,  $Y_i|\nu_i \sim N(\nu_i, \phi^2)$ ,  $\nu_i \sim N(\theta, \rho^2)$ , and  $\bar{Y}_n = n^{-1} \sum Y_i$ . Thus  $\bar{X}_n \sim N(\theta + \delta, (\phi^2 + \rho^2)/n)$ ,  $\bar{Y}_n \sim N(\theta, (\phi^2 + \rho^2)/n)$ , and  $\bar{Z}_n = \bar{X}_n - \bar{Y}_n \sim N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 2(\phi^2 + \rho^2)$ .

#### *Normal data, known mean performance $\theta_0$ for existing treatment*

Let  $X_1, X_2, \dots$  be the clinical outcomes for patients using the new treatment

in the trial and  $Z_i = X_i - \theta_0$ . If  $X_i \sim N(\theta, \sigma^2)$  then  $Z_i \sim N(\delta, \sigma^2)$ , and  $\bar{Z}_n \sim N(\delta, \sigma^2/n)$ .

*Binary data, paired comparisons*

Let  $p(X_i = 1|p_i) = p_i$ ,  $p(Y_i = 1|p_i) = p_i - \delta$ ,  $p_i$  be close to some central value  $p$  with high probability, and  $\delta \ll p$  with high probability. Thus, to a good approximation,  $Z_i = X_i - Y_i$  has mean  $\delta$  and variance  $2p(1 - p)$ , and for moderately large  $n$ , by the CLT,  $\bar{Z}_n \sim N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 2p(1 - p)$ .

*Binary data, treated group and (same size) control group, but unpaired*

Let  $p(X_i = 1|p_i) = p_i$ ,  $p_i \sim N(\theta + \delta, \rho^2)$ ,  $\bar{X}_n = n^{-1} \sum X_i$ ,  $p(Y_i = 1|q_i) = q_i$ ,  $q_i \sim N(\theta, \rho^2)$ ,  $\bar{Y}_n = n^{-1} \sum Y_i$ ,  $\theta$  be close to some central value  $p$  with high probability,  $\delta$  and  $\rho$  both be  $\ll p$ .

Using the CLT it follows that, to a good approximation,  $\bar{X}_n \sim N(\theta + \delta, (p(1 - p) + \rho^2)/n)$ ,  $\bar{Y}_n \sim N(\theta, (p(1 - p) + \rho^2)/n)$  and  $\bar{Z}_n = \bar{X}_n - \bar{Y}_n \sim N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 2p(1 - p) + 2\rho^2$ .

*Poisson data, paired comparisons*

Let  $X_i|\lambda_i \sim Poi(\lambda_i)$ ,  $Y_i|\lambda_i \sim Poi(\lambda_i + \delta)$ ,  $\lambda_i$  be close to some central value  $\lambda$  with high probability and  $\delta \ll \lambda$  with high probability. Thus, to a good approximation,  $Z_i = Y_i - X_i$  has mean  $\delta$  and variance  $2\lambda$ , and for large  $n$ , by the CLT,  $\bar{Z}_n = \bar{Y}_n - \bar{X}_n \sim N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 2\lambda$ .

*Poisson data, treated group and (same size) control group, but unpaired*

Let  $X_i|\lambda_{i1} \sim Poi(\lambda_{i1})$ ,  $\lambda_{i1} \sim N(\theta, \rho^2)$ ,  $\bar{X}_n = n^{-1} \sum X_i$ ,  $Y_i|\lambda_{i2} \sim Poi(\lambda_{i2})$ ,  $\lambda_{i2} \sim N(\theta + \delta, \rho^2)$ ,  $\bar{Y}_n = n^{-1} \sum Y_i$ ,  $\theta$  be close to some specified value  $\lambda$  with high probability, and  $\delta$  and  $\rho$  both be  $\ll \lambda$ . Using the CLT it follows that, to a good approximation  $\bar{X}_n \sim N(\theta, (\lambda + \rho^2)/n)$  and  $\bar{Y}_n \sim N(\theta + \delta, (\lambda + \rho^2)/n)$ . Thus  $\bar{Z}_n = \bar{Y}_n - \bar{X}_n \sim N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 2\lambda + 2\rho^2$ .

*Survival analysis, proportional hazards, treated group and (same size) control group, unpaired*

Binary responses defined for a clinical trial usually concern the occurrence of an event within a specified time. A survival analysis seeks greater power by using the time from allocation to treatment until some event as the patient response. This time is referred to as the patient's *survival time*. It is assumed that the event under consideration is undesirable, such as death.

Let  $h_X(t)$  and  $h_Y(t)$ , be the hazard rates for a treated patient, and for a control-group patient, at time  $t$  after admission to the trial. Let  $\log\{\frac{h_Y(t)}{h_X(t)}\} =$

$\delta$ , which is assumed to be independent of  $t$ . Let  $t_1, t_2, \dots, t_k$  be the times at which the numbers of events in the two groups are observed; these could be the occurrence times of every event, or occur at less frequent intervals. At each observation time we can construct a  $2 \times 2$  contingency table for surviving patients at risk since the last observation time, recording treatment/control and event/no event. Let  $o_i$  be the number of new observed events in the control group at  $t_i$ , and  $e_i$  the expected number assuming  $\delta = 0$ . Let  $n$  be the total number of events up to time  $t_k$ . Let  $\bar{Z}_n = 4 \sum_i (o_i - e_i) / n$ .

It may be shown (see, for example, Spiegelhalter *et al* (1994)) that  $\bar{Z}_n$  is approximately distributed as  $N(\delta, \sigma^2/n)$ , where  $\sigma^2 = 4$ . The approximation is good for  $n$  large enough for convergence to normality, but small compared with the original group sizes.

Further applications to, for example, binary data with  $\delta$  defined as a difference in log odds, Poisson data with  $\delta$  defined as the log of the ratio of Poisson rates, or to ordinal data, are straightforward. Most of the necessary theory is given by Whitehead (1997).

## 6. Registration of the New Treatment

One of the key events in drug development and production is the *registration* of the new drug (treatment) for gaining a licence for marketing purposes.

It has already been mentioned that our methodology differs from the usual Bayesian procedure, by recognizing the existence of two decision makers, namely: the pharmaceutical company conducting the trial, which decides on its size; and the public at large, who determine ultimate usage. We now show that it is straightforward to include a regulator as another decision-maker, whose approval is necessary for the treatment to be licenced.

Clearly if the new treatment under consideration is not approved by the regulatory authorities it will not be used at all. Apart from the safety of the new treatment, we assume that the regulatory authorities require statistically significant evidence that its performance level exceeds some minimum.

In the following sections we shall show how the problem of registration of the new treatment may be handled with our model.

Let us assume that the new treatment will be approved if and only if  $\mu' \geq L + 1.5\tau'$  where  $L$  is the minimum amount of improvement required by the regulatory authorities. For the consistency of the notations we work with  $L\sigma$  in place of  $L$  (see (4.4)). Thus the condition may now be written as  $\mu'/\sigma \geq L + 1.5\tau'/\sigma$ .

Initially, and for the sake of simplicity, we suppose that the regulator's prior distribution is the same as the prior distribution for the company.

### 6.1. Public Health Benefit Function

For the public health benefit function, with the benefit of using the new treatment proportional to  $\delta$ , one of the three following cases must occur.

#### Case 1. $L \leq A < B$

In this case the regulator's decision whether or not to grant a licence has no influence on the number of users. This is because the licence is granted for all those outcomes for which the number of users would have been positive with no intervention by the regulator.

The function representing the number of subsequent users of the new treatment is unchanged and is of the form shown below

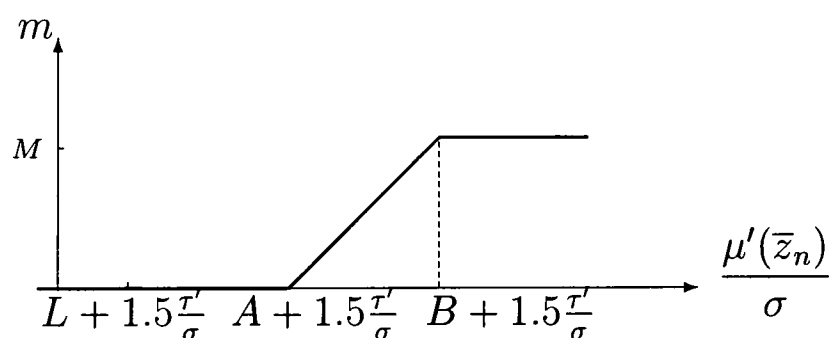


Figure 4.11: Number of subsequent users when  $L \leq A < B$ .

Obviously for this case the sample size problem is unchanged and the aim is still to find  $n^*$  which maximizes (4.5).

#### Case 2. $A < L < B$

This means that the minimum amount of improvement in performance required by the regulator to grant a licence is bigger than the one needed to

persuade some potential users to switch to the new treatment.

The function for the number of subsequent users of the new treatment is of the form shown below.

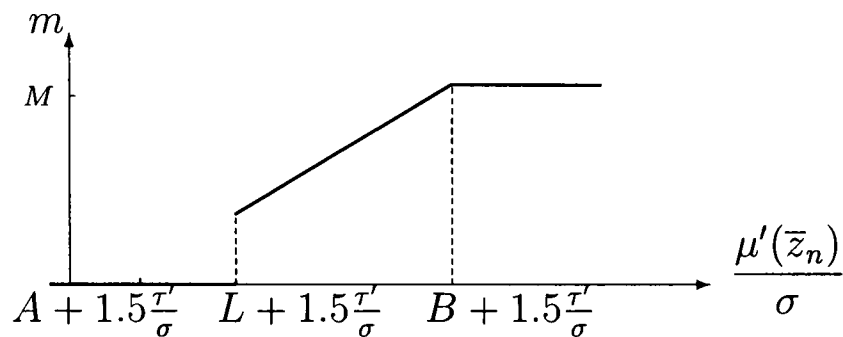


Figure 4.12: Number of subsequent users when  $A < L < B$ .

There are no subsequent users unless an approval is granted, in which case there are users, the minimum positive number being  $\frac{M}{B-A}(L-A)$ .

For this case the problem is changed and the aim is to find  $n^*$  which maximizes the following objective function

$$\begin{aligned}
 R(n) = & \frac{1}{B-A} \int_{h_1(L,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & - \frac{A}{B-A} \int_{h_1(L,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & - \frac{1}{B-A} 1.5(T^{-2} + n)^{-1/2} \int_{h_1(L,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \\
 & \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & + \int_{h_2(B,n)}^{\infty} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn.
 \end{aligned} \tag{4.10}$$

### Case 3. $A < B \leq L$

In this case the minimum amount of improvement in performance required by the regulator to grant a licence to the new treatment is more than enough to persuade all potential users to switch to it.



The function representing the number of subsequent users of the new treatment now has the form shown below

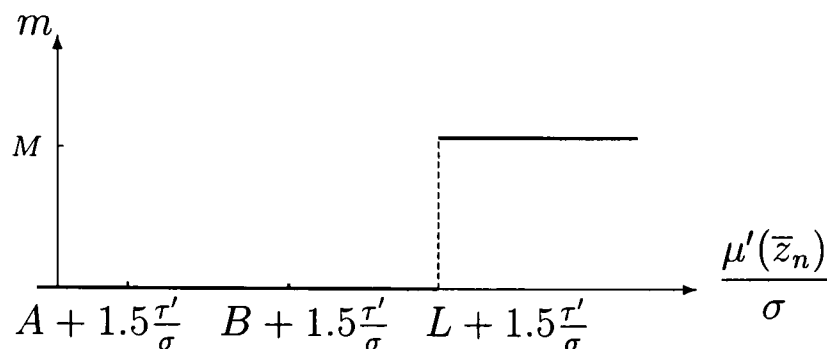


Figure 4.13: Number of subsequent users when  $A < B \leq L$ .

The aim is now to find  $n^*$  which maximizes the objective function

$$R(n) = \int_{h_2(L,n)}^{\infty} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn. \quad (4.11)$$

## 6.2. Commercial Benefit Function

The same three cases occur when the benefit function is from the standpoint of the company selling the new treatment.

**Case 1.**  $L \leq A < B$

The function representing the number of subsequent users of the new treatment,  $m$ , is unchanged and the objective function is as in (4.7).

**Case 2.**  $A < L < B$

$m$  now becomes as in figure 4.12 and the objective function is

$$\begin{aligned}
R(n) = & \frac{1}{B-A} \int_{h_1(L,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{h_1(L,n)}^{h_2(B,n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T^{-2} + n)^{-1/2} \int_{h_1(L,n)}^{h_2(B,n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{h_2(B,n)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn.
\end{aligned} \tag{4.12}$$

**Case 3.**  $A < B \leq L$

$m$  now becomes as in figure 4.13, and the objective function is

$$R(n) = \int_{h_2(L,n)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn. \tag{4.13}$$

□

## 7. The Regulator as a Third Decision-Maker

If the new treatment under consideration is not approved by the regulatory authorities, it will not be used at all. In section 6 we have shown how the problem of registration of the new treatment may be handled when it is assumed that the regulator uses the same prior distribution for  $\delta$  as the pharmaceutical company conducting the trial.

Here we assume that the regulator uses a prior distribution for  $\delta$  whose parameters are not necessarily the same as the one used by the pharmaceutical company. This will occur, in particular, if the regulator is more pessimistic than the company. To indicate that the prior distributions for  $\delta$  is the one for the company conducting the trial, from now on, we use  $\mu_c$ ,  $\tau_c$ ,  $\mu'_c$ , and  $\tau'_c$  in place of  $\mu$ ,  $\tau$ ,  $\mu'$ , and  $\tau'$ . The subscript  $r$  will indicate the corresponding parameters for the regulator's prior distribution.

Let us suppose that the regulator uses a prior density for  $\delta$  which is  $N(\mu_r, \tau_r^2)$ . Also, suppose that the new treatment will be approved if and only if

$$\mu'_r \geq L + 1.5\tau'_r, \tag{4.15}$$

in which  $L$  is the minimum amount of improvement required by the regulatory authorities. Using  $L\sigma$  in place of  $L$  in the above expression, (4.15) could be expressed as  $\bar{z}_n \geq H_r(L, n)$ , where

$$H_r(L, n) = \frac{\sigma(\sigma^2 + n\tau_r^2)(L + 1.5\frac{\tau_r'}{\sigma}) - \sigma^2\mu_r}{n\tau_r^2}.$$

### 7.1. The Regulator and the Public Health Benefit Function

For the public health benefit function, with the benefit of using the new treatment proportional to  $\delta$ , one of the three following cases must occur.

**Case 1.**  $H_r(L, n) \leq H_1(A, n) < H_2(B, n)$

In this case the regulator's decision whether or not to grant a licence has no influence on the number of users. This is because the licence is granted for all those outcomes for which the number of users would have been positive with no intervention by the regulator.

The function representing the number of subsequent users of the new treatment is unchanged and is of the form shown below

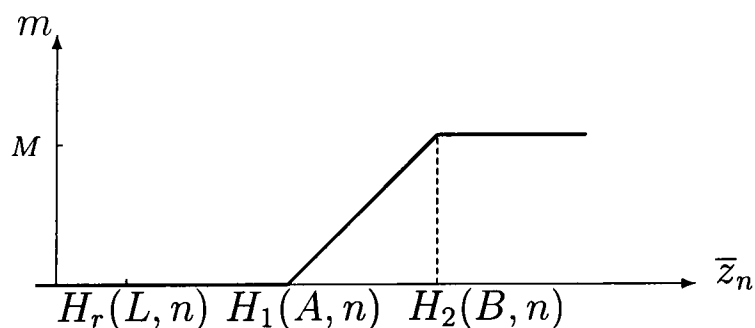


Figure 4.14: Number of subsequent users when  
 $H_r(L, n) \leq H_1(A, n) < H_2(B, n)$

Obviously for this case the sample size problem is unchanged and the aim is to find  $n^*$  which maximizes the objective function

$$\begin{aligned}
R(n) = 0 + \int_{H_1(A,n)}^{H_2(B,n)} \frac{1}{B-A} & \left[ \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n - A - 1.5 \frac{\tau_c'}{\sigma} + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right] \\
& \times \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n \\
& + \int_{H_2(B,n)}^{\infty} \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n - Cn,
\end{aligned} \tag{4.16}$$

where  $\bar{Z}_n \sim N(\mu_c, \frac{\sigma^2}{n} + \tau_c^2)$ .

Writing  $u = \frac{\bar{z}_n - \mu_c}{\sqrt{\frac{\sigma^2}{n} + \tau_c^2}}$ , the objective function could be written as

$$\begin{aligned}
R(n) = \frac{1}{B-A} \int_{h_1(A,n)}^{h_2(B,n)} & [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{h_1(A,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{h_1(A,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \\
& \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{h_2(B,n)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn,
\end{aligned} \tag{4.17}$$

where  $T_c = \frac{\tau_c}{\sigma}$ ,  $D_c = \frac{\mu_c}{\sigma}$ ,  $C = \frac{c}{Mb}$ ,  $R(n) = \frac{r(n)}{Mb}$ , and

$$\begin{aligned}
h_1(A,n) &= \frac{[A + 1.5(T_c^{-2} + n)^{-1/2} - D_c](T_c^{-2} + n)^{1/2}}{T_c \sqrt{n}}, \\
h_2(B,n) &= \frac{[B + 1.5(T_c^{-2} + n)^{-1/2} - D_c](T_c^{-2} + n)^{1/2}}{T_c \sqrt{n}}.
\end{aligned}$$

**Case 2.**  $H_1(A,n) < H_r(L,n) < H_2(B,n)$

In this case it is assumed that the minimum amount of improvement in

performance required by the regulator to grant a licence is bigger than the one assumed by the company conducting the trial.

Here there are some potential users for outcomes which are unacceptable to the regulator, and the function for the number of subsequent users of the new treatment is of the form shown below.

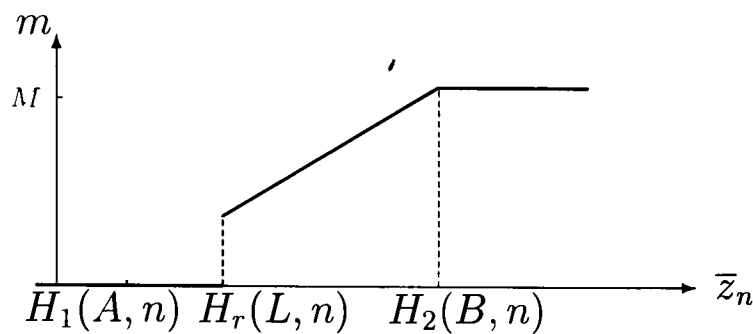


Figure 4.15: Number of subsequent users when  $H_1(A, n) \leq H_r(L, n) < H_2(B, n)$ .

As shown in figure 4.15, there are subsequent users if and only if an approval is granted.

For this case the problem is changed and the aim is to find  $n^*$  which maximizes the following objective function

$$\begin{aligned}
 R(n) = 0 + \int_{H_r(L, n)}^{H_2(B, n)} \frac{1}{B - A} & \left[ \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n - A - 1.5 \frac{\tau_c'}{\sigma} + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right] \\
 & \times \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n \\
 & + \int_{H_2(B, n)}^{\infty} \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n - Cn.
 \end{aligned} \tag{4.18}$$

Using the change of variable as in (4.17), the objective function may be written as

$$\begin{aligned}
R(n) = & \frac{1}{B-A} \int_{h_r(L,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{h_r(L,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{h_r(L,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \\
& \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{h_2(B,n)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn,
\end{aligned} \tag{4.19}$$

where

$$\begin{aligned}
h_r(L, n) = & \frac{(1 + nT_r^2)\{L + 1.5(T_r^{-2} + n)^{-1/2}\} - D_r - nT_r^2 D_c}{T_c T_r^2 \sqrt{n}(T_c^{-2} + n)^{1/2}}, \\
D_r = & \frac{\mu_r}{\sigma}, \quad T_r = \frac{\tau_r}{\sigma}.
\end{aligned}$$

**Case 3.**  $H_1(A, n) < H_2(B, n) \leq H_r(L, n)$

In this case the minimum amount of improvement in performance required by the regulator to grant a licence to the new treatment, is bigger than the maximum amount of the one assumed by the company conducting the trial.

For this case the function representing the number of subsequent users of the new treatment is changed to the form shown below

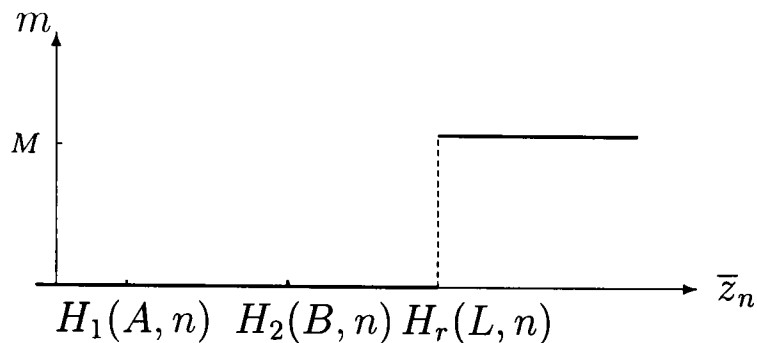


Figure 4.16: Number of subsequent users when  $H_1(A, n) < H_2(B, n) \leq H_r(L, n)$ .

As shown in figure 4.16, as soon as approval is granted all the expected subsequent users of the new treatment,  $M$ , will appear.

The aim is now to find  $n^*$  which maximizes the objective function

$$R(n) = \int_{H_r(L, n)}^{\infty} \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n - Cn, \quad (4.20)$$

which may be written as

$$R(n) = \int_{h_r(L, n)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn. \quad (4.21)$$

Now we can establish the following theorem.

**Theorem:** If the regulator's requirement is as in (4.15) and the benefit of using the new treatment is from the public health standpoint, then the scaled expected net benefit from carrying out a trial of size  $n$  is

$$\begin{aligned} R(n) = & \frac{1}{B-A} \int_{\max(h_1, h_r)}^{\max(h_2, h_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & - \frac{A}{B-A} \int_{\max(h_1, h_r)}^{\max(h_2, h_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{\max(h_1, h_r)}^{\max(h_2, h_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \\ & \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & + \int_{\max(h_2, h_r)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn. \end{aligned} \quad (4.22)$$

## 7.2. The Regulator and the Commercial Benefit Function

The same analysis as in 7.1 may be applied to show that when the benefit function is written from the standpoint of the pharmaceutical company conducting the trial, the registration problem may be dealt with by our general model by considering three different cases. These are listed below.

**Case 1.**  $H_r(L, n) \leq H_1(A, n) < H_2(B, n)$

The function representing the number of subsequent users of the new treatment,  $m$ , is unchanged and the scaled expected net benefit is

$$\begin{aligned}
 R(n) = 0 + \int_{H_1(A, n)}^{H_2(B, n)} \frac{1}{B - A} & \left[ \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n - A - 1.5 \frac{\tau'_c}{\sigma} \right. \\
 & \left. + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right] f(\bar{z}_n) d\bar{z}_n \\
 & + \int_{H_2(B, n)}^{\infty} f(\bar{z}_n) d\bar{z}_n - Cn,
 \end{aligned} \tag{4.23}$$

which may be written as

$$\begin{aligned}
 R(n) = \frac{1}{B - A} \int_{h_1(A, n)}^{h_2(B, n)} & [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & - \frac{A}{B - A} \int_{h_1(A, n)}^{h_2(B, n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & - \frac{1}{B - A} 1.5(T_c^{-2} + n)^{-1/2} \int_{h_1(A, n)}^{h_2(B, n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
 & + \int_{h_2(B, n)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn,
 \end{aligned} \tag{4.24}$$

where  $h_1(A, n)$  and  $h_2(B, n)$  are as in (4.17).

**Case 2.**  $H_1(A, n) < H_r(L, n) < H_2(B, n)$

$m$  now becomes as in figure 4.15 and the scaled expected net benefit is



$$\begin{aligned}
R(n) = 0 + \int_{H_r(L,n)}^{H_2(B,n)} \frac{1}{B-A} \left[ \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n - A - 1.5 \frac{\tau_c'}{\sigma} \right. \\
\left. + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right] f(\bar{z}_n) d\bar{z}_n \\
+ \int_{H_2(B,n)}^{\infty} f(\bar{z}_n) d\bar{z}_n - Cn.
\end{aligned} \tag{4.25}$$

This may be written as

$$\begin{aligned}
R(n) = \frac{1}{B-A} \int_{h_r(L,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
- \frac{A}{B-A} \int_{h_r(L,n)}^{h_2(B,n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
- \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{h_r(L,n)}^{h_2(B,n)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
+ \int_{h_2(B,n)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn.
\end{aligned} \tag{4.26}$$

**Case 3.**  $H_1(A, n) < H_2(B, n) \leq H_r(L, n)$

$m$  now becomes as in figure 4.16, and the scaled expected net benefit is

$$R(n) = \int_{H_r(L,n)}^{\infty} f(\bar{z}_n) d\bar{z}_n - Cn. \tag{4.27}$$

This may be written as

$$R(n) = \int_{h_r(L,n)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn. \tag{4.28}$$

Now we can establish the following theorem.

**Theorem:** If the regulator's requirement is as in (4.15) and the benefit of using the new treatment is expressed from the commercial standpoint, then the scaled expected net benefit for carrying out a trial of size  $n$  may be

written as

$$\begin{aligned}
R(n) = & \frac{1}{B-A} \int_{\max(h_1, h_r)}^{\max(h_2, h_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{\max(h_1, h_r)}^{\max(h_2, h_r)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{\max(h_1, h_r)}^{\max(h_2, h_r)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{\max(h_2, h_r)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn.
\end{aligned} \tag{4.29}$$

□

## 8. Two Case Studies

The tables and figures which follow set out a detailed description of using BeBay to calculate retrospectively ‘optimal’ sample sizes for trials 1 and 2 of those 6 trials described in chapter 1. These trials were recently carried out by a UK-based pharmaceutical company which for commercial reasons its name has been omitted. The results are then contrasted with the sample sizes which were actually used, calculated on the basis of the size and power of the resulting significance tests (the 5% and 80% chances mentioned there are the quantities which are defined to be the *size* and *power* of the test). Trial 1 was for normally distributed data, and trial 2 for binomially distributed data.

Since no direct information on prior distributions was available it was decided, after consulting the company concerned, to make the following reasonably realistic assumptions for each trial for the means  $\mu_c$  and  $\mu_r$ , and standard deviations  $\tau_c$  and  $\tau_r$ , of the company’s and the regulator’s prior distributions for  $\delta$ , the average difference in the outcomes for the two treatments being compared.

$\mu_c$  = the clinically relevant difference in outcomes used in the original calculation of sample size,  $\tau_c = \mu_c/2$ ,  $\mu_r = 0$ ,  $\tau_r = \mu_c$ . It was also assumed, and

again this seemed realistic, that  $A = 0.8\mu_c$ ,  $B = 1.2\mu_c$ , and that  $2\mu_c/3$  and  $\mu_c$  were low and high estimates for  $L$  (see section 7). The low estimate for the evidence required by the regulator amounts to an assumption that the regulator's decision will be positive even when the apparent improvement from using the new treatment is barely enough to persuade potential customers to use it. This may well be a realistic assumption for the both trials, as in every case the drug concerned was already on the market, so it is plausible that approval for the additional use may have been obtainable more easily than would have been the case for a new drug.

For trial 1, the value  $\sigma = 0.3$  was taken from the original calculation of sample size. For the binomial data  $\bar{Z}_n$  was redefined as  $\log[\frac{1}{n} \sum X_i / (1 - \frac{1}{n} \sum X_i)] - \log[\frac{1}{n} \sum Y_i / (1 - \frac{1}{n} \sum Y_i)]$  and  $\delta$  as  $\log[p_1 / (1 - p_1)] - \log[p_2 / (1 - p_2)]$ , where  $p_1 = P_r(\text{cure} \mid \text{new treatment})$  and  $p_2 = P_r(\text{cure} \mid \text{alternative treatment})$ . Now  $X_i = 1$  if the  $i$ 'th patient responds favourably to the new treatment, and otherwise  $X_i = 0$ , and  $Y_i$  is defined similarly for the  $i$ 'th patient receiving the alternative treatment. Inference was based on the fact that  $\bar{Z}_n$  has a distribution which is approximately normal with the mean  $\delta$  and the variance  $\frac{2(p_1(1-p_1)+p_2(1-p_2))}{np_1(1-p_1)p_2(1-p_2)}$  (see Coad and Rosenberger (1999)). For rare events, Spiegelhalter *et al* (1994) suggest an approximation of  $4/n$  for the variance of  $\bar{Z}_n$ .

The cost per patient in the trial was assumed, on the basis of the company's previous experience, to be £600 for out-patients in trial 1 and £4000 for hospitalised patients in trial 2. A more careful cost analysis would have included a set up cost for the trial in addition to the cost per patient. This would simply have reduced the objective function  $r(n)$  by the amount of the setup cost, leaving the optimal trial size unchanged, provided it remained worthwhile to carry out a trial. Each of the two trials was for an existing drug applied to a new class of patients (for example juveniles, or patients suffering from the effects of chemotherapy). Thus actual sales and profit figures were available. High and low estimates for the total benefit from a conclusively favourable trial were obtained by assuming either a 50% or a

5% increase in annual profit, and discounting the resulting cash flow to give a figure for its present value.

The function  $r(n)$  was maximized and the optimal (i.e. maximizing) sample size recorded. The probability of detecting the clinically relevant difference in outcomes (i.e. the power) of a hypothesis test with 5% probability of detecting a non-existent difference in performance between the new treatment and the placebo was also calculated. For each trial four calculations were carried out: for high and low estimates for the strength of the evidence required by the regulator, and for high and low estimates of the total benefit from a conclusively favourable trial.

The results, not surprisingly, show that the optimal sample size depends strongly on the expected benefit from a conclusively favourable outcome, and on the strength of the evidence required by the regulator. In practice it should be possible to estimate both of these much more accurately than the wide intervals, based on ball-park figures, which have been used here for illustrative purposes.

It is also noticeable that the optimal sample size based on our analysis is frequently much larger than the size used in practice. For large sample sizes an interim analysis is very desirable, as it has the potential both to reduce sampling costs and to advance the launch date of the new therapy. The results for trial 2 show that our analysis can also indicate small sample sizes.

## ● Trial 1; A trial for the suitability of an existing drug for juveniles, carried out by GP's

A trial for the suitability of a certain treatment for juveniles was carried out by a pharmaceutical company using GP's. Based on earlier experience the standard deviation of change in total score assumed to be 0.3. BeBay was run with the following parameter values.

$\sigma = 0.3$ . This was based on the company's previous experience with similar

trials.

$\mu_c = 0.15$ . The company view of the likely difference in the performance of the treatments

$\mu_r = 0$ . The regulator's view of the likely difference in the performance of the treatments.

$\tau_c = \mu_c/2$ . A convenient representation of prior ignorance. It means that

$$P(\delta < 0) = P(\delta > 2\mu) \approx 0.025.$$

$\tau_r = 2\tau_c = 0.15$ . The regulator's prior standard deviation.

$c = \text{£}600$ . The total cost per patient, to include payment to the GP and administration cost.

$A = 0.8\mu_c = 0.12$ . Maximum critical difference for no sale.

$B = 1.2\mu_c = 0.18$ . Minimum critical difference for maximum sale.

$L = 0.1, 0.15$ . Minimum regulator critical difference for granting a licence.

$Mb = \text{£}25M, \text{£}250M$ . A couple of estimates, arrived at along the following lines, for the total additional return if the trial gives a clear positive result.

Let

$I$  = Current annual income from the treatment,

$I(1 + x)$  = Future annual income if this trial gives a clear positive result,

$(1 + r)$  = (value of being able to show now that the trial result is clearly positive)/(value of being able to show in one year from now that the trial result is clearly positive).

Thus  $r$  is a discount rate. It is convenient to express it in 'real' terms, with the effect of inflation removed. Its main components are attributable to obsolescence because of advances in alternative therapies and to reduction of the remaining period of patent protection.

Thus the expected net benefit if the trial is clearly successful is

$$Mb = Ix(1 + \frac{1}{1+r} + \frac{1}{(1+r)^2} + \dots) = Ix(1 + \frac{1}{r}).$$

For example, if  $I = \pounds 15M$ ,  $x = 0.1$ , and  $r = 0.1$ , then  $Mb = \pounds 16.5M$ .

The following table illustrates the results of sample size calculations for this trial.

**Trial 1**  
Actual Number of Patients, 100

		<i>Low Regulator Requirement</i>	<i>High Regulator Requirement</i>
<i>Low Maximum Benefit</i>	Optimal (BeBay) No. of Patients	1319	1350
	Maximized expected net benefit	£10.11M	£9.09M
	Power	1.00	1.00
<i>High Maximum Benefit</i>	Optimal (BeBay) No. of Patients	6165	6270
	Maximized expected net benefit	£113.9M	£103.9M
	Power	1.00	1.00

As shown by the table the cost/benefit ratio  $\frac{c}{Mb}$  has a considerable influence on the optimal sample size.

● **Trial 2; A trial for the suitability of the drug for treating severely ill hospitalised patients**

This is another trial using the same treatment as in trial 1, but now we are observing a 0–1 quantity (some hair loss=0, no hair loss=1) for severely ill hospitalized patients. We work in term of the log odds-ratio  $\delta$  and assume central limit theorem normality to calculate the optimal sample size which maximizes the expected net benefit resulting from the trial.

$\sigma = 2.$

$\mu_c = 2.09.$  This is the log odds-ratio with success probabilities of 0.2 with placebo and 0.67 with the active cream, which we assume to be the expected value.

$\mu_r = 0, \tau_c = 1.045,$  and  $\tau_r = 2.09.$  As before.

$c = \pounds 4K.$  Larger costs per patient than the  $\pounds 600$  assumed for the other case, on the grounds that these are hospital patients.

$A = 1.672, B = 2.508,$  and regulator requirements (low)  $L = 1.5,$  (intermediate)  $L = 2.09.$

$Mb = \pounds 5M, \pounds 25M.$  As before.

The following table shows optimal sample sizes, maximized expected net benefit and the expected power for this trial.

<div>Trial 2</div> <div>Actual Number of Patients, 20</div>			
		<i>Low Regulator Requirement</i>	<i>High Regulator Requirement</i>
<i>Low Maximum Benefit</i>	Optimal (BeBay) No. of Patients	76	78
	Maximized expected net benefit	£1.57M	£1.35M
	Power	1.00	1.00
<i>High Maximum Benefit</i>	Optimal (BeBay) No. of Patients	228	233
	Maximized expected net benefit	£9.57M	£8.71M
	Power	1.00	1.00

Figure 4.17 shows how the expected net benefit varies as a function of the size of the trial. The relationship is shown for the same four cases as in previous tables for each of two trials. As shown by figure 4.17, when the regulator requires a milder condition for granting a licence to the new treatment the expected net benefit is increased.

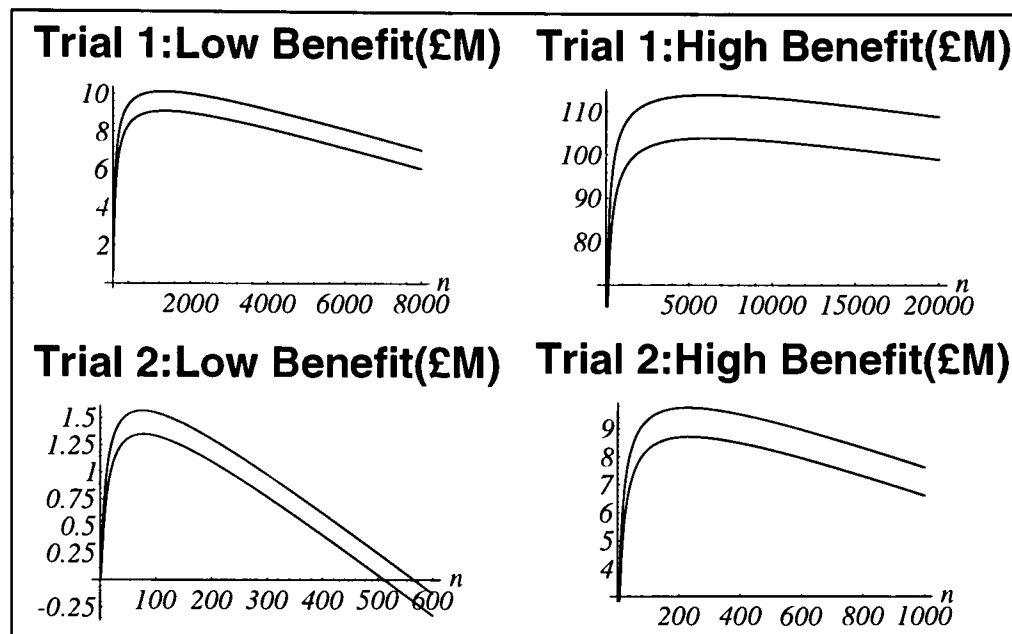


Figure 4.17: Expected net benefit as a function of  $n$  for the two trials. The upper and lower curves on each plot are for the lower and higher regulatory requirements respectively.

**Lemma:** Probability of approval of the new treatment is

$$P = 1 - \Phi(h_r(L, n)) \quad (4.30)$$

**Proof:** Note that

$$\text{Probability of approval} = P_r(\mu'_r(\bar{Z}_n)/\sigma > L + 1.5\tau'_r/\sigma)$$

$$= P_r(\bar{Z}_n > \frac{(L + 1.5\sqrt{\frac{\sigma^2\tau_r^2}{\sigma^2 + n\tau_r^2}})(\sigma^2 + n\tau_r^2) - \sigma^2\mu_r}{n\tau_r^2}).$$

Using the same change of variable as in (4.19), after some calculation we arrive at (4.30).  $\square$



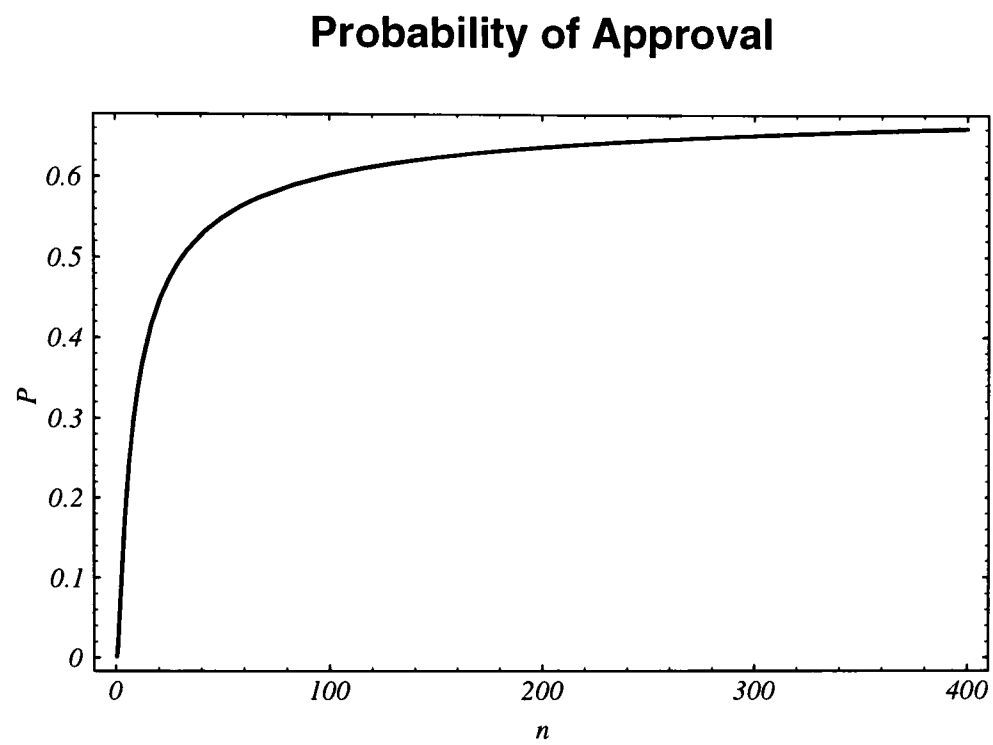


Figure 4.18: Probability of approval as a function of  $n$ .

Figure 4.18 shows the probability that the regulator approves the treatments in the second trial and grants a licence to it. This is an increasing function of  $n$ , as we might have expected.

# Chapter 5

## A Behavioural Bayes Approach to the Sample Size Question; Normal Distribution, Unknown Variance

In this chapter we discuss the sample size problem for the case where the data are assumed to come from a normal distribution for which the mean and variance are both unknown.

### 1. Introduction

Our general model (or BeBay approach) was formulated on the assumption of paired comparison of normally distributed effects with a known variance, but we have shown how it could be applied to a rather wide class of other cases without these restrictions (see section 5 of chapter 4). As mentioned before the known variance is often a realistic and applicable assumption but in this chapter this will be relaxed.

Frequently both the mean and the variance of the observations are unknown, rather than just the mean. So we consider the case where we have a set of clinical outcomes  $Z_i$ ,  $i = 1, 2, \dots$  which are distributed independently as  $N(\delta, \sigma^2)$  with  $\delta$  and  $\sigma^2$  both unknown. To deal with sample size determination for this case we consider a joint prior density for  $\delta$  and  $\sigma^2$ .

In section 2 we state the sample size problem and introduce the notation.

In section 3 we establish our objective function which is the expected net benefit. The optimal size of a clinical trial is obtained with a benefit function which seems appropriate from the public health standpoint.

Section 4 deals with the problem for which the benefit function is defined from the standpoint of the company selling the drug.

In section 5 the distribution of the number of subsequent users of the new treatment is discussed.

The regulatory authority's requirements for granting a licence to the new treatment are discussed in section 6.

Section 7 shows the consistency of the calculation by illustrating the convergence of the reward function for a sequence of unknown variance cases to the reward function for the limiting case in which  $\sigma^2$  is known.

## 2. Preliminaries

Consider the sample size question stated in chapter 4 in which  $Z_i = X_i - Y_i$  ( $i = 1, 2, \dots$ ) have a normal density which is  $N(\delta, \sigma^2)$  where both  $\delta$  and  $\sigma^2$  are unknown.

Following O'Hagan (1994), let us assume that the joint prior density function for  $(\delta, \sigma^2)$  is of the form shown below

$$\pi(\delta, \sigma^2) = k(a, g, w)(\sigma^2)^{-\frac{g+3}{2}} e^{-\frac{1}{2\sigma^2}\{a+(\delta-\mu)^2/w\}}, \quad (5.1)$$

where

$$k(a, g, w) = a^{\frac{g}{2}} 2^{-\frac{g+1}{2}} (\pi w)^{-1/2} \{\Gamma(g/2)\}^{-1}.$$

Thus the marginal prior density functions for  $\delta$  and  $\sigma^2$  are

$$\pi(\delta) = \frac{1}{B(1/2, g/2)} (aw)^{-1/2} \left\{1 + \frac{(\delta - \mu)^2}{aw}\right\}^{-\frac{g+1}{2}}, \quad (5.2)$$

where  $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ , and

$$\pi(\sigma^2) = \left(\frac{a}{2}\right)^{\frac{g}{2}} \frac{1}{\Gamma(g/2)} (\sigma^2)^{-\frac{g+2}{2}} \exp\left(-\frac{a}{2\sigma^2}\right). \quad (5.3)$$

The latter is a kind of inverse chi-squared distribution, because the distribution of  $a/\sigma^2$  is chi-squared with  $g$  degrees of freedom.

Applying Bayes' theorem the posterior density turns out to be

$$\pi^n(\delta, \sigma^2 | \mathbf{z}) = k(a_1, g_1, w_1)(\sigma^2)^{-\frac{g_1+3}{2}} e^{-\frac{1}{2\sigma^2}\{a_1 + (\delta - \mu_1)^2/w_1\}}, \quad (5.4)$$

where  $z_i = x_i - y_i$  ( $i = 1, 2, \dots, n$ ),  $\mathbf{z}$  is the vector of observations,  $\bar{z}_n = \frac{1}{n} \sum_i z_i$ ,  $s^2 = \sum_i (z_i - \bar{z}_n)^2$  and

$$\begin{aligned} w_1 &= \frac{w}{1 + nw}, & \mu_1 &= \frac{\mu + nw\bar{z}_n}{1 + nw} (= \mu'), \\ g_1 &= g + n, & a_1 &= a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1 + nw}. \end{aligned}$$

So the joint prior distribution function is conjugate in the sense that the posterior distribution has the same form as the prior.

The marginal posterior density function of  $\delta$  can be easily found by integrating over the joint posterior density function and using the substitution  $y = \{a_1 + (\delta - \mu)^2/w_1\}/(2\sigma^2)$ . We find that the posterior density of  $\delta$  is

$$\pi^n(\delta | \mathbf{z}) = (w_1 a_1)^{-1/2} \frac{1}{B(1/2, g_1/2)} \{1 + (\delta - \mu_1)^2/w_1 a_1\}^{-\frac{g_1+1}{2}}. \quad (5.5)$$

Therefore  $(\delta - \mu_1)/\sqrt{w_1 a_1/g_1}$  has a Student t-distribution with  $g_1$  degrees of freedom. So the posterior mean and variance of  $\delta$  (provided  $g_1 > 2$ ) are

$$\mu'(\bar{z}_n) = \frac{\mu + nw\bar{z}_n}{1 + nw}, \quad \tau'^2(\bar{z}_n, s^2) = \frac{w_1 a_1}{g_1 - 2}.$$

It is also easy to show that the posterior density function of  $\sigma^2$  is of the form

$$\pi^n(\sigma^2 | \mathbf{z}) = \frac{(\frac{a_1}{2})^{g_1/2}}{\Gamma(g_1/2)} (\sigma^2)^{-\frac{g_1+2}{2}} \exp(-\frac{a_1}{2\sigma^2}). \quad (5.6)$$

As in the known  $\sigma^2$  we suppose that  $m$ , the number of subsequent users, depends on the mean  $\mu'$  and the standard deviation  $\tau'$  of the posterior distribution for  $\delta$  as shown below.

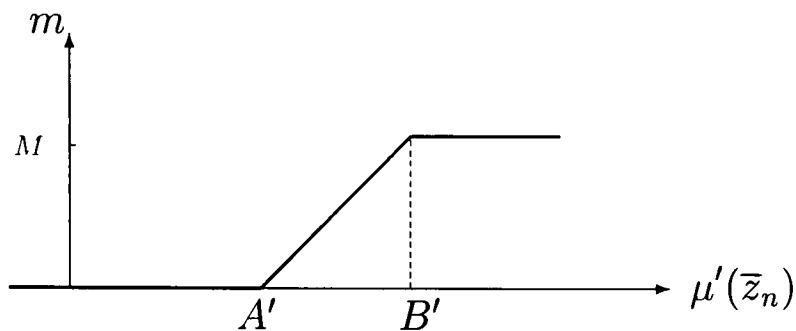


Figure 5.1: Number of subsequent users,  
where  $A' = A + 1.5\tau'(\bar{z}_n, s^2)$  and  $B' = B + 1.5\tau'(\bar{z}_n, s^2)$ .

$M$ ,  $A$ , and  $B$  are as described in chapter 4.

We are now in a position to write our objective function for our two versions of the benefit function in turn, and give some numerical examples.

### 3. Public Health Benefit Function

If the benefit of using the new treatment is  $b\delta$ , ( $b \geq 0$ ) then the expected net benefit can be written as

$$\begin{aligned} r(n) &= \int_{s^2} \int_{\bar{z}_n} \int_{\sigma^2} \int_{\delta} mb\delta f(\delta, \sigma^2, \bar{z}_n, s^2) d\delta d\sigma^2 d\bar{z}_n ds^2 - cn \\ &= \int_{s^2} \int_{\bar{z}_n} \int_{\sigma^2} \int_{\delta} mb\delta \pi^n(\delta, \sigma^2 | \bar{z}_n, s^2) f(\bar{z}_n, s^2) d\delta d\sigma^2 d\bar{z}_n ds^2 - cn, \end{aligned}$$

or finally

$$r(n) = \int_{s^2} \int_{\bar{z}_n} mb\mu' f(\bar{z}_n, s^2) d\bar{z}_n ds^2 - cn. \quad (5.7)$$

To calculate  $f(\bar{z}_n, s^2)$  we use the fact that

$$f(\bar{z}_n, s^2) = \int_{\sigma^2} \int_{\delta} \pi(\delta, \sigma^2) f(\bar{z}_n | \delta, \sigma^2) f(s^2 | \delta, \sigma^2) d\delta d\sigma^2,$$

where

$$\begin{aligned} f(\bar{z}_n | \delta, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2/n}} e^{-\frac{n}{2\sigma^2}(\bar{z}_n - \delta)^2}, \\ f(s^2 | \delta, \sigma^2) &= \left(\frac{1}{2\sigma^2}\right)^{\frac{n-1}{2}} \frac{1}{\Gamma(\frac{n-1}{2})} (s^2)^{\frac{n-1}{2}-1} e^{-\frac{1}{2\sigma^2}s^2}. \end{aligned}$$

After some calculations it turns out that

$$f(\bar{z}_n, s^2) = \frac{1}{B(1/2, (g+n)/2)} \left\{ \frac{a^{g/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \times (s^2)^{\frac{n-1}{2}-1} \left( a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw} \right)^{-\frac{g+n}{2}}. \quad (5.8)$$

Let the number of subsequent users of the new treatment be of the form shown in figure 5.1, so that

$$m = \begin{cases} 0 & \mu' < A' \\ \frac{M}{B-A}[\mu' - A'] & A' < \mu' < B' \\ M & B' < \mu'. \end{cases}$$

Thus equation (5.7) may be written

$$r(n) = \int_0^\infty \int_{H_1}^{H_2} \frac{Mb}{B-A} \{ \mu' - A - 1.5 \sqrt{\frac{w_1 a_1}{g_1 - 2}} \} \mu' f(\bar{z}_n, s^2) d\bar{z}_n ds^2 + Mb \int_0^\infty \int_{H_2}^\infty \mu' f(\bar{z}_n, s^2) d\bar{z}_n ds^2 - cn,$$

or

$$R(n) = \frac{1}{(B-A)B(1/2, (g+n)/2)} \left\{ \frac{a^{g/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \int_0^\infty \int_{H_1}^{H_2} \left\{ \frac{\mu + nw\bar{z}_n}{1+nw} - A - 1.5 \sqrt{\frac{\frac{n}{1+nw}(a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw})}{g+n-2}} \right\} \left( \frac{\mu + nw\bar{z}_n}{1+nw} \right) (s^2)^{\frac{n-1}{2}-1} \times \left( a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw} \right)^{-\frac{g+n}{2}} d\bar{z}_n ds^2 + \frac{1}{B(1/2, (g+n)/2)} \left\{ \frac{a^{g/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \int_0^\infty \int_{H_2}^\infty \left( \frac{\mu + nw\bar{z}_n}{1+nw} \right) \times (s^2)^{\frac{n-1}{2}-1} \left( a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw} \right)^{-\frac{g+n}{2}} d\bar{z}_n ds^2 - Cn, \quad (5.9)$$

where  $R(n) = \frac{r(n)}{Mb}$ ,  $C = \frac{c}{Mb}$ , and  $H_1$  and  $H_2$  are the values of  $\bar{z}_n$  for which, respectively,

$$\bar{z}_n = \frac{(A + 1.5 \sqrt{\frac{w_1 a_1}{g_1 - 2}})(1 + nw) - \mu}{nw}$$

and

$$\bar{z}_n = \frac{(B + 1.5\sqrt{\frac{w_1 a_1}{g_1 - 2}})(1 + nw) - \mu}{nw}.$$

Numerical calculations show that  $r(n)$  has a unique maximum as a function of  $n$ .

The optimal sample size  $n^*$  for the case shown by figure 5.2 is 44.

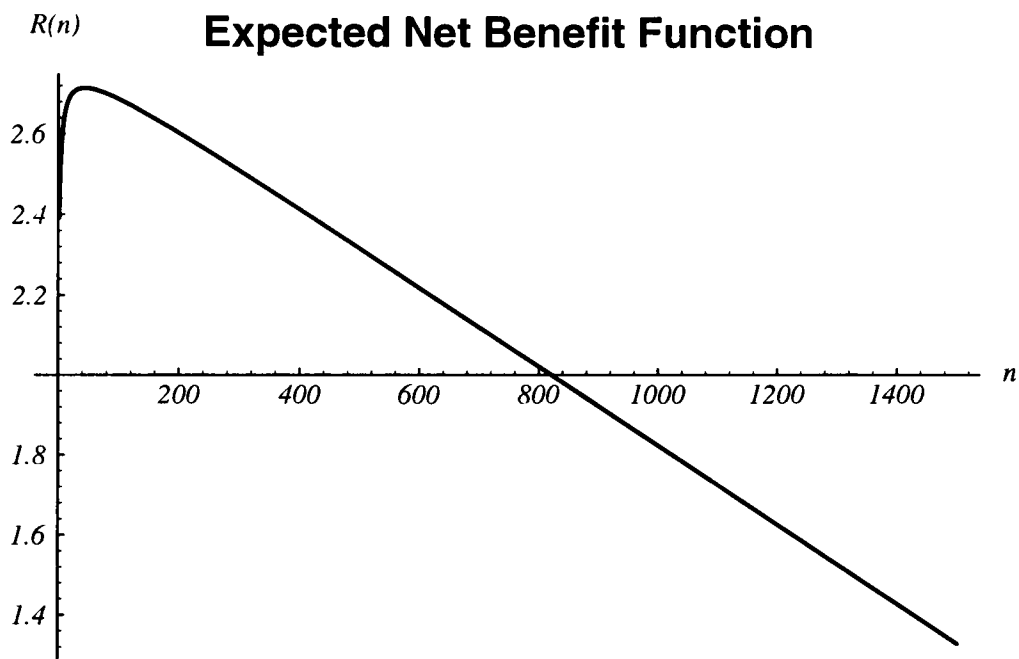


Figure 5.2: Expected net benefit when  $a = w = 1$ ,  $\mu = 3$ ,  $g = 5$ ,  $\frac{c}{Mb} = 0.001$ , and  $(A, B) = (2, 2.5)$ ; public health benefit function.

## 4. Commercial Benefit Function

If we assume that for each patient using the new treatment there is a benefit  $b(\geq 0)$  independent of  $\delta$ , then

$$r(n) = \int_{s^2} \int_{\bar{z}_n} mbf(\bar{z}_n, s^2) d\bar{z}_n ds^2 - cn. \quad (5.10)$$

By proceeding as in section 3 it can be shown that the objective (or expected net benefit) function is

$$\begin{aligned} R(n) = & \frac{1}{(B-A)B(1/2, (g+n)/2)} \left\{ \frac{a^{g/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \int_0^\infty \int_{H_1}^{H_2} \left\{ \frac{\mu + nw\bar{z}_n}{1+nw} \right. \\ & - A - 1.5 \sqrt{\frac{\frac{n}{1+nw}(a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw})}{g+n-2}} \left. \right\} (s^2)^{\frac{n-1}{2}-1} \\ & \times \left( a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw} \right)^{-\frac{g+n}{2}} d\bar{z}_n ds^2 \\ & + \frac{1}{B(1/2, (g+n)/2)} \left\{ \frac{a^{g/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \int_0^\infty \int_{H_2}^\infty (s^2)^{\frac{n-1}{2}-1} \\ & \times \left( a + s^2 + \frac{n(\bar{z}_n - \mu)^2}{1+nw} \right)^{-\frac{g+n}{2}} d\bar{z}_n ds^2 \\ & - Cn. \end{aligned} \quad (5.11)$$

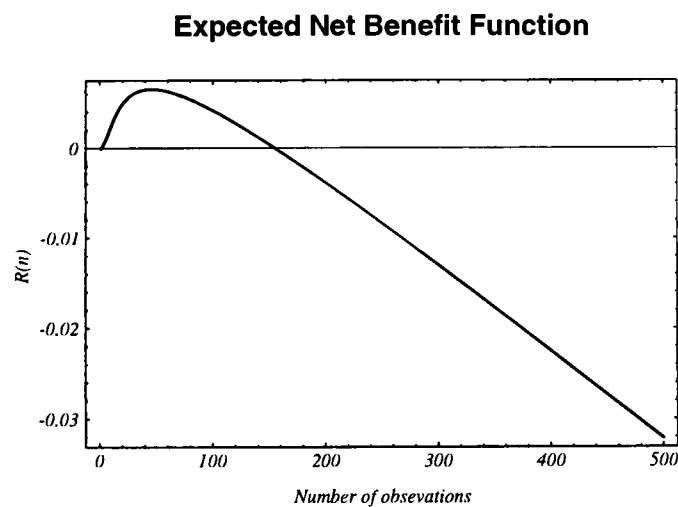


Figure 5.3: Expected net benefit when  $a = w = 1$ ,  $\mu = 1$ ,  $g = 5$ ,  $\frac{c}{Mb} = 0.0001$ , and  $(A, B) = (2, 2.5)$ ; commercial benefit function.



Figure 5.3 illustrates the variation of (5.11) as a function of  $n$ . The optimal sample size  $n^*$  for this case is 48. It should be noted that in (5.11)  $R(n)$ ,  $C$ ,  $H_1$ , and  $H_2$  are as in (5.9), though the meaning of  $b$  is now different.

## 5. Distribution of the Number of Subsequent Users of the New Treatment

The same reasoning as in the known variance case suggests that when  $n = 0$ , the number of subsequent users of the new treatment may be of the form shown below.

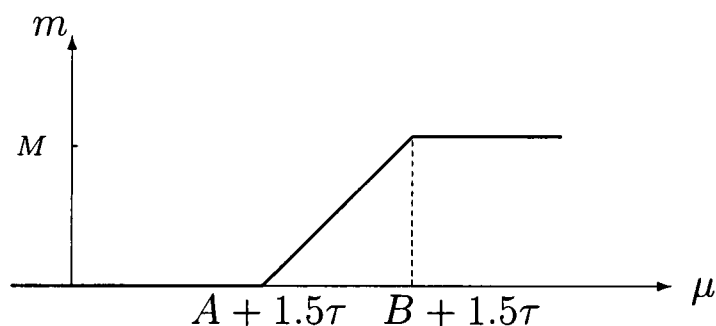


Figure 5.4: Number of subsequent users.

Here  $\mu$  is the prior mean and  $\tau = \sqrt{\frac{wa}{g-2}}$  is the standard deviation of the prior distribution for  $\delta$ . If the prior knowledge about  $\delta$  indicates that  $\mu < A$  and we decide not to conduct a trial ( $n = 0$ ) then  $\mu < A + 1.5\tau$ , so the number of subsequent users,  $m$ , and hence the expected benefit resulting from the trial, is equal to zero. Consequently the expected net benefit is equal to zero.

It should be noted that if  $\mu < A$  but  $\tau$  is large it may well be worth conducting a trial.

Proceeding as in (4.8) we have the following lemma.

**Lemma:** The prior distribution function of  $m$ , the number of subsequent users of the new treatment, is

$$F_m(y) = \begin{cases} 0 & y < 0 \\ F_g\left(\frac{H'(y)-\mu}{\sqrt{\frac{a(1+nw)}{ng}}}\right) & 0 \leq y < M \\ 1 & y \geq M. \end{cases} \quad (5.12)$$

Hg degrees of freedom and  $H'(y)$  is the value of  $\bar{z}_n$  for which

$$\bar{z}_n = \frac{\left[\frac{y}{M}(B-A) + A + 1.5\sqrt{\frac{w_1 a_1}{g_1-2}}\right](1+nw) - \mu}{nw}.$$

Using the above lemma we can show that the prior density function of  $m$  is

$$f_m(y) = \begin{cases} P_r(m=y) = F_g\left(\frac{H_1-\mu}{\sqrt{\frac{a(1+nw)}{ng}}}\right) & y = 0 \\ f_g\left(\frac{H'(y)-\mu}{\sqrt{\frac{a(1+nw)}{ng}}}\right) \frac{B-A}{nwM}(1+nw) & 0 < y < M \\ P_r(m=y) = 1 - F_g\left(\frac{H_2-\mu}{\sqrt{\frac{a(1+nw)}{ng}}}\right) & y = M, \end{cases} \quad (5.13)$$

where  $f_g(\cdot)$  denotes the density function of  $t$ -Student with  $g$  degrees of freedom. Using this density function one can find the expected number of subsequent users of the new treatment.

## 6. The Regulator as a Third Decision-Maker

As in the known variance case, we assume that the regulator uses a different (joint) prior density for  $(\delta, \sigma^2)$  from the one assumed by the company conducting the trial. This may be written as

$$\pi_r(\delta, \sigma^2) = k(a_r, g_r, w_r)(\sigma^2)^{-\frac{g_r+3}{2}} e^{-\frac{1}{2\sigma^2}\{a_r + (\delta - \mu_r)^2/w_r\}},$$

where

$$k(a_r, g_r, w_r) = a_r^{\frac{g_r}{2}} 2^{-\frac{g_r+1}{2}} (\pi w_r)^{-1/2} \{\Gamma(1/2)\}^{-1}.$$

The mean and the variance of the marginal prior density of  $\delta$  will be denoted by  $\mu_r$  and  $\tau_r^2$ , while  $\mu'_r$  and  $\tau_r'^2$  denote, respectively, the posterior mean and

the posterior variance for  $\delta$ . We assume that the new treatment will be approved if and only if

$$\mu'_r \geq L + 1.5\tau'_r, \quad (5.14)$$

where, once again,  $L$  is the minimum amount of improvement required by the regulatory authorities to grant a licence to the new treatment. This inequality may be written as  $\bar{z}_n \geq H_r$ , where  $H_r$  is the value of  $\bar{z}_n$  for which

$$\bar{z}_n = \frac{(L + 1.5\sqrt{\frac{w_{1r}a_{1r}}{g_{1r}-2}})(1 + nw_r) - \mu_r}{nw_r}. \quad (5.15)$$

Using the above notation, the function representing the number of subsequent users of the new treatment may now be written as

$$m = \begin{cases} 0 & \mu'_c < A + 1.5\tau'_c \\ \frac{M}{B-A}[\mu'_c - A - 1.5\tau'_c] & A + 1.5\tau'_c < \mu'_c < B + 1.5\tau'_c \\ M & B + 1.5\tau'_c < \mu'_c. \end{cases}$$

Once again we discuss the registration problem, first for the case where the benefit function is written from the public health standpoint.

## 6.1. The Regulator and the Public Health Benefit Function

As before there are three different cases to be considered. Note that each case depends on  $n$  and on  $s^2$ .

### Case 1. $H_r \leq H_1 < H_2$

In this case the regulator's decision whether or not to grant a licence to the new treatment has no influence on the number of users.

Thus

$$\begin{aligned}
R(n) = & \frac{1}{(B-A)B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \times \\
& \int_0^\infty \int_{H_1}^{H_2} \left\{ \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} - A - 1.5 \sqrt{\frac{\frac{n}{1+nw_c} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})}{g_c + n - 2}} \right\} \\
& \times \left( \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right) (s^2)^{\frac{n-1}{2}-1} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 \\
& + \frac{1}{B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \int_0^\infty \int_{H_2}^\infty \left( \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right) \\
& \times (s^2)^{\frac{n-1}{2}-1} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 - Cn.
\end{aligned} \tag{5.16}$$

**Case 2.**  $H_1 < H_r < H_2$

Now the objective function can be written as

$$\begin{aligned}
R(n) = & \frac{1}{(B-A)B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \int_0^\infty \int_{H_r}^{H_2} \left\{ \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right. \\
& \left. - A - 1.5 \sqrt{\frac{\frac{n}{1+nw_c} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})}{g_c + n - 2}} \right\} \left( \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right) (s^2)^{\frac{n-1}{2}-1} \\
& \times (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 \\
& + \frac{1}{B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \int_0^\infty \int_{H_2}^\infty \left( \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right) (s^2)^{\frac{n-1}{2}-1} \\
& \times (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 - Cn.
\end{aligned} \tag{5.17}$$

**Case 3.**  $H_1 < H_2 \leq H_r$

Now we have

$$R(n) = \frac{1}{B(1/2, (g_c + n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1 + nw_c}} \int_0^\infty \int_{H_r}^\infty \left( \frac{\mu_c + nw_c \bar{z}_n}{1 + nw_c} \right) (s^2)^{\frac{n-1}{2} - 1} \\ \times \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1 + nw_c} \right)^{-\frac{g_c + n}{2}} d\bar{z}_n ds^2 - Cn. \quad (5.18)$$

## 6.2. The Regulator and the Commercial Benefit Function

There are three different cases to be considered

**Case 1.**  $H_r \leq H_1 < H_2$

The objective function can be written as

$$R(n) = \frac{1}{(B - A)B(1/2, (g_c + n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1 + nw_c}} \times \\ \int_0^\infty \int_{H_1}^{H_2} \left\{ \frac{\mu_c + nw_c \bar{z}_n}{1 + nw_c} - A - 1.5 \sqrt{\frac{\frac{n}{1 + nw_c} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1 + nw_c})}{g_c + n - 2}} \right\} \\ \times (s^2)^{\frac{n-1}{2} - 1} \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1 + nw_c} \right)^{-\frac{g_c + n}{2}} d\bar{z}_n ds^2 \\ + \frac{1}{B(1/2, (g_c + n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1 + nw_c}} \int_0^\infty \int_{H_2}^\infty (s^2)^{\frac{n-1}{2} - 1} \\ \times \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1 + nw_c} \right)^{-\frac{g_c + n}{2}} d\bar{z}_n ds^2 - Cn. \quad (5.19)$$

**Case 2.**  $H_1 < H_r < H_2$

Thus

$$\begin{aligned}
R(n) = & \frac{1}{(B-A)B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \int_0^\infty \int_{H_r}^{H_2} \left\{ \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} - A \right. \\
& - 1.5 \sqrt{\frac{\frac{n}{1+nw_c} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})}{g_c + n - 2}} \left. \right\} (s^2)^{\frac{n-1}{2}-1} \\
& \times (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 \\
& + \frac{1}{B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \int_0^\infty \int_{H_2}^\infty (s^2)^{\frac{n-1}{2}-1} \\
& \times (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 - Cn.
\end{aligned} \tag{5.20}$$

**Case 3.**  $H_1 < H_2 \leq H_r$

Now we have

$$\begin{aligned}
R(n) = & \frac{1}{B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw}} \int_0^\infty \int_{H_r}^\infty (s^2)^{\frac{n-1}{2}-1} \\
& \times (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 - Cn.
\end{aligned} \tag{5.21}$$

The above results may now be summarized in the following theorems.

**Theorem:** If the regulator's requirement is as in (5.14) and the benefit of using the new treatment is expressed from the public health standpoint, then the optimal solution to the sample size problem stated in section 2 is  $n^*$  which maximizes the following objective (expected net benefit) function

$$\begin{aligned}
R(n) = & \frac{1}{(B-A)B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \times \\
& \int_0^\infty \int_{\max(H_1, H_r)}^{\max(H_2, H_r)} \left\{ \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} - A - 1.5 \sqrt{\frac{\frac{n}{1+nw_c} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})}{g_c + n - 2}} \right\} \\
& \times \left( \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right) (s^2)^{\frac{n-1}{2}-1} \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c} \right)^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 \\
& + \frac{1}{B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \int_0^\infty \int_{\max(H_2, H_r)}^\infty \left( \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} \right) \\
& \times (s^2)^{\frac{n-1}{2}-1} \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c} \right)^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 - Cn.
\end{aligned} \tag{5.22}$$

**Theorem:** If the benefit of using the new treatment is written from stand-point of the pharmaceutical company conducting the trial and if the regulator's requirement is as in (5.14) then the optimal solution to the sample size problem stated in section 2 is  $n^*$  which maximizes the following objective (expected net benefit) function

$$\begin{aligned}
R(n) = & \frac{1}{(B-A)B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \times \\
& \int_0^\infty \int_{\max(H_1, H_r)}^{\max(H_2, H_r)} \left\{ \frac{\mu_c + nw_c \bar{z}_n}{1+nw_c} - A - 1.5 \sqrt{\frac{\frac{n}{1+nw_c} (a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c})}{g_c + n - 2}} \right\} \\
& \times (s^2)^{\frac{n-1}{2}-1} \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw_c} \right)^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 \\
& + \frac{1}{B(1/2, (g_c+n)/2)} \left\{ \frac{a_c^{g_c/2}}{\Gamma(\frac{n-1}{2})} \right\} \sqrt{\frac{n}{1+nw_c}} \int_0^\infty \int_{\max(H_2, H_r)}^\infty (s^2)^{\frac{n-1}{2}-1} \\
& \times \left( a_c + s^2 + \frac{n(\bar{z}_n - \mu_c)^2}{1+nw} \right)^{-\frac{g_c+n}{2}} d\bar{z}_n ds^2 - Cn.
\end{aligned} \tag{5.23}$$

Figure 5.5 illustrates the case of figure 5.3 for low ( $L = 1.5$ ) and intermediate ( $L = 2.25$ ) regulator requirements. With the low requirement the expected net benefit is higher than the one for the intermediate requirement. In this example we assumed that  $\mu_r = 0$ ,  $a_r = w_r = 1$ , and  $g_r = 3$ . The optimal sample size for the lower expected net benefit function is  $n^* = 36$ .

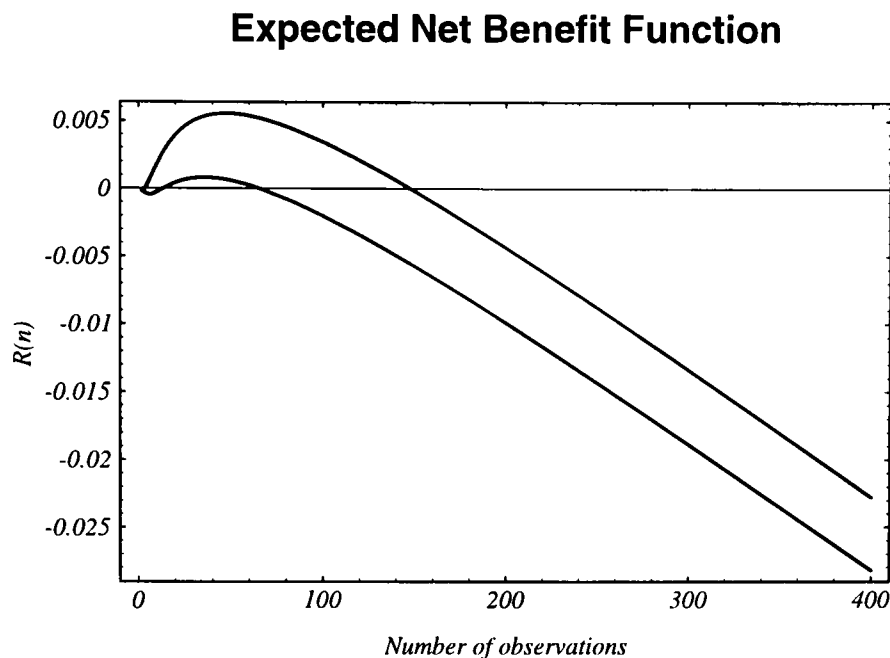


Figure 5.5: Expected net benefit with low and intermediate regulator requirements for the case of figure 5.3.

## 7. Convergence to the Known Variance Case

For a sequence of unknown variance cases for which the mean and variance of the prior distribution for  $\delta$  and the mean of the prior distribution for  $\sigma^2$  are held fixed, the variance of the prior distribution for  $\sigma^2$  is decreasing, and the prior distribution for  $\delta$  tends to normal, we would expect to find that  $n^*$  tends to the value for the known  $\sigma^2$  limit.

Recall that the prior density function of  $\delta$  is such that

$$\frac{\delta - \mu}{\sqrt{\frac{wa}{g}}} \sim t_g, \quad (5.24)$$

which converges in distribution to  $N(0, 1)$  as  $g \rightarrow \infty$ .

From (5.8) one can derive the (predictive) density function of  $\bar{Z}_n$ . This is

$$f(\bar{z}_n) = \frac{1}{B(1/2, g/2)} \left( \frac{a(1+nw)}{n} \right)^{-1/2} \left[ 1 + \frac{(\bar{z}_n - \mu)^2}{\frac{a(1+nw)}{n}} \right]^{-\frac{g+1}{2}}. \quad (5.25)$$



Putting  $t = \frac{\bar{Z}_n - \mu}{\sqrt{\frac{a(1+nw)}{ng}}}$  we have

$$f(t) = \frac{1}{B(1/2, g/2)} \frac{1}{\sqrt{g}} \left[1 + \frac{t^2}{g}\right]^{-\frac{g+1}{2}}, \quad (5.26)$$

so  $\frac{\bar{Z}_n - \mu}{\sqrt{\frac{a(1+nw)}{ng}}} \sim t_g$ .

$\sigma^2$  has a prior density which is an inverse chi-squared with the mean and the variance

$$E(\sigma^2) = \frac{a}{(g-2)}, \quad V(\sigma^2) = \frac{2a^2}{(g-2)^2(g-4)}. \quad (5.27)$$

With a tight prior distribution for  $\sigma^2$  we are close to the situation where  $\sigma^2$  is known to be equal to the mean of this distribution. Thus the optimal sample size calculated with a tight prior distribution for  $\sigma^2$  should result in a value close to the one for the known variance case. We can use this fact to check the consistency of our sample size calculations for the case where the variance of the observations is unknown.

Now let  $a$  and  $g$  tend to infinity, subject to  $\frac{a}{g-2} = \rho^2$  and  $\frac{wa}{g-2} = \tau^2$ , where  $\rho$  and  $\tau$  are fixed.

Thus

$$\frac{\bar{Z}_n - \mu}{\sqrt{\frac{a(1+nw)}{ng}}} \xrightarrow{D} N(0, 1),$$

$$\bar{Z}_n \xrightarrow{D} N(\mu, \tau^2 + \frac{\rho^2}{n}), \quad (5.28)$$

and

$$\sigma^2 \xrightarrow{p} \rho^2.$$

In other words the predictive distribution of  $\bar{Z}_n$  in the unknown variance case tends to the one for the known variance case.

Since the prior distribution of  $\delta$  and the predictive distribution of  $\bar{Z}_n$  in the unknown variance case tend to those for the known variance case, it is

not difficult to show that  $m\mu'$  converges in distribution to its distribution for the known variance case as the objective or expected net benefit function tends to the one for known variance, and it follows that the optimal sample size tends to the optimal sample for the known  $\sigma^2$  limit. We shall prove this through the following lemmas and theorems. We use the subscript  $u$  to indicate that the parameter belongs to the case of unknown variance and the subscript  $k$  to indicate the known variance case.

**Lemma:** The posterior mean of  $\delta$  for unknown variance converges in distribution to its value for known variance. In symbols

$$\frac{\mu + nw\bar{Z}_u}{1 + nw} \xrightarrow{D} \frac{\rho^2\mu + n\tau^2\bar{Z}_k}{\rho^2 + n\tau^2}. \quad (5.29)$$

**Proof:** The observation (5.28) may be written as  $\bar{Z}_u \xrightarrow{D} \bar{Z}_k$ . It follows that

$$\frac{\frac{a}{g-2} \mu + n \frac{wa}{g-2} \bar{Z}_u}{\frac{a}{g-2} + n \frac{wa}{g-2}} \xrightarrow{D} \frac{\rho^2\mu + n\tau^2\bar{Z}_k}{\rho^2 + n\tau^2}.$$

Thus

$$\frac{\mu + nw\bar{Z}_u}{1 + nw} \xrightarrow{D} \frac{\rho^2\mu + n\tau^2\bar{Z}_k}{\rho^2 + n\tau^2}, \quad (5.30)$$

and (5.29) follows. □

**Lemma:** The posterior variance of  $\delta$  for the unknown variance case converges in probability to its value for known variance. In symbols

$$\frac{\frac{w}{1+nw} \left( a + s^2 + \frac{n(\bar{Z}_u - \mu)^2}{1+nw} \right)}{n + g - 2} \xrightarrow{P} \frac{\rho^2\tau^2}{\rho^2 + n\tau^2} \quad (5.31)$$

**Proof:** First note that as  $g$  tends to infinity  $\bar{Z}_u \xrightarrow{P} \mu$ . So  $(\bar{Z}_u - \mu)^2 \xrightarrow{P} 0$ . Also note that  $\frac{s^2}{n-1} \xrightarrow{P} \rho^2$  as  $g \rightarrow \infty$ . So

$$a + s^2 + \frac{n(\bar{Z}_u - \mu)^2}{1 + nw} \xrightarrow{P} a + (n-1)\rho^2. \quad (5.32)$$

Since  $\frac{a}{g-2} = \rho^2$  and  $\frac{wa}{g-2} = \tau^2$  it follows that

$$\frac{w(a + (n-1)\rho^2)}{(1 + nw)(n + g - 2)} \rightarrow \frac{wa/(g-2)}{1 + nw} \approx \frac{\rho^2\tau^2}{\rho^2 + n\tau^2}, \quad (5.33)$$

as  $g \rightarrow \infty$ . The lemma follows from (5.32) and (5.33).

Now we are in a position to state following theorems.

**Theorem:** If  $a$  and  $g$  tend to infinity, subject to  $\frac{a}{g-2} = \rho^2$  and  $\frac{wa}{g-2} = \tau^2$ , where  $\rho^2$  and  $\tau$  are fixed, then  $E[m(\mu'_u)] \rightarrow E[m(\mu'_k)]$ .

**Proof:** Since  $m$  is a bounded continuous function of both  $\mu'_u$  and  $\tau'_u$ , and since  $\mu'_u \xrightarrow{D} \mu'_k$  and  $\tau'_u \xrightarrow{P} \tau'_k$  as  $a$  and  $g$  tend to infinity, the theorem holds (see, for example, theorem 2.2 of Durrett (1996), p.86). □

This theorem implies that the reward function of the unknown variance case with the commercial benefit function converges to the corresponding reward function of the known variance case:

$$r_u(n) \rightarrow r_k(n). \quad (5.34)$$

**Theorem:** Under the conditions of the previous theorem

$$E[m(\mu'_u)\mu'_u] \rightarrow E[m(\mu'_k)\mu'_k].$$

**Proof:** It follows from the continuous mapping theorem that since  $\mu'_u \xrightarrow{D} \mu'_k$

and  $m(x)x$  is a continuous function of  $x$  then  $m(\mu'_u)\mu'_u \xrightarrow{D} m(\mu'_k)\mu'_k$ . Now note that  $E(\mu'_u) = \mu = E(\mu'_k) < \infty$  and  $m(\cdot)$  is bounded. The theorem follows. □

The above theorem implies that the reward function of the unknown variance case with the public health benefit function converges to the corresponding reward function for the known variance case.

Before closing this section let us consider the following example which demonstrates the convergence of the optimal sample size of a sequence of the unknown variance cases to the one of the known  $\sigma^2$  limit.

### 7.1. An Example

Let us consider known variance with the following parameters for the commercial version of the objective function.

$$\sigma^2 = 4, \quad \tau = 1.045, \quad \mu = 2.09, \quad \frac{c}{Mb} = \frac{4000}{5000000}, \quad (A, B) = (1.67, 2.51).$$

It can be shown that the optimal sample size is  $n^* = 75.77125$  and that  $r(n^*) = 1.56795 \times 10^6$ .

A sequence of sets of parameters for unknown variance which converge to this case are shown below.

1)

$$g = 5, \quad a = 12, \quad w = 0.435, \quad \mu = 2.09, \quad (A, B) = (1.67, 2.51);$$

$$\boxed{n^* = 63.6099, \quad r(n^*) = 1.76210 \times 10^6,}$$

2)

$$g = 12, \quad a = 40, \quad w = 0.3135, \quad \mu = 2.09, \quad (A, B) = (1.67, 2.51);$$

$$\boxed{n^* = 71.5817, \quad r(n^*) = 1.68758 \times 10^6,}$$

3)

$$g = 32, \quad a = 120, \quad w = 0.2787, \quad \mu = 2.09, \quad (A, B) = (1.67, 2.51);$$

$$n^* = 74.6111, \quad r(n^*) = 1.66123 \times 10^6.$$

Here  $n^*$  tends to the known variance value 75.77125 as  $a$  and  $g$  tend to infinity.

□

# Chapter 6

## A Behavioral Bayes Approach to the Sample Size Question; Binomial Distribution

### 1. Introduction

Using the central limit theorem to justify the assumption of normality for the case with binary responses, in chapter four we discussed how we may apply our BeBay methodology to binomially distributed data. In our examples for binomial data, using the notation of chapter four,  $\bar{Z}_n$  was re-defined as  $\log[\frac{1}{n} \sum X_i / (1 - \frac{1}{n} \sum X_i)] - \log[\frac{1}{n} \sum Y_i / (1 - \frac{1}{n} \sum Y_i)]$  and  $\delta$  as  $\log\{p_N(1 - p_C)/(p_C(1 - p_N))\}$  where  $p_N = P_r(\text{cure} \mid \text{new treatment})$  and  $p_C = P_r(\text{cure} \mid \text{control or alternative treatment})$ . Thus  $p_N$  and  $p_C$  are the respective probabilities of success on the new and control (alternative or existing) treatments. Now  $X_i = 1$  if the  $i$ 'th patient responds favourably to the new treatment, and otherwise  $X_i = 0$ , and  $Y_i$  is defined similarly for the  $i$ 'th patient receiving the alternative treatment. Inference was based on the assumption that  $\delta$  has a normal prior density, using the fact that  $\bar{Z}_n$  has a distribution which is approximately normal with mean  $\delta$  and variance  $\frac{2(p_C(1-p_C)+p_N(1-p_N))}{np_C(1-p_C)p_N(1-p_N)}$  (see Coad and Rosenberger (1999)). For rare events, Spiegelhalter *et al* (1994) suggest an approximation of  $4/n$  for the variance of  $\bar{Z}_n$ .

This chapter extends the methodology to binomial data for which the central limit theorem does not provide an adequate approximation. As usual 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. Throughout this chapter we assume that there is just one group of patients in the trial and they receive the new treatment. In due course it is hoped to extend the BeBay methodology to trials for which there are control groups.

In the second 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 for both the public health and commercial benefit functions. The prior density function of  $m$  is obtained in section 5. The regulator's requirements for granting a licence to the new treatment are discussed in section 6.

## 2. Notation

In a clinical trial of size  $n$ , suppose  $X$  is the number of favourable outcomes for patients using the new treatment. If we assume that the patients (observations) are independent then  $X$  is binomially distributed with parameters  $n$  and  $p$  where  $p$  is the probability of favourable outcome for each individual.

With a binomial likelihood it is mathematically convenient, and often reasonably realistic, to make the assumption that  $p$  has a prior distribution  $\pi$  which is beta with parameters  $\alpha$  and  $\beta$ . Thus

$$\pi(p) = \frac{1}{B(\alpha, \beta)} p^{\alpha-1} (1-p)^{\beta-1} \quad \alpha, \beta > 0, \quad 0 \leq p \leq 1, \quad (6.1)$$

( $B(\alpha, \beta)$  is the beta function with parameters  $\alpha$  and  $\beta$ ). We may write this as  $p \sim \text{Beta}(\alpha, \beta)$ . If  $X$  takes the value  $x$ , then the posterior distribution of  $p$  is

$$\pi^n(p | x) = \frac{1}{B(\alpha + x, \beta + n - x)} p^{\alpha+x-1} (1-p)^{\beta+n-x-1}, \quad 0 \leq p \leq 1. \quad (6.2)$$

Recall that in binomial sampling, we assume that  $n$  is fixed and  $X$  is a random variable with  $x = 0, 1, 2, \dots, n$ .

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

$$\mu = \frac{\alpha}{\alpha + \beta}, \quad \tau^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)},$$

and

$$\mu' = \frac{\alpha + x}{\alpha + \beta + n}, \quad \tau'^2 = \frac{(\alpha + x)(\beta + n - x)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)}.$$

Note that the predictive distribution of  $X$  is a beta-binomial distribution with parameters  $\alpha$ ,  $\beta$  and  $n$  and density function

$$f_X(x) = \binom{n}{x} \frac{B(\alpha + x, \beta + n - x)}{B(\alpha, \beta)}, \quad x = 0, 1, 2, \dots, n. \quad (6.3)$$

Its cumulative distribution function will be denoted by  $F_X(x)$ . Before taking a sample of fixed size  $n$ , preposterior analysis studies the possible outcomes concerning the posterior distribution. Since  $X$  is a random variable, it follows that  $\mu'$  and  $\tau'^2$  are random variables because of their dependence on  $X$ .

Using (6.3), we may obtain the means and the variances of  $\mu'$  and  $\tau'^2$

$$E(\mu') = \frac{\alpha}{\alpha + \beta} = \mu, \quad V(\mu') = \frac{n\tau^2}{\alpha + \beta + n},$$

and

$$E\tau'^2 = \frac{\alpha\beta}{(\alpha + \beta)(\alpha + \beta + 1)(\alpha + \beta + n)},$$

$$\begin{aligned} V(\tau'^2) = & \{E(X^4) + 2(\alpha - n - \beta)E(X^3) + [\alpha^2 + (n + \beta)(\beta + n - 4\alpha)]E(X^2) \\ & + 2\alpha(\beta + n)(\beta + n - \alpha)E(X) + \alpha^2(\beta + n)^2\} \\ & \times [(\alpha + \beta + n)^4(\alpha + \beta + n + 1)^2]^{-1} - [E(\tau'^2)]^2, \end{aligned}$$

where  $E(X^j)$ ,  $1 \leq j \leq 4$ , is the  $j$ th moment of a random variable  $X$  having a beta-binomial distribution with parameters  $\alpha$ ,  $\beta$ , and  $n$ .

Using these results we may calculate the sample sizes required to satisfy precision conditions on criteria related to the posterior distribution (see, for example Pham-Gia and Turkkan (1992)). Here we follow our usual procedure and obtain the optimal sample size by maximizing the expected net benefit function. This is the expected benefit from conducting the trial minus its cost.



### 3. Number of Subsequent Users of the New Treatment

Following our approach in normal sampling we assume that the number of subsequent users of the new treatment is a function of the posterior mean and variance of  $p$ . As shown by the following graph the number of subsequent users of the new treatment is zero unless the posterior mean is greater than  $A + 1.5\tau'$ .

$$m = \begin{cases} 0 & \mu' < A + 1.5\tau' \\ \frac{M}{B-A}[\mu' - A - 1.5\tau'] & A + 1.5\tau' < \mu' < B + 1.5\tau' \\ M & B + 1.5\tau' < \mu', \end{cases} \quad (6.4)$$

in graphical terms

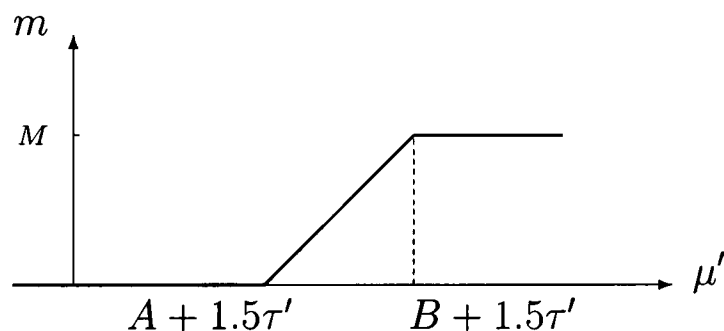


Figure 6.1: Number of subsequent users.

## 4. Utility Functions

The benefit per patient using the new treatment is assumed by the BeBay programme to be  $b_1p + b_2$  ( $b_1 \geq 0$ ). As before, our discussion focuses on the two extreme cases  $b_1 = 0$  and  $b_2 = 0$ . The case  $b_2 = 0$ , which might correspond to a public health benefit, is discussed first.

### 4.1. Public Health Benefit Function

To simplify the notation let us suppress the subscript and assume that the benefit per patient of using the new treatment is  $bp$ , where  $b (> 0)$  is a known constant.

The objective function is

$$r(n) = \sum_x \int_p m \, b p \, \pi^n(p | x) dp \, f_X(x) - cn,$$

where  $cn$  is the cost of a trial of size  $n$ . This may be written as

$$r(n) = \sum_x mb \frac{\alpha + x}{\alpha + \beta + n} \binom{n}{x} \frac{B(\alpha + x, \beta + n - x)}{B(\alpha, \beta)} - cn. \quad (6.5)$$

It is convenient to write (6.4) as a function of  $x$

$$m = \begin{cases} 0 & x < [H(A)] \\ \frac{M}{B-A} [\mu' - A - 1.5\tau'] & [H(A)] \leq x < [H(B)] \\ M & [H(B)] \leq x, \end{cases}$$

where  $[s]$  is the greatest integer less than or equal to  $s$ , and

$$H(i) = \frac{-b_1 + \sqrt{b_1^2 - 4a_1c_1}}{2a_1}, \quad i = A, B$$

with

$$\begin{aligned} a_1 &= \alpha + \beta + n + 3.25, \\ b_1 &= 2.25(\beta + n - \alpha) - 2(\alpha - i(\alpha + \beta + n))(\alpha + \beta + n + 1), \\ c_1 &= (\alpha - i(\alpha + \beta + n))^2(\alpha + \beta + n + 1) - 2.25\alpha(\beta + n). \end{aligned}$$

So (6.5) may now be written as

$$\begin{aligned} r(n) &= 0 + \sum_{x=[H(A)]}^{[H(B)]} \frac{Mb}{B-A} \left[ \frac{\alpha + x}{\alpha + \beta + n} - A - 1.5\tau' \right] \left( \frac{\alpha + x}{\alpha + \beta + n} \right) f_X(x) \\ &\quad + \sum_{x=[H(B)]}^n Mb \left( \frac{\alpha + x}{\alpha + \beta + n} \right) f_X(x) - cn, \end{aligned}$$

or finally

$$\begin{aligned}
 R(n) = & \sum_{x=[H(A)]}^{[H(B)]} \frac{1}{B-A} \left[ \frac{\alpha+x}{\alpha+\beta+n} - A - 1.5\tau' \right] \left( \frac{\alpha+x}{\alpha+\beta+n} \right) f_X(x) \\
 & + \sum_{x=[H(B)]}^n \left( \frac{\alpha+x}{\alpha+\beta+n} \right) f_X(x) - Cn,
 \end{aligned} \tag{6.6}$$

where  $f_X(x)$  is defined in (6.3),  $R(n) = \frac{r(n)}{Mb}$ , and  $C = \frac{c}{Mb}$ .

## 4.2. Commercial Benefit Function

A similar analysis to the above, now with a constant benefit  $b_2 = b$  per patients, leads us to the following objective (or expected net benefit) function

$$\begin{aligned}
 R(n) = & \sum_{x=[H(A)]}^{[H(B)]} \frac{1}{B-A} \left[ \frac{\alpha+x}{\alpha+\beta+n} - A - 1.5\tau' \right] f_X(x) \\
 & + \sum_{x=[H(B)]}^n f_X(x) - Cn,
 \end{aligned} \tag{6.7}$$

where  $R(n)$ ,  $H(A)$ ,  $H(B)$  and  $C$  are as in (6.6), although the meaning of  $b$  is now different.

□

**Remark:** If  $p \sim B(\alpha, \beta)$  with the mean  $E(p) = \mu$  and the variance  $V(p) = \tau^2$  then

$$\alpha = \left( \frac{\mu}{\tau} \right)^2 (1 - \mu) - \mu, \quad \beta = \left( \frac{1 - \mu}{\tau} \right)^2 \mu + (\mu - 1). \tag{6.8}$$

## 4.3. Example (Trial 2 of chapter 1)

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 population 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 recently carried out by a 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 proportion of successes is as 47%. Consequently the mean  $\mu$  of the prior distribution for  $p$  was assumed to be 0.47. The prior standard deviation,  $\tau$ , for  $p$  was assumed to be  $1/2\min(\mu, 1 - \mu)$ . It was also assumed, and this seemed realistic, that  $A = 0.8\mu$  and  $B = 1.2\mu$ . 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 provided the trial was successful. The total cost  $c$  per patient in the trial including hospital expenses and administration costs, is £4000.

Running BeBay with the following parameters gives the optimal sample size.

$\mu = 0.47$ . The company view of the likely improvement in condition.

$\tau = 1/2\min(\mu, 1 - \mu) = 0.235$ . A convenient representation of prior ignorance.

(From equations (6.8) it follows that  $\alpha = 1.65$  and  $\beta = 1.85$ )

$A = 0.8\mu = 0.376$  = maximum critical difference for no sale.

$B = 1.2\mu = 0.564$  = minimum critical difference for maximum sale.

$Mb = £5,000,000$  = the expected net benefit if the trial is clearly successful.

The function  $r(n)$ , as given by equation (6.7), was maximized and the optimal (i.e. maximizing) sample size  $n^*$  for this trial was found to be 52, resulting in an expected net benefit  $r(n^*)$  of £1.67815M.

In fact the company concerned based the sample size on calculations of the power and size of the resulting significance test leading, control rules of classical approach to the inclusion of 40 patients in the trial. We shall see later

that, as in the normal distribution case, the important issues influencing the optimal sample size are the expected benefit from a conclusively favourable outcome and the strength of the evidence required by the regulator.

□

## 5. Distribution of the Number of Subsequent Users of the New Treatment

**Lemma:** The prior distribution of the number  $m$  of subsequent users of the new treatment, is

$$F_m(y) = \begin{cases} 0 & y < 0 \\ \sum_{x=0}^{[H'(y)]} \binom{n}{x} \frac{B(\alpha+x, \beta+n-x)}{B(\alpha, \beta)} & 0 \leq y < M \\ 1 & M \leq y, \end{cases} \quad (6.9)$$

where

$$\begin{aligned} H'(y) = & 1/(2(\alpha + \beta + n + 3.25))(2.25(\beta + n - \alpha) - 2(\alpha + \beta + n + 1) \\ & \times (\frac{y}{M}(B - A) + A)(\alpha + \beta + n) + 2\alpha(\alpha + \beta + n + 1) \\ & + \left\{ ((2.25(\beta + n - \alpha) - 2(\alpha + \beta + n + 1)(\frac{y}{M}(B - A) + A) \right. \\ & \times (\alpha + \beta + n) + 2\alpha(\alpha + \beta + n + 1))^2 - 4(\alpha + \beta + n + 3.25) \\ & \times \left( \left( \alpha^2 - 2\alpha(\frac{y}{M}(B - A) + A) + (\frac{y}{M}(B - A) + A)^2 \right. \right. \\ & \left. \left. \times (\alpha + \beta + n)^2 \right) (\alpha + \beta + n + 1) - 2.25\alpha(\beta + n) \right) \left. \right\}^{1/2}. \end{aligned} \quad (6.10)$$

**Proof:** If  $0 \leq y < M$  then as shown by figure 6.2 we may write

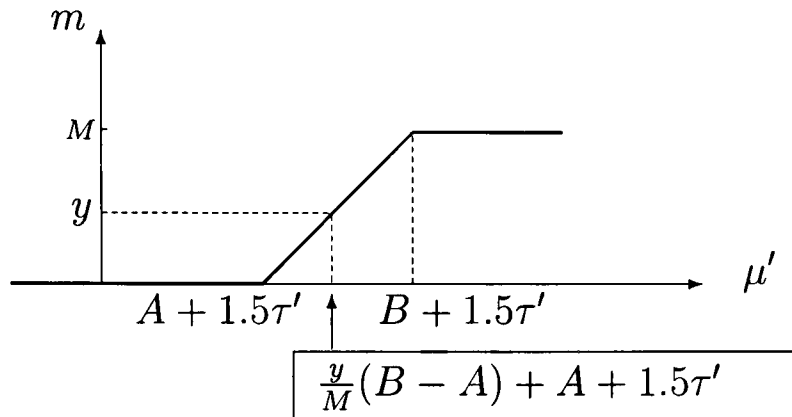


Figure 6.2: Number of subsequent users.

$$F_m(y) = P_r(m \leq y)$$

$$\begin{aligned}
 &= P_r \left\{ \mu' \leq \frac{y}{M}(B - A) + A + 1.5\tau' \right\} \\
 &= P_r \left\{ \frac{\alpha + x}{\alpha + \beta + n} \leq \frac{y}{M}(B - A) + A + \right. \\
 &\quad \left. 1.5 \left[ \frac{(\alpha + x)(\beta + n - x)}{(\alpha + \beta + n)^2(\alpha + \beta + n + 1)} \right]^{1/2} \right\}.
 \end{aligned}$$

Using the fact that, if  $0 < a < b$  then  $a^2 < b^2$ , the above expression can be written as

$$\begin{aligned}
 F_m(y) = P_r \bigg\{ &2.25(\alpha + x)(\beta + n - x) - (\alpha + \beta + n + 1) \\
 &[\alpha + x - (\frac{y}{M}(B - A) + A)(\alpha + \beta + n)]^2 \geq 0 \bigg\}.
 \end{aligned}$$

Solving this quadratic expression for  $x$  leads us to (6.9).

□

It should be noted that

$$\begin{aligned}
 P_r(m = 0) &= P_r\{\mu' < A + 1.5\tau'\} \\
 &= F_X([H(A)]),
 \end{aligned}$$

and

$$\begin{aligned} P_r(m = M) &= P_r\{\mu' \geq B + 1.5\tau'\} \\ &= 1 - F_X([H(B)]). \end{aligned}$$

Using these equations together with the above lemma, the prior density function of  $m$  at any point of its support is given by the equation

$$f_m(y) = F_m([H'(y)]) - F_m([H'(y^-)]), \quad y = 1, 2, \dots, M - 1, \quad (6.11)$$

where  $F_m([H'(y^-)])$  is  $\lim_{z \uparrow y} F_m([H'(z)])$ .

## 6. The Regulator as a third decision-Maker

We assume that the regulator also uses a beta prior density for  $p$ , but with different parameters from the one used by the company conducting the trial. Parameters for prior and posterior distributions for the company and for the regulator will be distinguished by the subscripts  $c$  and  $r$ .

We assume that the new treatment will be approved if and only if

$$\mu'_r \geq L + 1.5\tau'_r, \quad (6.12)$$

where  $L$  is the minimum probability of success  $p$  required by the regulator to grant a licence to the new treatment. This inequality may be written as  $x \geq H_r(L)$ , where  $H_r(L)$  is the value of  $x$  for which

$$x = (L + 1.5\tau'_r)(\alpha_r + \beta_r + n) - \alpha_r. \quad (6.13)$$

Taking account of the regulator in this way the function representing the number of subsequent users of the new treatment becomes

$$m = \begin{cases} 0 & \mu'_c < A + 1.5\tau'_c \\ \frac{M}{B-A}[\mu'_c - A - 1.5\tau'_c] & A + 1.5\tau'_c < \mu'_c < B + 1.5\tau'_c \\ M & B + 1.5\tau'_c < \mu'_c. \end{cases}$$

We shall discuss the registration of the new treatment first for the public health benefit function.

## 6.1. The Regulator and the Public Health Benefit Function

As discussed in chapters 4 and 5 there are three different cases to be considered. These may be summarized in the following theorem.

**Theorem:** If the regulator's requirement is as in (6.12) and the benefit of using the new treatment is expressed from the public health standpoint, then the optimal solution  $n^*$  to the sample size problem stated in section 2 is the value of  $n$  which maximizes the following objective (expected net benefit) function

$$R(n) = \sum_{x=\max([H(A)], [H_r(L)])}^{\max([H(B)], [H_r(L)])} \frac{1}{B-A} \left[ \frac{\alpha_c + x}{\alpha_c + \beta_c + n} - A - 1.5\tau'_c \right] \left( \frac{\alpha_c + x}{\alpha_c + \beta_c + n} \right) f_X(x) \\ + \sum_{x=\max([H(B)], [H_r(L)])}^n \left( \frac{\alpha_c + x}{\alpha_c + \beta_c + n} \right) f_X(x) - Cn, \quad (6.14)$$

where  $f_X(x)$  is as in (6.3),  $R(n) = \frac{r(n)}{Mb}$ ,  $C = \frac{c}{Mb}$ , and  $H(i)$  ( $i = A, B$ ) are as in (6.6), and  $H_r(L)$  is defined in (6.13).

## 6.2. The Regulator and the Commercial Benefit Function

Proceeding as before we can state the following theorem.



**Theorem:** If the regulator's requirement is as in (6.12) and the benefit of using the new treatment is expressed from the standpoint of the drug company conducting the trial, then the optimal solution to the sample size problem stated in section 2 is  $n^*$  which maximizes the following objective (expected net benefit) function

$$R(n) = \sum_{x=\max([H(A)], [H_r(L)])}^{\max([H(B)], [H_r(L)])} \frac{1}{B-A} \left[ \frac{\alpha_c + x}{\alpha_c + \beta_c + n} - A - 1.5\tau'_c \right] f_X(x) + \sum_{x=\max([H(B)], [H_r(L)])}^n f_X(x) - Cn, \quad (6.15)$$

where the notation is as in (6.14), though the meaning of  $b$  is now different (see section 4).

### 6.3. Example (see section 4.3.)

We assume that the regulator is not as optimistic as the company conducting the trial, and suppose he estimates the performance of the new treatment as  $\mu_r = p_C = 0.2$ , i.e. no improvement on the control (placebo) treatment. Since the regulator has less information than the company, let us assume  $\tau_r = \min(\mu_r, 1 - \mu_r) = 0.2$ . So his prior belief about  $p$ , the probability of success of the new treatment, may be summarized as  $p \sim \text{Beta}(0.6, 2.4)$ . If we assume that the regulator is prepared to grant a licence to the cream for the public use if the minimum probability of success is greater than or equal  $L = \frac{A+B}{2} = 0.47$ , then the optimal BeBay sample size  $n^*$  turns out to be 43 resulting in a net expected benefit of £1.55694M.

The following table shows BeBay optimal sample sizes and maximized expected benefit.

It is noticeable that when the regulator requires a milder condition for granting a licence to the new treatment the expected net benefit is increased.

**Trial 2**  
Actual Number of Patients, 20

		<i>Low Regulator Requirement</i>	<i>High Regulator Requirement</i>
<i>Low Maximum Benefit</i>	Optimal (BeBay) No. of Patients Binomial Calculation	52	43
	Maximized expected net benefit	£1.68M	£1.56M

# Chapter 7

## BeBay; A User–Guide

In this chapter we introduce the commands which might be used to run the computer-based planning tool, BeBay, for producing the optimal size of a trial.

BeBay is written in *Mathematica* which is a software platform on which the user can run packages built for specific applications and also it can be used as a high-level programming language which allows the user to create programs. It is therefore necessary for the user of BeBay to have access to *Mathematica*.

*Mathematica* is a general software system and language intended for mathematical and other applications. It can be run on a wide variety of computer systems such as Apple Macintosh, CONVEX, DEC, VAX (Ultrix and VMS) and RISC, Hewlett–Packard, Appollo, IBM 386(onwards)–based PC compatibles (MS–Dos and Microsoft Windows) and IBM RISC, MIPS, NeXT, Silicon Graphics, Sony and Sun (and SPARC compatibles) (see, for example, <http://www.wolfram.com> for more detailed information about the software).

### 1. Introduction

BeBay is a computer program for calculating the optimal size of a clinical trial using our general model described in previous chapters. It helps the drug company, or other organization conducting the trial, to estimate how many patients should be recruited in the trial. It is applicable to trials for drugs applied to new classes of patients, and to phase III clinical trials. In all cases the question at issue is the number of patients “*in total*” who should

be included. In general these will be at more than one centre, and will constitute more than one trial.

The objective (or the net expected benefit) function for the data from a continuous population may be written as a multiple integral of the form shown below

$$\int m(b_1\delta + b_2)f(\gamma(\mathbf{z}), \Theta)d\mathbf{z} - cn,$$

in which  $m$  is the number of subsequent users of the new treatment under consideration,  $\delta$  is the average amount of improvement of a treated patient,  $b_1\delta + b_2$  is the benefit for each patient using the new treatment,  $\mathbf{z}=(z_1, z_2, \dots, z_n)$  are data,  $\gamma(\mathbf{z})$  is the estimate of  $\delta$  based on the data,  $\Theta$  is the vector of unknown parameters (including  $\delta$ ), and  $cn$  is the cost of a trial with  $n$  patients. For discrete data the multiple integral becomes a multiple summation.

Thus the program calculates the expected net benefit of conducting a trial as the expected benefit resulting from it minus the cost of the trial. It consists of a set of routines which may be used for the calculations described in chapters 4, 5, and 6.

BeBay has two distinct stages of operation:

- 1) Data input,
- 2) Results.

Section 2 is a guide for using BeBay with normally distributed data. In section 3 the necessary routines and commands for BeBay sample size determination for binomially distributed data are discussed. Section 4 is a brief description about truncation error in numerical integration. Section 5 is a discussion about the numerical precision. A summary of the filenames is given in section 6.

## 2. Normally Distributed Data

*Mathematica* allows a user to work with either a text-based interface or a notebook interface. With a text-based interface, which we have used for our programming, when *Mathematica* is ready for input, it prints a prompt of

the form `In[n] :=`. When the user has entered the input, *Mathematica* processes it, and, if appropriate, prints out the result, with a label of the form `Out[n] =`. In a notebook interface, the input and output appear as elements in the notebook document, potentially mixed with text and graphics. Once again, the prompt for the input is of the form `In[n] :=` and the label for the output is `Out[n] =`.

With a text-based interface, the user can interact with the computer primarily by typing text on the keyboard. This kind of interface is available for *Mathematica* on many computer system, for instance those running the Unix or VMS operating system. On such systems *Mathematica* is called by typing the command `math` at an operating system prompt. When *Mathematica* has started, it will print the prompt `In[1] :=`, signifying that it is ready for the input. Ending a line (usually by pressing `Return` or `Enter`) tells *Mathematica* that you have finished giving your input. To exit *Mathematica*, you may either type `CONTROL-D` or `Quit[ ]` at an input prompt.

With a notebook interface, the user can start a session by double-click on the *Mathematica* icon. When *Mathematica* starts up, it usually gives a blank notebook. You can enter the input into the notebook, then press `SHIFT-RETURN` to make *Mathematica* process the input. After you send your input from the notebook, *Mathematica* will label the input with `In[n] :=`. It labels the corresponding output `Out[n] =`. You can use the `Quit` menu item in the notebook interface to exit from *Mathematica*.

There are nine files containing the routines needed for running BeBay. These are listed in section 6. Before starting a *Mathematica* session, these files must be loaded on a working directory (of user's choice).

We are now ready to discuss the data input process. Let us start with the known variance case (note that this includes the variants discussed in section 5 of chapter 4).

The next section gives a brief introduction to the meaning of the variables which must be entered as the *input* parameters along with some explanations about the BeBay *input* procedure for normally distributed data with known variance. Section 2.2. explains how the user may obtain the *output* associated with the *input*. There is a detailed example in section 2.3., based on a recent clinical trial, illustrating sections 2.1. and 2.2. We invite the user to work through the illustration sessions at the end of each section and see what happens.

## 2.1. Data Input; Known Variance

(see chapter 4 section 7)

The user of BeBay is asked to supply data outlining the characteristics of the trial under consideration. These may be summarized as

$A$  ,  $B$  ,  $c$  ,  $mc$  ,  $mr$  ,  $tc$  ,  $tr$  ,  $s$  ,  $L$  ,  $Mb1$  ,  $Mb2$  .

The parameters have the following meanings.

- A: The amount of apparent improvement of the new treatment above which some users switch to the new treatment.
- B: The amount of apparent improvement of the new treatment above which all users switch to the new treatment.
- c: Cost per patient in trial
- mc: Company's prior mean for  $\delta$  the average amount of improvement with the new treatment
- mr: Regulator's prior mean for  $\delta$  the average amount of improvement with the new treatment
- tc: Company's prior standard deviation for the average amount of improvement with the new treatment
- tr: Regulator's prior standard deviation for the average amount of improvement with the new treatment
- s: standard deviation of distribution of patients' responses
- L: The minimum amount of apparent improvement required by the regulator to grant a licence for the new treatment
- Mb1: First parameter of the total benefit function
- Mb2: Second parameter of the total benefit function

Note that  $\delta Mb1 + Mb2$  is the total benefit if all the potential users switch to the new treatment. If one of the parameters Mb1 and Mb2 is zero we get one of the two reduced forms of the benefit function discussed in sections 7.1. and 7.2. of chapter 4.

A routine which calculates the objective (or the expected net benefit) function, for which the benefit per patient of using the new treatment is  $b_1\delta + b_2$  (where  $b_1 = Mb_1/M$  and  $b_2 = Mb_2/M$  and  $M$  is as defined in section 2 of chapter 4), has been written and saved on the file `BeBayNK.all`. The user has to provide the above eleven data and then use the following procedure to run BeBay on `BeBayNK.all` and calculate the optimal size of the trial,  $n^*$ , which maximizes the expected net benefit from it.

## 2.2. Results; Known Variance

When the data entry process has been completed the user needs to run BeBay for the known variance data by using the command

```
In[12] := <<BeBayNK.all
```

This reads in the input file `BeBayNK.all`, evaluates the objective function in the known variance case for the eleven parameter values entered above, and writes the expected net benefit as a function of  $n$ .

The following command plots the objective function as a function of  $n$ , the number of patients in the trial. It should be noted that  $n$  varies between 0 and  $N$ , which must be specified.

```
In[13] := gt1 = Plot[%, {n, 0, N}]
```

Here `%` refers to the the previous output (for this case output of the `In[12]`). Note that if one wishes to label the axes he can instead use the following command

```
gt1 = Plot[%, {n, 0, N}, AxesLabel  
->{" n ", " Total Benefit (e.g. Million Pounds) "}]
```

This is not recommended at this stage. `In[15]` illustrates how to produce the final graph with title and with labelled axes.

The following command gives the optimal sample size, which maximizes the expected net benefit function,  $r(n)$ . The lower and the upper bounds can be determined from the previous graph.

```
In[14]:= FindMinimum[ - %, {n, {Lower bound, upper bound}}]
```

Here %% refers to the output of the second previous operation (once again, the output of In[12]). To produce a graph with a title and with labelled axes the following command may be used

```
In[15]:= g1= Show[ gt1, Frame -> True, FrameLabel ->
              {" Number of Observations", " Total Benefit
              (e.g. Million Pounds) ", " ", " "},
              DefaultFont -> {"Times-Italic", 8}, PlotLabel
-> FontForm["Title for the Graph", {"Helvetica-Bold",
12}]]
```

If the same calculations have been done for another set of parameters (e.g. different values for  $Mb_i$ ,  $i = 1, 2$ , or a different value for  $L$ ) then we may have another graph  $g_2$ , say. The following command illustrates how to show them beside each other

```
In[16]:= Show[GraphicsArray[{g1, g2}]]
```

Using the following command we can have both the graphs plotted on the same set of axes.

```
In[17]:= Show[g1, g2].
```

### 2.3. Illustration I (Trial 5 of Chapter 1)

In this section we provide an example to illustrate how the user may start from In[1] in section 2.1. and carry on to get the optimal sample size and different graphs. Data are taken from trial 5 of chapter 1.

```
In[1]:= A = 0.33
In[2]:= B = 0.49
In[3]:= c = 4000
```



```

In[4]:= mc = 0.41
In[5]:= mr = 0
In[6]:= tc = 0.21
In[7]:= tr = 0.41
In[8]:= s = 2
In[9]:= L = 0.27
In[10]:= Mb1 = 0
In[11]:= Mb2 = 15000000

```

Now we can run BeBay on this set of parameters by reading in the input file BeBayNK.all;

```

In[12]:= <<BeBayNK.all

```

This produces the following output which is the formula for the expected net benefit as a function of  $n$  or  $r(n)$  (see equation (4.29)).

```

Out[12]= -4000 n + 1.5 10 (6.25
                                1      0.5
> NIntegrate[Max[0, (-0.33 - 0.63 (-----)
                                4 + 0.0441 n
                                1.64      0.0441 n u
> +----- + -----) /
    4 + 0.0441 n      4 + 0.0441 n
                        2
    (-0.41 + u) / (2 (0.0441 + 4/n))
> (Sqrt[2] E Sqrt[(0.0441 + -)
                                4
                                n
> Pi]]\ , {u, Max[(22.6757 (-1.64 + ...

```

Applying the command

```

In[13]:= gt1 = Plot[%, {n, 0, 1000}]

```

where % refers to Out[12], we get the following graph for  $r(n)$  for  $n \in [0, 1000]$ . Note that here we called this graph gt1 to show that it is temporarily plotted. We shall call the final graph for this part g1.

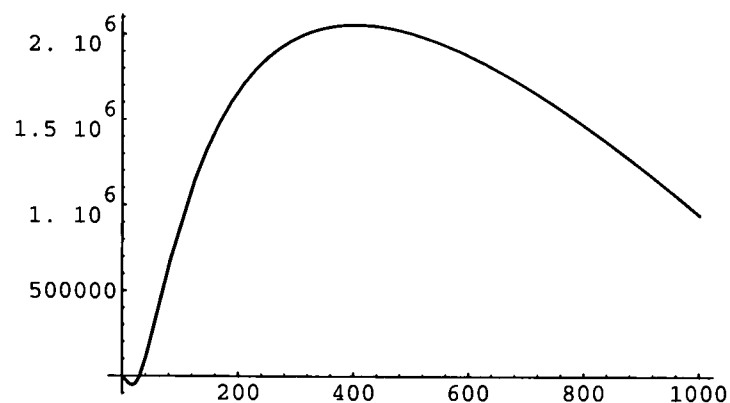


Figure 7.1: Expected net benefit as a function of trial size.

Now we can use the following command to determine the optimal sample size  $n^*$ . According to the graph  $n^*$  is around 400 so we may choose {Lower bound, upper bound} in In[14] as {200, 600} or any other (conservative or tight) interval containing the graph's peak.

```
In[14]:= FindMinimum[ - %, {n, {200, 600}}]
```

%% refers to the output of two operations ago, Out[12]. This gives

```

6
Out[14]= {-2.05906 106 , {n -> 402.929}}
```

which means that  $n = 402.929$  minimizes  $\{-r(n)\}$  (or, equivalently, maximizes  $r(n)$ ). So the maximum expected net benefit is  $2.05906 \times 10^6$  and it occurs at  $n^* = 402.929$ . To assess the *precision* of the calculation the user may use the following command

```
In[15]:= Precision[%]
```

On our machine (Sparc Server 1000) the precision was output as

```
Out[15]= 16
```

This means the total number of digits in the calculation of  $n$  is 16.

To assess the *accuracy* of  $n$  we may use

```
In[16]:= Accuracy[%]
```

This gives

```
Out[16]= 10
```

and means that the number of significant decimal digits to the right of the decimal point in every  $n$  (including  $n^*$ ) is 10. We shall discuss the numerical precision further in section 5.

Now we may want to title the graph and label its axes.

```
In[17]:= g1= Show[gt1, Frame -> True, FrameLabel
-> {" n", " r(n) ", " ", " "}, DefaultFont ->
{"Times-Italic", 8}, PlotLabel ->
FontForm["Graph Title", {"Helvetica-Bold", 12}]]
```

This produces the following graph

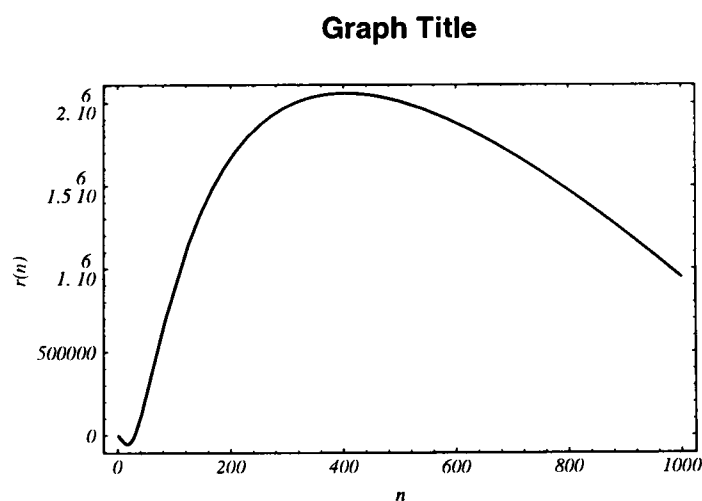


Figure 7.2: The previous graph with title and with labelled axes(g1).

If the same calculations have been done with a different value for regulator requirement  $L$ , say  $L = 0.41$ , we may have another graph,  $g2$  as shown below

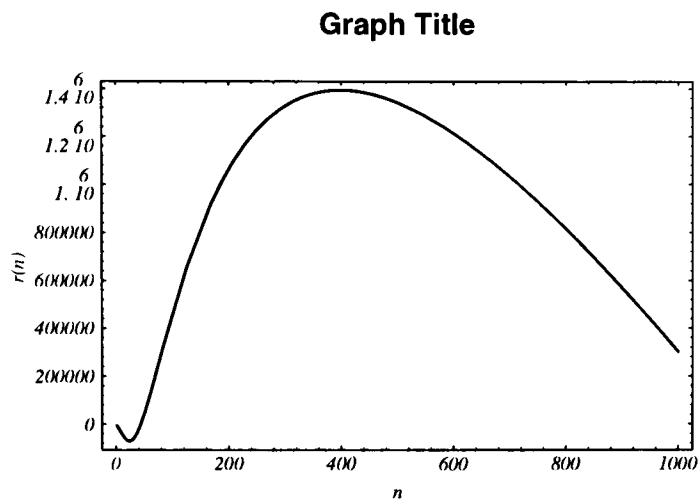


Figure 7.3: Expected net benefit function with a new value for  $L$  ( $=0.41$ ).

we may have, for example,

```
In[18]:= Show[graphicsArray[{gt1, gt2}]]
```

to get the following graph

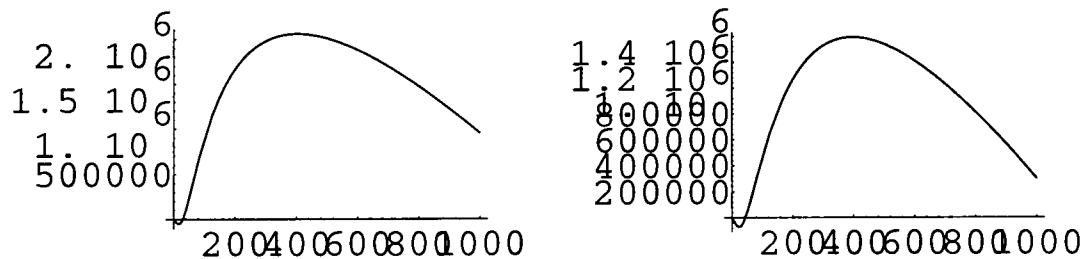


Figure 7.4:  $gt1$  and  $gt2$  beside each other.

Alternatively

```
In[19]:= Show[g1, g2]
```

plots both the  $g1$  and  $g2$  on one set of axes.

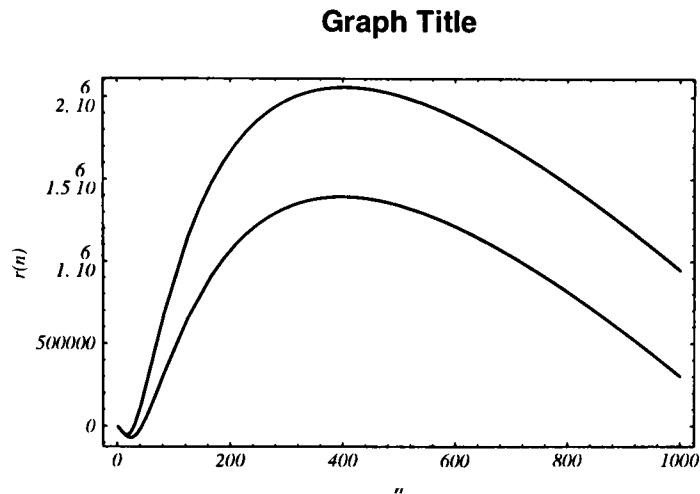


Figure 7.5: g1 and g2 on one set of axes.

## 2.4. Data Input; Unknown Variance

(see chapter 5 section 6)

As in the known variance case, we give some brief explanations first and then provide an illustration.

The user is asked to supply data outlining the characteristics of the trial. These may be summarized as:

A , B , c , mc , mr , ac , ar , gc , gr , wc , wr ,  
L , Mb1 , Mb2.

Some of these parameters have already been defined in section 2.1. of this chapter. The following list gives the meanings of those which have not been introduced. Note that the variance of observation  $\sigma^2$  has a prior distribution where  $a/\sigma^2 \sim \chi_{(g)}^2$  and the the amount of improvement with the new treatment  $\delta$  has a prior density where  $(\delta - \mu)/\sqrt{wa/g} \sim t_{(g)}$ .

ac: One of the parameters of the company's prior density for  $\sigma^2$ .

ar: One of the parameters of the regulator's prior density for  $\sigma^2$ .

gc: Another parameter of the company's prior density for  $\sigma^2$ .

gr: Another parameter of the regulator's prior density for  $\sigma^2$ .

wc: Parameter of the company's prior density for  $\delta$ .

wr: Parameter of the regulator's prior density for  $\delta$ .

It should be noted that the mean and the mode of the prior density for  $\sigma^2$  are, respectively,  $mode = a/(g - 2)$  and  $mean = a/(g + 2)$ . So we may write  $a = 4 (mode) \cdot (mean)/(mean - mode)$  and  $g = 2 (mode + mean)/(mean - mode)$ . Also the conditional prior density of  $\delta$  given  $\sigma^2$  has the variance  $w\sigma^2$ .

A routine which calculates the objective (or the expected net benefit) function, for which the benefit of using the new treatment is  $b_1\delta + b_2$ , has been written and saved on the file `BeBayNU.all`. The user must provide the above fourteen data and then use the following procedure to calculate the optimal size of the trial,  $n^*$ , which maximizes the expected net benefit from the trial.

## 2.5. Results; Unknown Variance

When the data entry process has been completed the user must call `BeBayNU.all` by the following command and continue as in the known variance case (see section 2.2.).

```
In[15] := <<BeBayNU.all
```

## 2.6. Illustration II (Case 2 of Chapter 5)

Applying `BeBayNU.all` to the following set of values for the parameters, we shall illustrate how one may get the desired results. Data are taken from case 2 of section 4 of chapter 5.

*Data entry:*

```
In[1] := A = 2
In[2] := B = 2.5
In[3] := c = 1000
In[4] := mc = 1
In[5] := mr = 0
In[6] := wc = 1
In[7] := wr = 1
In[8] := ac = 1
In[9] := ar = 1
In[10] := gc = 5
In[11] := gr = 3
```

```
In[12]:= L = 1.5
In[13]:= Mb1 = 0
In[14]:= Mb2 = 10000000
```

As before we need to read in the input file for this set of parameter values.

```
In[15]:= <<BeBayNU.all
```

Now we are in the position to get our outputs including

```
In[16]:= gt3 = Plot[%, {n, 0, 100}]
```

which gives

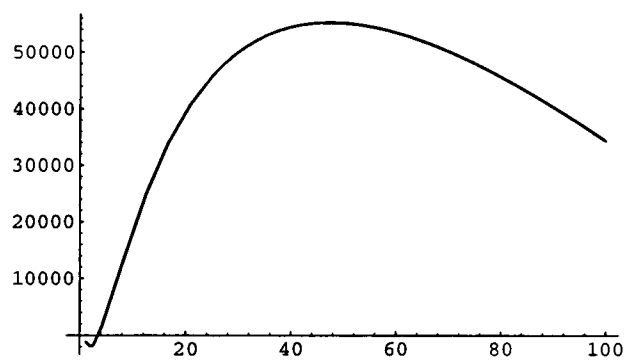


Figure 7.6: Expected net benefit as a function of trial size.

or

```
In[17]:= FindMinimum[- %%, {n, 20, 80}]
```

which gives

```
Out[17]= {-55309.8, {n -> 47.3909}}
```

and this means that the optimal size  $n^*$  is 47.3909 resulting in a net expected benefit of 55309.8.

Now let us enter

```
In[18]:= L= 2.25
```

which changes the value of L (from 1.5) to 2.25 and run the program again

```
In[19]:= <<BeBayNU.all
```

We can now calculate another set of results. For instance

```
In[20]:= FindMinimum[- %, {n, 20, 80}]]
```

produces

```
Out[20]= {-7931.51, {n -> 35.6886}}
```

or checking the precision and accuracy of the result

```
In[21]:= Precision[%]
```

```
Out[21]=16
```

```
In[22]:=Accuracy[%%]
```

```
Out[22]=10
```

Now

```
In[23]:= gt4 = Plot[ %%%, {n, 0, 100}]
```

which means plot a graph for the output of four operations ago (%%%)



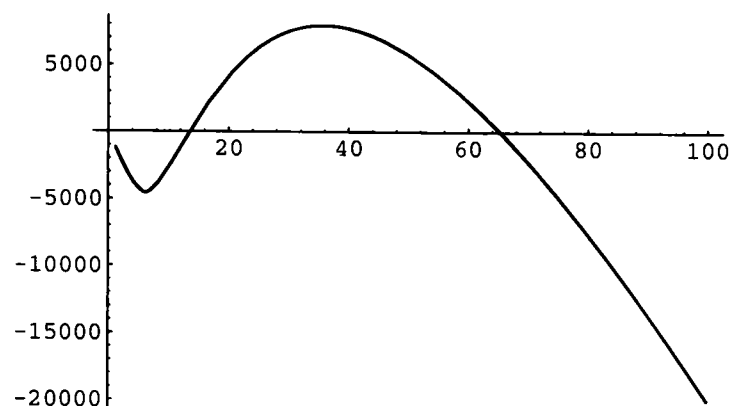


Figure 7.7: Expected net benefit as a function of trial size.

Proceeding as in In[17] to In[19] of the previous section gives the following graph.

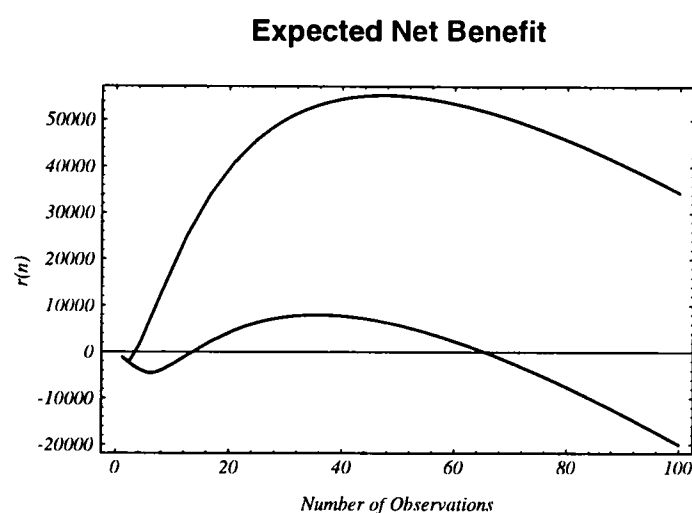


Figure 7.8: g3 and g4 on one set of axes.

### 3. Binomially Distributed Data

As in previous sections, we give some brief explanations first and then provide an illustration.

The user of the BeBay is asked to supply data outlining the characteristics of the trial under consideration. These may be summarized as

$A$  ,  $B$  ,  $c$  ,  $mc$  ,  $mr$  ,  $tc$  ,  $tr$  ,  $L$  ,  $Mb1$  ,  $Mb2$

All these parameters have already been introduced.

It should be noted here that if  $p$  has a beta prior density with the mean  $\mu$  and the variance  $\tau^2$  then the parameters of the beta density,  $\alpha$  and  $\beta$  may be found from  $\alpha = (\frac{\mu}{\tau})^2(1 - \mu) - \mu$  and  $\beta = (\frac{1-\mu}{\tau})^2\mu + (\mu - 1)$ . A routine which calculates the objective (or the expected net benefit) function, for which the benefit of using the new treatment is  $b_1p + b_2$  has been saved on the file called `BeBayB.all`. The user must provide the above ten data and then use the following procedure to calculate the optimal size of the trial,  $n^*$ , which maximizes the expected net benefit from the trial.

### 3.1. Results; Binary Data

When the data entry process has been completed the user needs to run BeBay for the binomially distributed data by using the command

```
In[11] := <<BeBayB.all
```

This reads in the input file `BeBayB.all`, evaluates the objective function for the ten parameter values entered above, and writes the expected net benefit as a function of  $n$ . Using the following command we can calculate  $r(n)$  for different values of  $n$  between 0 and  $N$ .

```
In[12] := Table[ %, {n, 0, N}]
```

The following command gives the maximum value of the objective function  $r(n)$

```
In[13] := Max[ % ]
```

To find the optimal sample size  $n^*$  use the command

```
In[14] := Position[ %, % ]
```

This means find the position of previous output, % (or `Out[13]`) in the output of two operations ago, %% (or `Out[12]`).

### 3.2. Illustration III (Trial 2, see chapter 6)

In this section we provide an example to illustrate how the BeBay user may start from In[1] and carry on to get the optimal sample size for a trial with binary data. Data are taken from example (1) of chapter 6.

```
In[1]:= A = 0.376
In[2]:= B = 0.564
In[3]:= c = 4000
In[4]:= mc = 0.47
In[5]:= mr = 0.2
In[6]:= tc = 0.235
In[7]:= tr = 0.2
In[8]:= L = 0.47
In[9]:= Mb1 = 0
In[10]:= Mb2 = 5000000
```

Now read in BeBayB.all by using the following command

```
In[11]:= <<BeBayB.all
```

This produces the following output which is the expected net benefit  $r(n)$  as a function of  $n$

```
Out[11]= -4000 n + 5000000 (Sum[3.93702 Beta[1.65 + x,
>                               1.86064 + n - x]
>
>   Binomial[n, x], {x, Max[Floor[(2.25 (1.8 + n) -
>   2 (4. + n) (0.6 - 0.47 (3. + n)) +
>
>   Sqrt[(2.25(1.8 + n)- 2(4. + n)(0.6 - 0.47(3. + n)))2 -
>   4 (6.25 + n) (-1.35 (2.4 + n) +
>
>   ....
```

Applying the following command the user can produce a table of the expected net benefit for different values of  $n$  between 0 and 100.

```
In[12]:= Table[%, {n, 0, 100}]
```

Again % refers to the previous output.

```
Out[12]= {0, -4000, -8000, 1.3448 106, 1.17317 106,
> 1.5304 106, 1.38505 106, 1.24017 106, 1.50089 106,
> 1.36601 106, 1.44284 106, 1.45642 106, 1.33782 106,
> 1.437 106, 1.41333 106, 1.48152 106, 1.41548 106,
> .
> .
> .
> 1.47624 106, 1.4584 106, 1.4722 106, 1.45406 106}
```

The maximum of these numbers may be found by

```
In[13]:= Max[ % ]
```

```
Out[13]= 1.55649 106
```

Note that if we check the precision and accuracy of calculation (as in previous sections; `In[14]:= Precision[%]` and `In[15]:= Accuracy[%]`) we get, respectively, 16 and 10. So on our machine the BeBay program automatically uses these significant digits in the calculation. As we shall see later it is possible to change the precision and the accuracy of calculation.

To find  $n^*$  which gives the above (optimal) value for  $r(n)$  we may use

```
In[14]:= Position[ %, % ]
```

which gives

```
Out[14]= {{44}}
```

This means  $n^* = 43$ . Note that we calculated  $r(n)$  for  $n = \{0, 1, \dots, N\}$ . So  $n^* = \text{Position} - 1$ . If we want to have the expected net benefit at the optimal sample size  $n^* = 43$  with a precision of eleven significant figures we may use

```
In[15]:= N[%, 11]
```

```
Out[15]= 1.5564921674 106
```

where `N[%, 11]` means “show the result of `Out[13]`, numerically, with eleven significant digits”.

## 4. Assessing the Truncation Error

Truncation error is a consequence of doing a finite number of steps in a calculation that would require an infinite number of steps to do exactly.

Numerical integration is an example of an operation that is affected by truncation error. The `NIntegrate` function, which represents the numerical integration routine used in the BeBay program, works by evaluating the integral at a finite number of points, and by using smoothing functions to approximate the integrand between those points. The difference between those smoothing functions and the actual integrand leads to truncation error. It is not possible by numerical techniques alone to get an exact value of the size of truncation error in the value  $n^*$ . However, the user of BeBay can obtain an estimate of this error at  $n = n_0$ , where  $n_0$  is known. There are two routines `BeBayNKT.all` and `BeBayNUT.all`, for normally distributed data with known and with unknown variance, respectively. Using these routines for a specified  $n$  with a particular set of parameter values, as shown in the following examples, the user may obtain an estimate of truncation error.

For example, let us consider trial 5 of chapter 1. The user of BeBay may estimate this error at  $n = n_0$ . For instance to estimate the error at the optimal sample size  $n^* = 403$  the following commands may be used

```
In[1] := A = 0.33
In[2] := B = 0.49
In[3] := c = 4000
In[4] := mc = 0.41
In[5] := mr = 0
In[6] := tc = 0.21
In[7] := tr = 0.41
In[8] := s = 2
In[9] := L = 0.27
In[10] := Mb1 = 0
In[10] := Mb2 = 15000000
In[11] := n = 403
In[12] := <<BeBayNKT.all
```

This gives

```
Out[12]= 7.54097 10-5
```

which means the upper bound for the truncation error is  $7.54097 \times 10^{-5}$ . The following command shows it with 16 significant figures.

```
In[13] := N[%, 16]
Out[13]= 0.00007540965452790261
```

Note that if we run BeBayNK.all with the above set of parameters we get  $2.05906 \times 10^6$  (or more precisely  $2.059064061851345 \times 10^6$ ).

In writing BeBayNKT.all we tried to compare the results of calculation using the fundamental theorem of calculus, where possible, to the numerical integration used in BeBayNK.all.

Now let us consider Illustration II. If we want to have an estimate of truncation error at the optimal sample size  $n^* = 48$  we may use the following commands

```
In[1]:= A = 2
In[2]:= B = 2.5
In[3]:= c = 1000
In[4]:= mc = 1
In[5]:= mr = 0
In[6]:= wc = 1
In[7]:= wr = 1
In[8]:= ac = 1
In[9]:= ar = 1
In[10]:= gc = 5
In[11]:= gr= 3
In[12]:= L = 1.5
In[13]:= Mb1 = 0
In[14]:= Mb2 = 10000000
In[15]:= n = 48
In[16]:= <<BeBayNUT.all
```

this gives

```
Out[16]= -48000 +  $\frac{187375 \cdot \text{Sqrt}[3]}{\text{Pi}}$ 
```

where `Pi` and `Sqrt[3]` represent, respectively,  $\pi$  and  $\sqrt{3}$ . In `Out[16]` both the  $\pi$  and  $\sqrt{3}$  are left symbolically instead of being approximated by a number with a finite significant figures. This means that `BeBayNUT.all` tries, where possible, to avoid the *representation* error<sup>1</sup> as well as the truncation error. The following command represents the previous output to 16 significant figures ( $\sqrt{3}$  and  $\pi$  are also being approximated to 16 significant figures).

```
In[17]:= N[%]

Out[17]= 55305.
```

---

<sup>1</sup>*Representation* error, or *roundoff* error, is the error associated with the fact that the computer keeps only a finite number of digits in calculations with irrational numbers.

If we want to see this output with a precision of 16, claimed by default precision on Sparc Server 1000, we may use the command

```
In[18]:= N[%, 16]
```

which means show the previous output, %, to 16 significant figures. This gives

```
Out[18]= 55305.01816793824
```

We can get other demonstrations. For example `In[17]:= N[%, 7]` gives `Out[17]= 55305.02` or `In[17]:= N[%, 13]` gives `Out[17]= 55305.01816794`. Now the upper bound for the truncation error for the above example may be obtained by comparing `Out[16]` and `Out[17]`

```
In[19]:= %% - %%%
```

```
Out[19]= -1.45519 10-11
```

in which %% and %%% refer, respectively, to the outputs of two and three operations ago.

As shown above the procedure for obtaining the truncation error in the unknown variance case is different from the one in the known variance case. This is because of the double integrals in the objective function (see (5.22) and (5.23) in section 6 of chapter 5), which have to be calculated numerically.

Truncation error in numerical integration can be reduced by evaluating the integrand at a greater number of points. Such a direct approach, however, is not very convenient. What we really want is to control the error itself, or equivalently, the number of significant digits in the result. In `NIntegrate` function used in BeBay program we can change the number of significant digits in the result by running BeBay on `BeBayNKP.all` for normal data with known variance and `BeBayNUP.all` for normal data with unknown variance. The following two examples illustrate how the user can change the precision of calculation to achieve more precise results. It should be noted



here that we used the option `PrecisionGoal` in writing `BeBayNKP.all` and `BeBayNUP.all`. This option is the number of significant figures that BeBay tries to get in the final answer. The default value for this option in *Mathematica* software is `Automatic`.

Changing the value for this option allows the user to see how the result of computations is sensitive to the significant figures, which BeBay tries to get in the final output. If the result changes rapidly when the value for the `PrecisionGoal` is increased then it is advised to use high values.

Suppose that the user wants to have the net expected benefit function for trial 5, discussed in Illustration I, at  $n = 100$  with different values for the `PrecisionGoal` (say 5 and 12).

```
In[1]:= A = 0.33
In[2]:= B = 0.49
In[3]:= c = 4000
In[4]:= mc = 0.41
In[5]:= mr = 0
In[6]:= tc = 0.21
In[7]:= tr = 0.41
In[8]:= s = 2
In[9]:= L = 0.27
In[10]:= Mb1 = 0
In[11]:= Mb2 = 15000000
In[12]:= n = 100
In[13]:= PG = 5
In[14]:= <<BeBayNKP.all
```

This gives

```
Out[14]= 888687.
```

To see this value computed to 16 significant figures we may use

```
In[15]:= N[%, 16]
Out[15]= 888687.141674561
```

Now let us change the `PrecisionGoal` for this calculation to 12

```
In[16] := PG = 12
```

We may run `BeBay` on `BeBayNKP.all` and get the following results

```
In[17] := <<BeBayNKP.all
Out[17]= 888687.
In[18] := N[%, 16]
Out[18]= 888687.141692357
```

As we see the difference between `Out[18]` and `Out[15]` is  $1.77962 \times 10^{-5}$ . Since increasing the value for the `PrecisionGoal` increases both the duration of the calculation and the memory used one may conclude that it was not worth changing the value from 5 to 12 for this example.

Now let us consider Illustration II. The following commands allow us to compare the expected net benefit calculated with `PrecisionGoal = 5` and with `PrecisionGoal = 7` at  $n = 48$ .

```
In[1] := A = 2
In[2] := B = 2.5
In[3] := c = 1000
In[4] := mc = 1
In[5] := mr = 0
In[6] := wc = 1
In[7] := wr = 1
In[8] := ac = 1
In[9] := ar = 1
In[10] := gc = 5
In[11] := gr = 3
In[12] := L = 1.5
In[13] := Mb1 = 0
In[14] := Mb2 = 10000000
In[15] := n = 48
In[16] := PG = 5
In[17] := <<BeBayNUP.all
```

This gives

```
Out[17]= 58394.18459931144
```

Now with 7 digits

```
In[18]:= PG = 7  
In[19]:= <<BeBayNUP.all  
Out[19]= 58394.19015326887
```

So the difference between the results of the two calculations is  $5.55396 \times 10^{-3}$ . One might think it was worth using `PrecisionGoal = 7` for this example.

Although it is possible to change the precision of calculation when using BeBay, the default precision set up in *Mathematica* is normally sufficient. For instance, on Sparc Server 1000 the default precision is between  $[6, \infty]$ , where  $\infty$  is the precision of an integer number.

## 5. Numerical Precision

As already mentioned, the *precision* of a real number is the number of decimal digits in it which are treated as significant for computation. The *accuracy* of a real number is the number of decimal digits which appear to the right of the decimal point. Precision is thus a measure of the *relative* error in a number, while accuracy is a measure of *absolute* error.

In general, *Mathematica* distinguishes two kinds of approximate real numbers: *arbitrary-precision* ones, and *machine-precision* ones. Arbitrary-precision numbers can contain any number of digits, and their precision is adjusted during numerical computations. Machine-precision numbers, on the other hand, contain a fixed number of digits, and their precision remains unchanged throughout the computations.

Machine-precision numbers work by making direct use of the numerical capabilities of the underlying computer system. For instance on the Sparc Server 1000, which we have been using for our numerical calculations the machine precision is 16 decimal digits:

```
In[1]:= $MachinePrecision
```

```
Out[1]= 16
```

When one does calculations with arbitrary-precision numbers, *Mathematica* keeps track of precision at all points. It treats arbitrary-precision numbers as representing the values of quantities where a certain number of digits are known, and the rest are unknown. In general, an arbitrary-precision number  $x$  is taken to have `Precision[x]` digits which are known exactly, followed by an infinite number of digits which are completely unknown.

When one does a computation `BeBay`, which is written in *Mathematica*, keeps track of which digits in the result could be affected by unknown digits in the input. It sets the precision of the result so that no affected digits are ever included. This procedure ensures that all digits returned by `BeBay` are correct, whatever the values of the unknown digits may be.

We have written routines, which incorporate the `WorkingPrecision` option of *Mathematica*, for normally distributed data with known variance and with unknown variance. These are, respectively, `BeBayNKW.all` and `BeBayNUW.all`. The default value for the `WorkingPrecision` option, which is the number of digits to use in internal computations, is `$MachinePrecision`. Running `BeBay` on these files allows the user to specify the precision.

The following example calculates the expected net benefit of illustration I at the optimal sample size,  $n^* = 403$ , using 20-digits-precision.

```
In[1]:= A = 0.33
In[2]:= B = 0.49
In[3]:= c = 4000
In[4]:= mc = 0.41
In[5]:= mr = 0
In[6]:= tc = 0.21
In[7]:= tr = 0.41
In[8]:= s = 2
In[9]:= L = 0.27
In[10]:= Mb1 = 0
In[11]:= Mb2 = 15000000
In[12]:= n = 403
```

```
In[13] := WP = 20
In[14] := <<BeBayNKW.all
```

This gives the output

```
Out[14]= 2.059064061926754 106
```

Typical machine-precision numbers in *Mathematica* are represented as “double-precision-floating-point numbers” in the underlying computer system. On most current computers, such numbers contain a total of 64 binary bits, typically yielding 16 decimal digits of mantissa. If the user wants to use the default value for the `WorkingPrecision`, (s)he may skip `In[13]` and run `BeBayNK.all` in the above example.

The main advantage of using the built-in floating-point capabilities of the computer is speed. Arbitrary-precision numerical calculations, which do not make such direct use of these capabilities, are usually many times slower than machine-precision calculations.

Using 5-digits-precision in the following example, the result of calculating the net expected benefit for illustration II at the optimal sample size,  $n^* = 48$ , is shown.

```
In[1] := A = 2
In[2] := B = 2.5
In[3] := c = 1000
In[4] := mc = 1
In[5] := mr = 0
In[6] := wc = 1
In[7] := wr = 1
In[8] := ac = 1
In[9] := ar = 1
In[10] := gc = 5
In[11] := gr = 3
In[12] := L = 1.5
In[13] := Mb1 = 0
In[14] := Mb2 = 10000000
In[15] := n = 48
```

```
In[16]:= WP = 5
In[17]:= <<BeBayNUW.all
```

which gives

```
Out[17]= 58394.18946
```

The result is accurate to 11 digits:

```
In[18]:= Accuracy[%]
Out[18]= 11
```

Although the user of the BeBay program can change the working precision, if Machine-precision uses 64 binary bits, which yields to 16 decimal digits in the mantissa, it is almost certainly adequate to work with Machine-precision. This reduces both the duration of the calculation and the memory used (for more details visit [http://www.wri.com/support /Math/Numerics](http://www.wri.com/support/Math/Numerics)).

## 6. A Summary of the BeBay Program Files

In the following list we give the names of the BeBay program files, with brief descriptions of their contents. The user can call these files in a *Mathematica* session using the command `In[n]:= <<filename`.

**BeBayB.all** BeBay program for binary data with the benefit function of the form  $b_1p + b_2$

**BeBayNK.all** BeBay program for normally distributed data with known variance and with the benefit function of the form  $b_1\delta + b_2$

**BeBayNKP.all** BeBay program using the option `PrecisionGoal = PG` at  $n = n_0$  for calculating the expected net benefit in normally distributed data with known variance

**BeBayNKT.all** BeBay program estimating the truncation error at  $n = n_0$

for normally distributed data with known variance

BeBayNKW.a11 BeBay program using the option `WorkingPrecision = WP` at  $n = n_0$  for calculating the expected net benefit in normally distributed data with known variance

BeBayNU.a11 BeBay program for normally distributed data with unknown variance and with the benefit function of the form  $b_1\delta + b_2$

BeBayNUP.a11 BeBay program using the option `PrecisionGoal = PG` at  $n = n_0$  for calculating the expected net benefit in normally distributed data with unknown variance

BeBayNUT.a11 BeBay program estimating the truncation error at  $n = n_0$  for normally distributed data with unknown variance

BeBayNUW.a11 BeBay program using the option `WorkingPrecision = WP` at  $n = n_0$  for calculating the expected net benefit in normally distributed data with unknown variance

# Chapter 8

## Desirable Further Work

In this chapter some desirable further work is discussed. We shall start with some more general assumptions on the function representing the number of subsequent user of the new treatment by introducing a continuous bounded function in section 1. We shall briefly discuss the reasons for considering this function and show how the objective (or the expected net benefit) function for the known variance case may be changed.

In any trial some patients receive the inferior treatment. It might be worth investigating benefit (cost) functions which are a mixture of the two already mentioned in the general model to deal with this. In section 2 we briefly discuss the problem for normally distributed data with known variance where the benefit of using the new treatment is written from the public health standpoint.

We have already mentioned how we may apply the BeBay methodology to the case where the size of the control group is not the same as the size of the treated group (see section 5 of chapter 4). In section 3 we point to a Bayesian argument to discuss how we may use the prior information on each treatment to decide how to distribute the total sample size,  $n$ , between the control and treated groups.

Section 4 discusses the regulator's requirement for granting a licence to the new treatment from a frequentist perspective. It is important to work out the details of applying BeBay in this situation as regulators often do use the frequentist approach for assessing the difference between the performances of two treatments.

To extend our model to Poisson data, we have supposed that  $n$  is large



and applied the Central Limit Theorem. Further investigation is needed to obtain a methodology which is valid for smaller values of  $n$ .

We have also assumed  $n$  is large and applied the Central Limit Theorem to extend our methodology to survival analysis. Further investigation is needed to develop BeBay for smaller values of  $n$ .

In chapter 6 we discussed how we may apply BeBay methodology to trials with binary outcomes for which there is just one group of patients in the trial. Further investigation is needed to adapt the methodology to the case where there is a control group in the trial.

Another potential substantial task is to work out how the BeBay methodology may be applied to 2 or 3 stage trials.

## 1. More General Assumptions

While our assumptions for constructing the function used to describe  $m$  are rational and easy to interpret, it should be noted that the function  $m$  has two points where the derivative does not exist. It might be worth considering using a function for  $m$  of the form

$$m = \frac{M}{1 + e^{-\delta}},$$

which is a kind of logistic function.

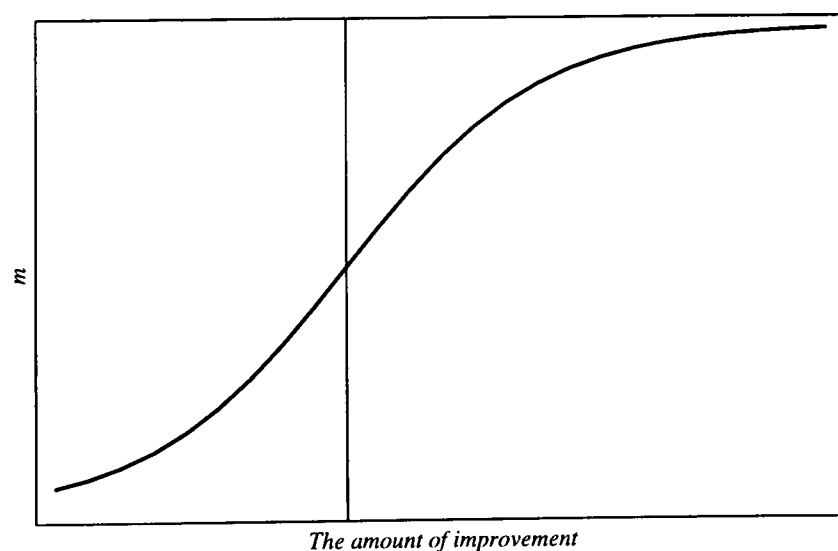


Figure 8.1: Number of subsequent users of the new treatment.

In the known variance case if we work with the scaled random variable  $\frac{\mu'}{\sigma}$  then we may write  $m$  as

$$m = \frac{M}{1 + e^{-\frac{\mu'(\bar{z}_n)}{\sigma}}}. \quad (8.1)$$

The objective function for the public health version of the benefit function may be written as

$$\begin{aligned} r(n) &= \int_{\bar{z}_n} \frac{M}{1 + e^{-\frac{\mu'(\bar{z}_n)}{\sigma}}} \left( \frac{b\mu'(\bar{z}_n)}{\sigma} \right) f(\bar{z}_n) d\bar{z}_n - cn \\ &= \frac{Mb}{\sigma} \int_{-\infty}^{\infty} \frac{\mu + \frac{n\tau^2}{\sigma^2 + n\tau^2}(\bar{z}_n - \mu)}{1 + e^{-\frac{1}{\sigma}[\mu + \frac{n\tau^2}{\sigma^2 + n\tau^2}(\bar{z}_n - \mu)]}} f(\bar{z}_n) d\bar{z}_n - cn \\ &= Mb \int_{-\infty}^{\infty} \frac{\frac{\mu}{\sigma} + \sqrt{n} \frac{\tau}{\sigma} (\frac{\tau^2}{\sigma^2} + n)^{-1/2} (\frac{\bar{z}_n - \mu}{\sqrt{(\sigma^2 + n\tau^2)/n}})}{1 + e^{-[\frac{\mu}{\sigma} + \sqrt{n} \frac{\tau}{\sigma} (\frac{\tau^2}{\sigma^2} + n)^{-1/2} (\frac{\bar{z}_n - \mu}{\sqrt{(\sigma^2 + n\tau^2)/n}})]}} f(\bar{z}_n) d\bar{z}_n - cn. \end{aligned} \quad (8.2)$$

Putting  $u = \frac{\bar{z}_n - \mu}{\sqrt{\frac{\sigma^2}{n} + \tau^2}}$  and using the change of variables,  $C = \frac{c}{Mb}$ ,  $D = \frac{\mu}{\sigma}$ ,  $R(n) = \frac{r(n)}{Mb}$ , and  $T = \frac{\tau}{\sigma}$ , (8.2) becomes

$$R(n) = \int_{-\infty}^{\infty} \frac{D + \sqrt{n}T(T^{-2} + n)^{-1/2}u}{1 + e^{-[D + \sqrt{n}T(T^{-2} + n)^{-1/2}u]}} \phi(u) du - Cn. \quad (8.3)$$

Thus the scaled objective function,  $R(n)$ , reduces to a function of  $C$ ,  $D$ , and  $n$ .

This form for the function  $m$  may also be used with the commercial version of the objective function and when the observation variance  $\sigma^2$  is not known.

It should be noted that this change in  $m$  is expected to make calculations, especially in the unknown variance case, more complicated and they take longer to carry out.

## 2. A Note On the Cost Structure

In our general model for normally distributed data with known variance we assumed that the public health version of the objective function could be written as

$$r(n) = \int_{\bar{z}_n} \int_{\delta} m\left(\frac{b\delta}{\sigma}\right) d\Pi^n(\delta|\bar{z}_n) f(\bar{z}_n) d\bar{z}_n - (cn + d),$$

where  $d$  is the setup cost.

It is obvious that in any trial some patients receive the worse treatment. Suppose these number  $kn$ , ( $0 \leq k \leq 1$ ). At the same time there is a benefit for the  $(1 - k)n$  patients receiving the better treatment in the trial. If these contributions to cost and benefit are included the objective function becomes

$$r(n) = \int_{\bar{z}_n} \int_{\delta} [m + (1 - k)n]\left(\frac{b\delta}{\sigma}\right) d\Pi^n(\delta|\bar{z}_n) f(\bar{z}_n) d\bar{z}_n - c_1 n - c_2(kn) - d,$$

where  $c_1$  is the cost per patient in the trial and  $c_2$  is the cost for a patient receiving the inferior treatment. The objective function now becomes

$$r(n) = \int_{\bar{z}_n} \int_{\delta} m\left(\frac{b\delta}{\sigma}\right) d\Pi^n(\delta|\bar{z}_n) f(\bar{z}_n) d\bar{z}_n + \frac{b(1 - k)n}{\sigma} \mu - (c_1 + kc_2)n - d.$$

Proceeding as before it may be shown that

$$\begin{aligned} R(n) = & \frac{1}{B - A} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & - \frac{A}{B - A} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & - \frac{1}{B - A} 1.5(T^{-2} + n)^{-1/2} \int_{h_1(A,n)}^{h_2(B,n)} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \\ & \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & + \int_{h_2(A,n)}^{\infty} [D + \sqrt{n}T(T^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\ & + \left\{ \frac{1 - k}{M} \right\} nD - \left\{ \frac{c_1}{Mb} + \frac{c_2}{Mb}k \right\} n - \frac{d}{Mb}, \end{aligned} \tag{8.4}$$

where  $D = \frac{\mu}{\sigma}$ ,  $R(n) = \frac{r(n)}{Mb}$ ,  $T = \frac{\tau}{\sigma}$ ,

$$h_1(A, n) = \frac{[A + 1.5(T^{-2} + n)^{-1/2} - D](T^{-2} + n)^{1/2}}{T\sqrt{n}},$$

$$h_2(A, n) = \frac{[B + 1.5(T^{-2} + n)^{-1/2} - D](T^{-2} + n)^{1/2}}{T\sqrt{n}}.$$

The same approach may be applied with the commercial benefit function and to the case of unknown variance. It may be of some interest to write a program taking these new parameters together with the regulator's ones into account.

### 3. Numbers of Patients on Each of the Two Treatments; Known Variance

In our general method for finding the optimal size of trial we assume that  $n$  patients are allocated to both the new treatment and control groups. In this section we shall apply a Bayesian argument to show how we may use the prior information on the two treatment effects to decide how many patients are needed in each of these two groups.

**Question:** Suppose that  $\theta_p$  and  $\theta_t$  are, respectively, the unknown responses for patients using placebo and for those using the new treatment. We assume  $\theta_p \sim N(\mu_p, \tau_p^2)$ ,  $\theta_t \sim N(\mu_t, \tau_t^2)$ , and we are interested in  $\delta = \theta_t - \theta_p$ . The question is how we may determine the number of patients on each of these two groups.

The prior density for  $\delta$  is

$$N(\mu_t - \mu_p, \tau_t^2 + \tau_p^2).$$

Let  $x_1, x_2, \dots, x_{n_t}$  and  $y_1, y_2, \dots, y_{n_p}$  be, respectively, the outcomes for patients using the new treatment and for those using placebo, and let  $X_i \sim N(\theta_t, \sigma^2)$  and  $Y_j \sim N(\theta_p, \sigma^2)$ . If  $\bar{x}_t$  and  $\bar{y}_p$  are the sample means then the posterior distribution of  $\delta$  is

$$N\left(\frac{\mu_t \tau_t^{-2} + \bar{x}_t n_t \sigma^{-2}}{\tau_t^{-2} + n_t \sigma^{-2}} - \frac{\mu_p \tau_p^{-2} + \bar{y}_p n_p \sigma^{-2}}{\tau_p^{-2} + n_p \sigma^{-2}}, (\tau_t^{-2} + n_t \sigma^{-2})^{-1} + (\tau_p^{-2} + n_p \sigma^{-2})^{-1}\right). \quad (8.5)$$

One possible answer to the above question is to find  $n_t$  and  $n_p$  so as to minimize the posterior variance of the parameter of interest,  $\delta$ . Applying the Lagrange multiplier technique the question becomes

$$\begin{aligned} \min_{n_t, n_p} f(n_t, n_p) &= (\tau_t^{-2} + n_t \sigma^{-2})^{-1} + (\tau_p^{-2} + n_p \sigma^{-2})^{-1}, \\ \text{subject to: } g(n_t, n_p) &= n_t + n_p - n = 0. \end{aligned} \quad (8.6)$$

To find the stationary points of  $f(n_t, n_p)$  we need to solve the following system of equations

$$\begin{cases} \frac{\partial f}{\partial n_t} + \lambda \frac{\partial g}{\partial n_t} = 0 \\ \frac{\partial f}{\partial n_p} + \lambda \frac{\partial g}{\partial n_p} = 0 \\ n_t + n_p - n = 0, \end{cases}$$

or

$$\begin{cases} -(\tau_t^{-2} + n_t \sigma^{-2})^{-2} \sigma^{-2} + \lambda = 0 \\ -(\tau_p^{-2} + n_p \sigma^{-2})^{-2} \sigma^{-2} + \lambda = 0 \\ n_t + n_p - n = 0. \end{cases} \quad (8.7)$$

Solving for  $n_t$  and  $n_p$  in the first two equations and putting them into the third one gives

$$\lambda^2 = \frac{n + (\tau_t^{-2} + \tau_p^{-2}) \sigma^2}{2 \sigma^6}.$$

So

$$n_t = (\tau_p^{-2} - \tau_t^{-2}) \frac{\sigma^2}{2} + \frac{n}{2}, \quad n_p = (\tau_t^{-2} - \tau_p^{-2}) \frac{\sigma^2}{2} + \frac{n}{2}, \quad (8.8)$$

or finally

$$n_t = \max\{0, (\tau_p^{-2} - \tau_t^{-2}) \frac{\sigma^2}{2} + \frac{n}{2}\}, \quad n_p = \max\{0, (\tau_t^{-2} - \tau_p^{-2}) \frac{\sigma^2}{2} + \frac{n}{2}\}.$$

### 3.1. Discussion

If the prior information on the placebo and on the new treatment are the same (ie;  $\tau_p^2 = \tau_t^2$ ) then  $n_t = n_p = n/2$ . This means a sample of size  $n/2$  should be taken from each of the two groups.

If the prior information on the placebo is more than for the new treatment (ie;  $\tau_p^2 < \tau_t^2$ ) then (8.8) implies  $n_p < n/2 < n_t$ . This means that we have to make the treated group larger than the control group. If the prior information on the placebo is less than for the new treatment (ie;  $\tau_p^2 > \tau_t^2$ ) then (8.8) implies  $n_t < n/2 < n_p$ . So we need to make the control group larger than the treated group.

It might be worth trying to investigate the question for the unknown variance case and for the binomially distributed data to work out the conditions on the parameter values under which we may find that  $\frac{n_p}{n_t} \approx 1$ .

## 4. Frequentist Regulatory Authorities

Assuming that the regulator uses a frequentist approach for assessing the difference between the performances of the two treatments, as often happens, it is important to consider how the BeBay methodology may be adapted. In this section we discuss an approach to this problem for normally distributed data with known variance. It is worth investigating the problem for other cases like unknown variance and binomially distributed data.

### 4.1. Known Variance

Let us assume that the regulator wishes to test the hypothesis  $H_0$  that  $\delta = 0$  against the alternative  $H_1$  that  $\delta > 0$ . If the data are normally distributed and if the size of the test is  $\alpha$  then the regulator's criterion for rejecting  $H_0$  in favour of  $H_1$  is

$$\text{reject } H_0 \text{ if and only if } \bar{z}_n > z_\alpha \sqrt{\frac{\sigma^2}{n}}. \quad (8.9)$$

In other words, the regulator grants a licence to the new treatment, provided that (8.9) holds. Note that here the paired comparison case of normally dis-

tributed data, discussed in chapter 4, is considered (see section 5 of chapter 4).

Recall that the function representing the number of subsequent users of the new treatment,  $m$ , is of the form shown below

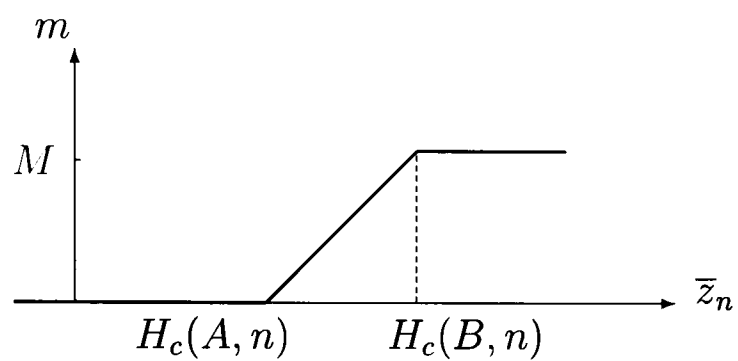


Figure 8.2: Number of subsequent users

where

$$H_c(A, n) = \frac{(\sigma^2 + n\tau_c^2)(A + 1.5\tau_c') - \sigma^2\mu_c}{n\tau_c^2},$$

$$H_c(B, n) = \frac{(\sigma^2 + n\tau_c^2)(B + 1.5\tau_c') - \sigma^2\mu_c}{n\tau_c^2}, \quad (8.10)$$

and  $\mu_c$  and  $\tau_c$  are, respectively, the prior mean and the prior standard deviation of  $\delta$  assumed by the drug company conducting the trial.

#### 4.1.1. Public Health Benefit Function

As before there are three cases to be considered

$$\text{Case 1: } z_\alpha \sqrt{\sigma^2/n} \leq H_c(A, n) < H_c(B, n)$$

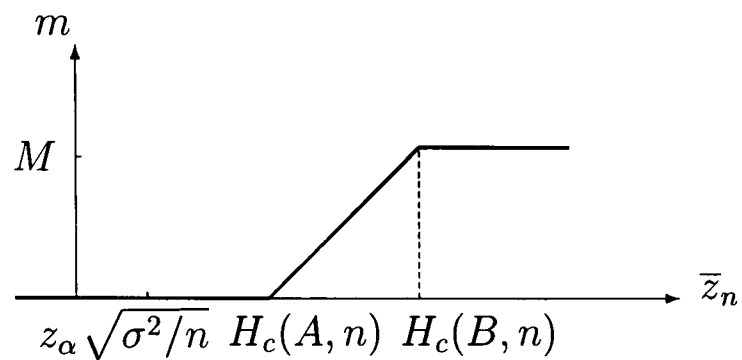


Figure 8.3: Number of subsequent users;  
Low Frequentist Regulator Requirement.

The objective function for this case is the same as the one with no regulator so the problem of sample size determination is unchanged.

$$\text{Case 2: } H_c(A, n) \leq z_\alpha \sqrt{\sigma^2/n} < H_c(B, n)$$

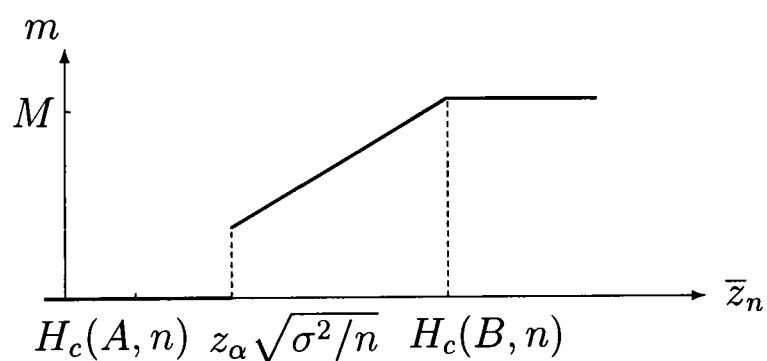


Figure 8.4: Number of Subsequent users;  
Intermediate Frequentist Regulator Requirement.

As shown in figure 8.4 there are no subsequent users unless an approval is granted. For this case the problem is to find  $n^*$  which maximizes the following objective function



$$\begin{aligned}
R(n) = 0 + \int_{z_\alpha \sqrt{\sigma^2/n}}^{H_c(B,n)} \frac{1}{B-A} & \left[ \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n - A - 1.5 \frac{\tau_c'}{\sigma} + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right] \\
& \times \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n \\
& + \int_{H_c(B,n)}^{\infty} \left( \frac{n\tau_c^2}{(\sigma^2 + n\tau_c^2)\sigma} \bar{z}_n + \frac{\sigma^2 \mu_c}{(\sigma^2 + n\tau_c^2)\sigma} \right) f(\bar{z}_n) d\bar{z}_n - Cn.
\end{aligned} \tag{8.11}$$

Putting  $u = \frac{\bar{z}_n - \mu_c}{\sqrt{\frac{\sigma^2}{n} + \tau_c^2}}$  the objective function may be written as

$$\begin{aligned}
R(n) = \frac{1}{B-A} \int_{f_r(\alpha,n)}^{h_2(B,n)} & [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{f_r(\alpha,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{f_r(\alpha,n)}^{h_2(B,n)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \\
& \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{h_2(B,n)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn,
\end{aligned} \tag{8.12}$$

where  $T_c = \frac{\tau_c}{\sigma}$ ,  $D_c = \frac{\mu_c}{\sigma}$ ,  $C = \frac{c}{Mb}$ ,  $R(n) = \frac{r(n)}{Mb}$ , and

$$\begin{aligned}
f_r(\alpha, n) &= \frac{z_\alpha - \sqrt{n}T_c}{\sqrt{1 + nT_c}}, \\
h_1(A, n) &= \frac{[A + 1.5(T_c^{-2} + n)^{-1/2} - D_c](T_c^{-2} + n)^{1/2}}{T_c \sqrt{n}}, \\
h_2(B, n) &= \frac{[B + 1.5(T_c^{-2} + n)^{-1/2} - D_c](T_c^{-2} + n)^{1/2}}{T_c \sqrt{n}},
\end{aligned}$$

It should be noted that we used the above change of variable to convert

the inequalities  $z_\alpha \sigma / \sqrt{n} < \bar{z}_n < \frac{\sigma(\sigma^2 + n\tau_c^2)(B + 1.5\tau'_c/\sigma) - \sigma^2\mu_c}{n\tau_c^2}$  to  $f_r(\alpha, n) < u < h_2(B, n)$ .

In this case the minimum amount of improvement in performance required by the regulator to grant a licence is bigger than the minimum improvement needed for some users to switch to the new treatment.

Case 3:  $H_c(A, n) < H_c(B, n) \leq z_\alpha \sqrt{\sigma^2/n}$

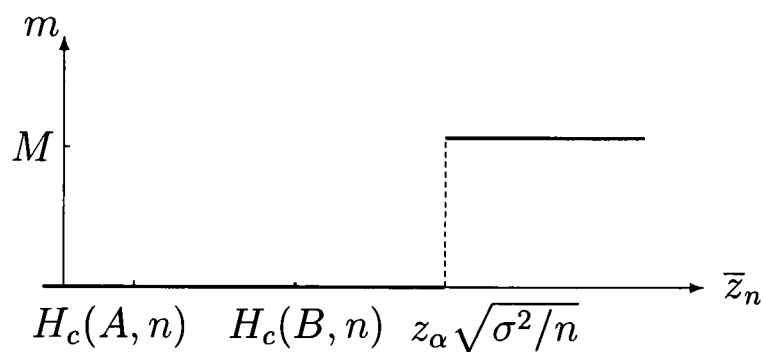


Figure 8.5: Number of subsequent users;  
High Frequentist Regulator Requirement

In this case the minimum amount of improvement in performance required by the regulator to grant a licence to the new treatment is so high that if it is achieved all the potential users of the new treatment do use it in preference to the old treatment.

Using the same change of variable as in (8.12) the objective function for this case is

$$R(n) = \int_{f_r(\alpha, n)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn. \quad (8.13)$$

Now we can establish the following theorem.

**Theorem:** If the regulator's requirement is as in (8.9) and the benefit of using the new treatment is from the public health standpoint, then the scaled

expected net benefit from carrying out a trial of size  $n$  is

$$\begin{aligned}
R(n) = & \frac{1}{B-A} \int_{\max(h_1, f_r)}^{\max(h_2, f_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{\max(h_1, f_r)}^{\max(h_2, f_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{\max(h_1, h_r)}^{\max(h_2, f_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \\
& \quad \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{\max(h_2, h_r)}^{\infty} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn.
\end{aligned} \tag{8.14}$$

#### 4.1.2. Commercial Benefit Function

Applying the same analysis as for public health version of the benefit function we can state the following theorem.

**Theorem:** If the regulator's requirement is as in (8.9) and the benefit of using the new treatment is expressed from the commercial standpoint, then the scaled expected net benefit for carrying out a trial of size  $n$  may be written as

$$\begin{aligned}
R(n) = & \frac{1}{B-A} \int_{\max(h_1, f_r)}^{\max(h_2, f_r)} [D_c + \sqrt{n}T_c(T_c^{-2} + n)^{-1/2}u] \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{A}{B-A} \int_{\max(h_1, f_r)}^{\max(h_2, f_r)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& - \frac{1}{B-A} 1.5(T_c^{-2} + n)^{-1/2} \int_{\max(h_1, h_r)}^{\max(h_2, f_r)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \\
& + \int_{\max(h_2, f_r)}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du - Cn.
\end{aligned} \tag{8.15}$$

□

# Bibliography

- ADCOCK, C. J. (1987) A Bayesian Approach to Calculating Sample Size for Multinomial sampling. *The Statistician* **36**, 155–159.
- 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 Sizes – some comments on the paper by Joseph, Wolfson, and Berger. *The Statistician* **44**, 155–161.
- ADCOCK, C. J. (1997) Sample Size Determination: A Review. *The Statistician* **46**, 261–283.
- BERGER, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis*. Springer-Verlag New York.
- BILLINGSLEY, P. (1986) *Probability and Measure, 2nd edn.* John Wiley and Sons, New York.
- CLAXTON, K. AND POSNETT, J. (1996) An Economic Approach to Clinical Trial Design and Research Priority–Setting. *Health Economics* **5**, 513–524.
- COAD, D. S. AND ROSENBERGER, W. F. (1999) A Comparison of the Randomized Play–The–Winner Rule and the Triangular Test for Clinical Trials with Binary Responses. *Statistics in Medicine* **18**, 761–769.
- COCHRAN, W. G. (1963) *Sampling techniques, 2nd edn.* John Wiley, New York.
- COCHRAN, W. G. AND COX G.M. (1957) *Experimental Designs*. John Wiley, New York.

- COHEN, J. (1977) *Statistical Power Analysis for the Behavioural Sciences*. Academic Press, New York.
- DE GROOT, M. (1985) *probability and statistics*, 2nd edn. Addison–Wesley, Reading.
- DESU, M. M. AND RAGHAVARAO, D. (1990) *Sample Size Methodology*. Academic Press, Boston.
- DURRETT, R. (1996) *Probability: Theory and Examples*, 2nd edn. Duxbury Press Belmont.
- FEIGEL, P. (1978) A Graphical Aid for Determining Sample Size when Comparing Two Independent Proportions. *Biometrics* **34**, 111–122.
- FRASER, D. A. S. AND GUTTMANN, I. (1956) Tolerance Regions. *Ann. Math. Statist.* **27**, 162–179.
- GARDNER, M. J. AND ALTMAN, D. G. (1986) Confidence Intervals Rather than p-values: Estimation Rather than Hypothesis Testing. *British Medical Journal* **292**, 746–750.
- GITTINS, J. C. (1989). *Multi-armed Bandit Allocation Indices*. John Wiley and Sons, Chichester.
- GITTINS, J. C. AND PEZESHK, H. (2000) How Large Should A Clinical Trial Be? *The Statistician* **49**, part 2, pp. 177–187.
- GITTINS, J. C. AND PEZESHK, H. (2000) A Behavioural Bayes Method for Determining the Size of a Clinical Trial. *Drug Information Journal* Vol. 34, pp. 355–363.
- GOLDSTEIN, M. (1981) A Bayesian Criterion for Sample Size. *Ann. Statist.* **9**, 670–672.
- GOULD, A. L. (1993) Sample Sizes for Event Rare Equivalence Trials using Prior Information. *Statistics in Medicine* **12**, 1209–1223.
- GRUNDY, P. M., HEALY, M. J. R., AND REES, D. H. (1956) Economic Choice of the Amount of Experimentation. *J. R. Statist. Soc. A* **18**, 32–48.
- HASEMAN, J. K. (1978) Exact Sample Sizes for use with the Fisher–Irwin Test for  $2 \times 2$  Tables. *The Biometrics* **34**, 106–109.

- HUTTON, J. L. AND OWENS, R. G. (1993) Bayesian Sample Size Calculations and Prior Beliefs about Child Sexual Abuse. *The Statistician* **42**, 399–404.
- ISAACS, G. L., CHRIST, D. E., NOVICK, M. R., AND JACKSON, P. H. (1974) *Tables for Bayesian Statisticians*. University of Iowa press, Iowa.
- JOSEPH, L., WOLFSON, D. B., AND DU BERGER, R. (1995) Sample Size Calculations for Binomial Proportions via a Highest Posterior Density Intervals. *Statistics in Medicine* **44**, 143–154.
- JOSEPH, L., DU BERGER, R., AND BÉLISLE, P. (1997) Bayesian and Mixed Bayesian/Likelihood Criteria for Sample Size Determination. *Statistics in Medicine* **16**, 769–781.
- JOSEPH, L., AND BÉLISLE, P. (1997) Bayesian Sample Size Determination for Normal Means and Differences Between Normal Means. *The Statisticians* **46**, 209–226.
- KASS, R. E., AND RAFTERY, A. E. (1995) Bayes Factors. *J. Am. Statist. Ass.* **90**, 773–795.
- KOOPMANS, L., AND QUALLS, C. (1971) Fixed Length Confidence Intervals for Parameters of the Normal Distribution Based on Two-Stage Sampling Procedure. *Rocky Mountain J. Math.* **1**, 578–602.
- KRAEMER, H. C. AND THIEMANN, S. (1987) *How Many Subjects?: Statistical Power Analysis in Research*. Sage, Newbury Park.
- LACHIN, J. M. (1981) Introduction to Sample Size Determination and Power analysis for Clinical Trials. *Controlled Clinical Trials* **2**, 93–113.
- LEMESHOW, S., HOMSER, JR., D. W., AND STEWART, J. P. (1981) A Comparison of Sample Size Determination Methods in the Two Group Trial where the Underlying disease is rare. *Commun Statist Simula Computa, Bio* (**5**), 437–449.
- LEMESHOW, S., HOMSER, JR., D. W., KLAR, J., AND LWANGA, S. K. (1990) *Adequacy of Sample Size in Health Studies*. John Wiley and Sons, Chichester.
- LINDLEY, D. V. (1997) The Choice of Sample Size. *The Statistician* **46**, 129–138.

- LIPSEY, M. W. (1990) . *Design Sensitivity, Statistical Power for Experimental Research* Sage, Newbury Park.
- LIU, W. (1997) On Some Sample Size Formulae for Controlling both Size and Power in Clinical Trial. *The Statistician* **46**, 239–251.
- LLOYD, J. AND RAVEN, A. (1994) *Handbook of Clinical Research*. Churchill Medical Communications, London.
- MILLER, A. J. (1990) *Subject Selection in Regression*. Chapman and Hall, London.
- MORRIS, W.T. (1968) *Management Science: A Bayesian Introduction*. Prentice–Hall, Englewood Cliff, NJ.
- MOSHMAN, J. (1958) A Method for Selecting the Size of the Initial Sample Size in the Stein’s Two–Sample Procedure. *Ann Math. Statist.* **29**, 1271–1275.
- MULLER, K. E., LAVANGE, L. M., RAMEY, S. L., AND RAMEY, C. T. (1992) Power Calculations for the General Linear Model including Repeated Measure Analysis. *J. Am. Statist. Ass.* **87**, 1209–1226.
- MURTHY, M. N. (1967) *Sampling Theory and Methods*. Statistical Publishing, Calcutta .
- O’HAGAN, A. (1994) *Kendall’s Advanced Theory of Statistics Volume 2B Bayesian Statistics*. Edward Arnold.
- PEZESHK, H. AND GITTINS, J. C. (1999) Sample Size Determination in Clinical Trials. *Student* **3**, No.1, 19–26.
- PEZESHK, H. AND GITTINS, J. C. (1999) A Behavioural Bayesian Method for Sample Size Determination in Clinical Trials. *International Workshop on Objective Bayesian Methodology, Valencia; Spain*, <http://www.ox.ac.uk/~pezeshk/valec3.pdf>, pp 1–10.
- PHAM–GIA, T. AND TURKKAN, N. (1992) Sample Size Determination in Bayesian Analysis. *The Statistician* **41**, 389–392.
- RAIFFA, H. AND SCHLAIFER, R. (1961) *Applied Statistical Decision theory*. Division of Research Harvard Business.
- RAO, R. (1973)’ *Linear Statistical Inference and its Applications*, 2nd edn. John Wiley and Sons, New York.

- SADOOGHI-ALVANDI, M. (1986) The Choice of Subsample Size in Two-Stage Sampling. *J. Amer. Statist. Assoc.* **81**, 555–558.
- SEELBINDER, B. M. (1953) On Stein's Two-Stage Sampling Scheme. *Ann. Math. Statist.* **24**, 640–649.
- SENN, S. J. (1995) A Personal View of Some Controversies in Allocating Treatment to Patients in Clinical Trials. *Statistics in Medicine* **14**, 2661–2674.
- SENN, S. J. (1997) *Statistical Issues in Drug Development*. John Wiley and Sons, Chichester.
- SIGNORINI, D. F. (1991) Sample Size for Poisson Regression. *Biometrika* **78**, 446–450.
- SINGER, J. (1997) Estimating Sample Size for Continuous Outcomes, Comparing more than two Parallel Groups with Unequal Sizes. *Statistics in Medicine* **16**, 2805–2811.
- SISON, C. P. AND GLAZ, J. (1995) Simultaneous Confidence Intervals and Sample Size Determination for Multinomial Proportions. *J. Am. Statist. Ass.* **90**, 366–369.
- SMITH, T. M. F. (1976) The Foundation of Survey Sampling: A Review. *J. R. Statist. Soc. A* **139**, 183–195.
- SPIEGELHALTER, D. J. AND FREEDMAN, L. S. (1986) A Predictive Approach to Selecting the Size of a Clinical Trial, Based on Subjective Clinical Opinion. *Statistics in Medicine* **5**, 1–13.
- SPIEGELHALTER, D. J., FREEMAN, L. S. AND PARMAR, M. K. B. (1994) Bayesian Approaches to Randomized Trials. *J. R. Statist. Soc. A* **157**, 357–416.
- STALLARD, N. (1998) Sample Size Determination for phase II Clinical Trials based on Bayesian Decision Theory. *Biometrics* **54**, 279–294.
- STEIN, C. (1945) A two-Sample Test for a Linear Hypothesis where Power is Independent of Variance. *Ann. Math. Statist.* **16**, 243–258.
- THE EUROPEAN AGENCY FOR THE EVALUATION OF MEDICAL PRODUCTS (1996) ICH Topic E8 General Considerations for Clinical Trials. *Note for Guidance on General Considerations for Clinical Trials*. CPMP/ICH/291/95.



- THIGPEN, C. C. (1987) A Sample Size Problem in Linear Regression. *Am. Statistn.* **41**, 214–215.
- THOMPSON, S. K. (1987) Sample Size for Estimating Multinomial Proportions.. *Am. Statistn.* **41**, 42–46.
- TORTORA, R. D. (1978) A Note on Sample Size Estimation for Multinomial Proportions. *Am. Statist.* **32**, 100–101.
- WEISS, R. (1997) Bayesian Sample Size Calculations for Hypothesis Testing. *The Statistician* **46**, 185–191.
- WHITEHEAD, J. (1997) *The Design and Analysis of Sequential Clinical Trials*. John Wiley and Sons, Chichester.
- WHITTMORE, A. S. (1981) Sample Size for Logistic Regression with Small Response Probability. *J. Am. Statist. Ass.* **76**, 27–32.
- WOLFRAM, S. (1991) *Mathematica: a system for doing mathematics by computer*. Addison Wesley, Redwood City.