Statistical Analysis for Two-stage Adaptive Designs with Different Study Endpoints
關於具有不同目標變量的兩階段自適應設計的統計分析

Submitted to
Department of Management Science
管理科學系
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
哲學博士學位

by

Lu Qingshu
盧清樹

June 2009
二零零九年六月
Abstract

In clinical research, adaptive trial design has attracted much attention in recent years due to its flexibility in modifying some aspects of an on-going trial. Without undermining the validity and integrity of the trial, modifications of a clinical trial based on accrued information are necessary to improve efficiency of drug development and achieve ethical gain. Examples of modifications include sample size re-calculation, change in inclusion/exclusion criteria, adjustment of study dose, early termination of the trial, and modification of statistical hypotheses. In some cases, the study endpoints in different stages of a clinical trial may be different due to long treatment duration or some other reasons. For example, a biomarker (or a surrogate endpoint) may be adopted at the first stage, and a survival endpoint is considered at the second stage when the treatment duration is too long. The change in study endpoint offers a challenge to combine data from both stages for a valid final analysis at the end of the trial.

In this study, statistical methods utilizing data collected from both stages of a two-stage adaptive design are proposed for statistical analysis at the end of the trial, assuming that there is a well established relationship between the two different study endpoints. Three types of data from different study endpoints, i.e., continuous data, event data, and survival data, are considered in this study. For each type of data, method is proposed for combining data of two different study endpoints. Especially, for continuous data, we assume a functional relationship between two study endpoints, by which the “predicted” values of the primary study endpoint are obtained from data collected at the first stage. Those “predicted” data as well as the observed data of the primary study endpoint at the second stage are utilized in the final analysis. For illustration, data of both stages from two normal populations are included to assess the population mean of the primary endpoint using the Graybill-Deal estimator, assuming that a linear relationship is established between the two study endpoints. For the event data, we assume occurrence of an event of interest is determined by an underlying lifetime distribution, by which the data observed from the two stages with different durations are included into the likelihood function. For the time-to-event data, the analysis is straightforward when the study endpoints are different in the sense of study duration.
For the three types of data, tests for equality and equivalence between two treatments, test for superiority of the test treatment over the control treatment, and test for non-inferiority of the test treatment against the control are considered. Sample size calculations based on the proposed tests are addressed to achieve a pre-specific power. Sample size allocations at the two stages and between two treatments are also discussed. Simulations are conducted to investigate performance of the tests, including the type I error rate and power. Superiority of inclusion of data from both stages is shown by theoretical and numerical results.

Keywords: Adaptive design, Study endpoint, Biomarker, Sample size calculation, Graybill-Deal estimator, Equality, Superiority/Non-inferiority, Equivalence, Event data, Time-to-event data, Cox proportional hazards model, Censoring.
# TABLE OF CONTENTS

Abstract ........................................................................................................................i

Acknowledgement ................................................................................................... iii

List of Figures ........................................................................................................... vii

List of Tables ........................................................................................................... viii

1 Introduction ..........................................................................................................1

1.1 Adaptive Designs .................................................................................................1

1.2 Different Adaptations of Trial .............................................................................3

1.3 Aims and Scope of the Study ..............................................................................5

1.3.1 Methods for Combining Data from Different Endpoints .........................6

1.3.2 Hypothesis Testing ...................................................................................... 7

1.3.3 Sample Size Calculation............................................................................ 9

1.4 Layout of the Dissertation ..................................................................................11

2 Different Continuous Study Endpoints .............................................................12

2.1 Continuous Data ...............................................................................................12

2.2 GD Estimators ....................................................................................................13

2.3 Estimation of Treatment Mean $\mu$ .......................................................................15

2.4 Comparisons between Two Treatments ..........................................................16

2.4.1 Test for Equality .................................................................................... 17

2.4.2 Test for Superiority ............................................................................... 18

2.4.3 Test for Non-Inferiority ......................................................................... 18

2.4.4 Test for Equivalence ........................................................................... 19

2.5 Sample Size Determination ........................................................................... 19

2.6 Numerical Studies ............................................................................................ 22

2.6.1 Test of Treatment Mean $\mu$ .................................................................. 22

2.6.2 Test for Equality of Two Treatment Means .......................................... 24

2.7 Discussions ....................................................................................................... 26

3 Event Data with Different Durations from Underlying Exponential Distributions .........................................................................................................................30

3.1 Introduction .................................................................................................... 30

3.2 Model Formulation and Estimation ................................................................ 30

3.3 Hypothesis Testing and Sample Size Calculation ..........................................32

3.3.1 Test for Equality .................................................................................. 32

3.3.2 Test for Superiority ........................................................................... 33

3.3.3 Test for Non-Inferiority .................................................................... 34

3.3.4 Test for Equivalence ....................................................................... 34

3.4 Numerical Studies .......................................................................................... 35

3.4.1 Testing of One Treatment ................................................................... 36

3.4.2 Comparison between Two Treatments .................................................. 39

3.5 Concluding Remarks ...................................................................................... 44

4 Event Data with Different Durations from Underlying Weibull Distributions .................................47

4.1 Extension to the Weibull Distribution ................................................................47

4.2 Model Formulation and Estimation ................................................................47
4.3 Hypothesis Testing ........................................................................................................ 50
  4.3.1 Test for Equality .................................................................................................... 50
  4.3.2 Test for Superiority ............................................................................................. 51
  4.3.3 Test for Non-Inferiority ....................................................................................... 51
  4.3.4 Test for Equivalence .......................................................................................... 52
4.4 Sample Size Calculation ............................................................................................. 52
4.5 Bootstrap Method ....................................................................................................... 54
4.6 Numerical Studies .................................................................................................... 57
  4.6.1 Testing of One Treatment ................................................................................... 57
  4.6.2 Comparison between Two Treatments ............................................................... 66

5 Survival Data with Different Durations from the Exponential Distributions ................. 72
  5.1 Survival Data ............................................................................................................ 72
  5.2 Model Formulation and Estimation ........................................................................ 72
  5.3 Comparison between Two Treatments ..................................................................... 74
    5.3.1 Test for Equality ............................................................................................... 74
    5.3.2 Test for Superiority .......................................................................................... 75
    5.3.3 Test for Non-Inferiority .................................................................................. 76
    5.3.4 Test for Equivalence ....................................................................................... 76
  5.4 Numerical Studies ................................................................................................... 77
    5.4.1 Testing of One Treatment ............................................................................... 77
    5.4.2 Comparison between Two Treatments ............................................................. 81
  5.5 Concluding Remarks ............................................................................................... 83

6 Survival Data with Different Durations from the Weibull Distributions ............. 86
  6.1 Weibull Model and Estimation ............................................................................... 86
  6.2 Hypothesis Testing .................................................................................................. 89
    6.2.1 Test for Equality ............................................................................................. 89
    6.2.2 Test for Superiority ......................................................................................... 90
    6.2.3 Test for Non-Inferiority .................................................................................. 90
    6.2.4 Test for Equivalence ...................................................................................... 91
  6.3 Sample Size Calculation .......................................................................................... 92
  6.4 Numerical Studies .................................................................................................. 93
    6.4.1 Testing of One Treatment ............................................................................... 94
    6.4.2 Comparison between Two Treatments ............................................................. 97

7 Survival Data from Different Durations with Allowance for Non-uniform... 101
  7.1 Non-uniform Entry of Subjects and Censoring ....................................................... 101
  7.2 Model Formulation and Estimation ......................................................................... 101
  7.3 Hypothesis Testing .................................................................................................. 104
    7.3.1 Test for Equality ............................................................................................. 104
    7.3.2 Tests for Superiority, Non-Inferiority and Equivalence ................................. 105
  7.4 Sample Size Calculation ......................................................................................... 105
  7.5 Numerical Studies .................................................................................................. 107

8 Survival Data with Different Durations under the Cox Proportional Hazards Model .................... 112
8.1 The Cox Proportional Hazards Model ................................................................. 112
8.2 Hypothesis Testing ............................................................................................. 113
  8.2.1 Test for Equality ....................................................................................... 113
  8.2.2 Test for Superiority/ Non-inferiority ......................................................... 116
  8.2.3 Test for Equivalence ............................................................................... 117
8.3 Numerical Studies ............................................................................................. 118
  8.3.1 Test for Equality ....................................................................................... 118
  8.3.2 Test for Superiority ............................................................................... 119
8.4 Concluding remarks .......................................................................................... 121

9 Discussions ........................................................................................................ 123
References ............................................................................................................... 126
List of Figures

Figure 4.1 Optimal $c$ as $L/\lambda_r$ changes..............................................64

Figure 4.2 Range of $c$ where sample size is not larger than 120% of the minimum sample size .......................................................... 65
List of Tables

Table 2.1 Sample sizes for different values of $\rho$ ........................................ 21
Table 2.2 $N/M$ for different values of $\rho$ and $r$ ........................................ 24
Table 2.3 $N$ for test of the common mean ........................................ 25
Table 2.4 The total sample size $N_T$ ........................................ 27

Table 3.1 Total sample size $N$ for test in (3.12) ........................................ 37
Table 3.2 Type I error rate for test in (3.12) ........................................ 38
Table 3.3 Power for test in (3.12) ........................................ 39
Table 3.4 Type I error rates of the three tests ........................................ 42
Table 3.5 Powers of the three tests ........................................ 42
Table 3.6 Total sample size $N_T$ for equality testing ........................................ 45
Table 3.7 Power for sample size given in Table 3.6 ........................................ 46

Table 4.1 Minimum sample size $N^*$ and optimal duration $L^*$ ............... 58
Table 4.2 Total sample size $N$ based on the normal approximation ............... 59
Table 4.3 Comparison of Type I error rate ........................................ 61
Table 4.4 Sample size $N$ based on the bootstrap method ........................................ 62
Table 4.5 Range of optimal $c$ with $\beta_1 \leq 1$ and $L = 1$ .............................. 63
Table 4.6 Range of optimal $c$ with $\beta_1 > 1$ and $L = 1$ .............................. 63
Table 4.7 Minimum sample size $N_T^*$ and optimal duration $L^*$ ............... 68
Table 4.8 Total sample size $N_T$ based on the normal approximation ............... 69
Table 4.9 Powers of test base on the normal approximation with sample size given in Table 4.8 ........................................ 70
Table 4.10 Total sample size $N_T$ based on the bootstrap method ............... 71

Table 5.1 Total sample size $N$ ........................................ 79
Table 5.2 Type I error rate for sample size given in Table 5.1 ....................... 80
Table 5.3 Power for sample size given in Table 5.1 ........................................ 81
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4</td>
<td>Total sample size $N_T$ for equality testing</td>
<td>84</td>
</tr>
<tr>
<td>5.5</td>
<td>Powers of test with sample size given in Table 5.4</td>
<td>85</td>
</tr>
<tr>
<td>6.1</td>
<td>Sample size $N$</td>
<td>95</td>
</tr>
<tr>
<td>6.2</td>
<td>Type I error rate with sample sizes in Table 6.1</td>
<td>96</td>
</tr>
<tr>
<td>6.3</td>
<td>Powers of test with sample sizes in Table 6.1</td>
<td>97</td>
</tr>
<tr>
<td>6.4</td>
<td>Total sample size $N_T$ for equality testing</td>
<td>99</td>
</tr>
<tr>
<td>6.5</td>
<td>Powers of test with sample size in Table 6.4</td>
<td>100</td>
</tr>
<tr>
<td>7.1</td>
<td>Total sample size $N_T$ for equality testing</td>
<td>110</td>
</tr>
<tr>
<td>7.2</td>
<td>Powers of test for equality testing with sample size in Table 7.1</td>
<td>111</td>
</tr>
<tr>
<td>8.1</td>
<td>A comparison between $\bar{n}_i$ and $n_i$</td>
<td>119</td>
</tr>
<tr>
<td>8.2</td>
<td>$n_i$ for superiority testing</td>
<td>120</td>
</tr>
<tr>
<td>8.3</td>
<td>Power of test for superiority testing</td>
<td>121</td>
</tr>
</tbody>
</table>