REPUBLIC OF TURKEY
HACETTEPE UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

THE SAMPLE SIZE CALCULATION IN CLINICAL TRIALS AND COMPARISONS WITH CLASSICAL APPROACHES

Ebru ÖZTÜRK

Program of Biostatistics
MASTER OF SCIENCE

ANKARA
2018

REPUBLIC OF TURKEY
HACETTEPE UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

THE SAMPLE SIZE CALCULATION IN CLINICAL TRIALS AND COMPARISONS WITH CLASSICAL APPROACHES

Ebru ÖZTÜRK

Program of Biostatistics
MASTER OF SCIENCE

ADVISOR
Prof. Ergun KARAAĞAOĞLU
CO-ADVISOR
Assit. Prof. Nimet Anıl DOLGUN

ANKARA
2018

The Sample Size Calculation in Clinical Trials and Comparisons with Classical Approaches

Ebru ÖZTÜRK

This study has been approved and accepted as a Master/PhD dissertation in the program of "Biostatistics" by the examining committee, whose members are listed below, on *03.01.2018*.

Chairman of the Committee :	Prof. C. Reha ALPAR Hacettepe University
Advisor of the Dissertation :	Prof. A. Ergun KARAAĞAOĞLU Hacettepe University
Member :	Prof. Mehtap AKÇİL OK Başkent University
Member :	Assoc. Prof. Jale KARAKAYA KARABULUT Hacettepe University
Member :	Assit. Prof. Sevilay KARAHAN Hacettepe University

This dissertation has been approved by the committee above in conformity to the regulations and by laws of Hacettepe University Graduate Programs.

10 Ocak 2018

Prof. Diclehan ORHAN, MD, PhD

Institute Director

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini üniversiteye teslim etmeyi taahhüt ederim.

o Tezimin/Raporumun 03.01.2020 tarihine kadar erişime açılmasını ve fotokopi alınmasını (İç kapak, özet, İçindekiler ve Kaynakça hariç) istemiyorum. (Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir, kaynak gösterilmek şartıyla bir kısmı veya tamamının fotokopisi alınabilir)

11./1./2018

Ebru ÖZTÜRK

ETHICAL DECLARATION

In this thesis study, I declare that all the information and documents have been obtained in the base of the academic rules and all audio-visual and written information and results have been presented according to the rules of scientific ethics. I did not do any distortion in data set. In case of using other works, related studies have been fully cited in accordance with the scientific standards. I also declare that my thesis study is original except cited references. It was produced by myself in consultation with supervisor (Prof. Ergun KARAAĞAOĞLU) and written according to the rules of thesis writing of Hacettepe University Institute of Health Sciences.

4./1./2018

Ebru ÖZTÜRK

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my advisor Prof. Ergun Karaağaoğlu for support, guidance, motivation and patience during thesis study. It has always been an honour for me to be his student.

I would like to present my grateful thanks to my examining committee members: Prof. Reha Alpar, Prof. Mehtap Akçil Ok, Assoc. Prof. Jale Karakaya and Assist. Prof. Sevilay Karahan for their valuable time to review my study.

I would also like to acknowledge my dear fellow, Elif Akça with my best wishes and hopes to be together in the future

I would like to thank Duygu Aydın Haklı, Merve Başol and Dinçer Göksülük for their nice friendship and kind supports. Special thanks to Dinçer Göksülük for sharing LATEX template with me. I would thank to all members of Hacettepe University Department of Biostatistics.

My warm and sincere thanks to my best friend, Öznur Demirel for her warm friendship, inspirational soul and supporting my dreams.

Finally, I would like to represent to my deepest thanks to my dear family for their unconditional love, support and patience. Without them, I could not success to finish this thesis.

ABSTRACT

Öztürk, E. Sample Size Calculation in Clinical Trials and Comparisons with Classical Approach. Hacettepe University Institute of Health Sciences, MSc. Thesis in Biostatistics, Ankara, 2018. Clinical trials are well-planned studies. One of the earlier steps in clinical trials is the determination of sample size. The question "how many subjects should be used?" must be answered carefully by considering many important aspects of the study. Some of the clinical trials might be too expensive. Besides, in some of them, finding subjects may be difficult to include clinical trials. For such these reasons, it is not efficient including either too few or too many subjects in clinical trials. Therefore, sample size calculation is an important issue in clinical trials due to ethical, economic and scientific reasons. There are several factors that affect the sample size such as study design, trial objectives or clinical important difference. In this thesis, we give an overview of sample size calculation in clinical trials. Parallel group and cross-over study designs are taken into account. We also considered equality, superiority, non-inferiority and equivalence trials for two samples. First, we gave proofs of sample size calculations with both known and population variance. Then, we show numeric examples to clarify sample size calculation. Additionally, we share how these calculations are carried out RStudio. We also create simulation scenarios under different distributions, trial objectives, sample size with the clinical important difference and specified effect sizes to compare observed power. We show that the observed power is highest in non-inferiority trials compared to superiority and equality trials based on same clinical important difference, Type I error, study design and sample size. The observed power is higher in cross-over design compared to parallel group design with same clinical important difference, Type I error, trial objective and sample size. The responses are created under different distributions, however; there is no considerable effect of different distributions on observed power.

Key Words: Sample size calculations, clinical trials, observed power, parallel group design, cross-over design.

ÖZET

Öztürk, E. Klinik Denemelerde Örneklem BüyüklüğüHesaplamaları ve Klasik Hesaplama Yöntemleri ile Karşılaştırılması. Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Programı Yüksek Lisans Tezi, Ankara, 2018. Klinik araştırmalar iyi planlanması gereken çalışmalardır. Çalışmanın başında, örneklem büyüklüğünün hesaplanması çalışmanın önemli adımlarındandır. "Çalışmada en az kaç kişi bulunmalıdır?" sorusu çalışmanın önemli yönleri ele alınarak dikkatlice cevaplandırılmalıdır. Bazı klinik araştırmaların maliyeti çok yüksek olabilir. Bununla birlikte, bazı klinik araştırmalarda ise çalışmaya katılacak kişilerin bulunması zor olabilir. Bu ve buna benzer nedenlerden dolayı klinik denemelerde örneklem büyüklüğü çok az ya da çok fazla olmamalıdır. Örneklem büyüklüğünün belirlenmesi ekonomik, etik ve bilimsel nedenlerle önemli bir konudur. Örneklem büyüklüğünü deneme amacı, deseni ve etki büyüklüğü gibi çeşitli faktörler etkilemektedir. Bu tezde, klinik araştırmalarda örneklem büyüklüğü hesaplamasına genel bir bakış sunulmuştur. Klinik denemelerde birincil değişkene ait ölçümlerin sürekli ve iki yanıtlı veri tipinde olması durumu ele alınmıştır. Klinik deneme düzeni olarak paralel düzen ve çapraz geçişli düzen üzerinde durulmuştur. Bilinen ve bilinmeyen evren varyansları ile örneklem büyüklüğünün nasıl hesaplanacağı ispatlar ve sayısal örneklerle gösterilmiştir. Bunlara ek olarak, bu hesaplamaların RStudio programı kullanılarak nasıl yapıldığına değinilmiştir. Farklı dağılımlar, deneme amaçları ve örneklem büyüklükleri,gözlenen güç ve belirli etki büyüklükleri ile karşılaştırmak için simülasyon senaryoları oluşturulmuştur. Aşağı olmayış denemelerinde gözlenen güç üstünlük ve eşitlik denemelerine göre en yüksektir. Aynı Tip I hata, deneme amacı ve örneklem büyüklüğüne sahip çapraz geçişli denemelerde gözlenen güç paralel grup denemelerine göre daha yüksektir. Farklı dağılımların gözlenen güç üzerinde önemli bir etkisi bulunamamıştır.

Anahtar Kelimeler: Örneklem büyüklüğü hesaplaması, klinik deneme, gözlenen güç, paralel grup deneme deseni, çapraz geçişli deneme

TABLE OF CONTENTS

FIGURES

TABLES

1 INTRODUCTION

Clinical trial is the study conducted on people prospectively that the participants receive the treatments (initiatives) determined within the framework of a protocol. These treatments or initiatives might be a medical product such as a drug or device, or a process (diet, etc.) that may change the behavior of the participant. A new medical approach is compared against a standard approach that already exists or against a placebo. Sometimes, two methods currently in use can be compared in some clinical trials.

When researchers design a clinical trial, first question is that how many subjects are needed in this trial. When a new product or approach is tested it is unknown whether it is beneficial, harmful, or different from existing alternatives (including placebo). For this reason, accurate determination of sample size in clinical trials is very important in order to ensure that people benefit from better treatment or are protected against worse treatment.

There are three main reasons underlying the determination of sample size which are economic reasons (1), ethical reasons (1) and scientific reasons (2), respectively.

- **Economic reasons**:An undersized study, the clinically significant results might not be obtained due to inadequate data. Therefore, the study might result in waste of resources. In spite of the fact that statistical significance could be achieved with a smaller sample size, a study with an oversized sample causes waste of resources by taking more subjects (1).

- **Ethical reasons**: An oversized study might be unethical as large number of subjects might be exposed to treatment that has unknown capability. For an undersized study, it may not be possible to notice clinically meaningful difference due to too small sample size. Therefore, an undersized study also might be unethical based on usage of subjects and other resources. In another words, subjects may be exposed to treatment with unknown efficacy (1).

- **Scientific reasons**: An undersized study, clinically important results might not be detected statistically through small sample size. On the other hand, an oversized study might enable to obtain statistically significant results without considering clinical importance (2).

In conclusion, the determination of the sample size in clinical trials is important in terms of the validity, reliability and ethical integrity of the study

design.

Several factors play a role in determining sample size in clinical trials. These are trial objective, primary end point, study design, effect size, variability in population, type I error, type II error and other factors such as drop out ratio (3).

In this study, we will discuss the case where the primary end points in the clinical trials are continuous and binary. Parallel group and cross-over study designs will be emphasized on hypothesis testing in clinical trials. **The purpose of this study is to highlight the importance of sample size calculation for clinical trials. Besides, figuring out how the sample size changes according to the trial objectives (equality, superiority, etc.) and study designs (parallel group and cross-over designs) with respect to known and unknown population variance. Moreover, simulation scenarios are created under different distributions, trial objectives, sample size with different clinically important difference and specified effect sizes that is worth detecting observed power.**

The plan of this thesis is as follows: In Chapter 2, general information of methods is given. In Chapter 3, material and methods are discussed. In Chapter 4, the results are demonstrated. Finally, the discussion and conclusion follow in Chapter 5 and Chapter 6.

2 GENERAL INFORMATION

In this chapter, the background information relating to the factors that affect sample size in clinical trials is presented.

2.1 Study Design

According to Good Clinical Practice Guidelines (GCP), the overall plan and design of a clinical trial should be clearly defined in the protocol (4). A well-organized protocol involves study design, trial objectives, subject inclusion and exclusion criteria, treatment schedule, trial endpoints and statistical methods. Study design and trial objectives are directly related with statistical methods (5). Therefore, determining study design is crucial in clinical trials. In this study, we point out parallel group and cross-over designs.

2.1.1 Parallel Group Design

Each subject is randomly assigned and receive only one treatment in parallel group design (5). Suppose one wants to compare a standard treatment with a new treatment within a parallel group design. To reach this aim, some of the subjects are assigned to only new treatment and rest of the them are assigned to only standard treatment. At the end of the trial, the results of the two independent groups are obtained for comparison. In parallel group design, more than two treatments can be compared such as new treatment, standard treatment and placebo. However, in this thesis we focus only on comparing two different treatments. In Figure 2.1, an example of two-group parallel design with one new treatment and one standard treatment is shown.

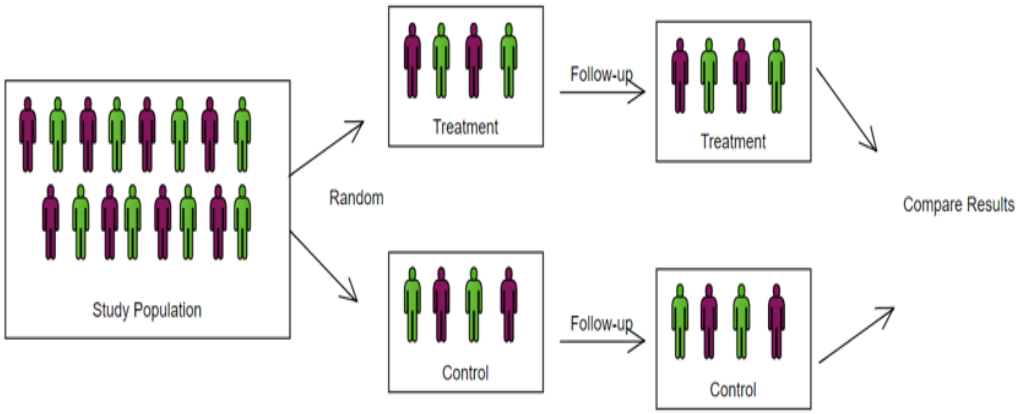

Figure 2.1. Parallel Group Design

2.1.2 Cross-over Design

Each subject receives each treatment in different periods in a cross-over design. In order to determine the order, according to which the treatments will be received by the subjects, treatment sequences are created. At the beginning of the study, subjects are randomly assigned to one of the treatment sequences. If comparing the effect of a new treatment with the one of a standard one by a cross-over design is of interest, some of the patients take the new treatment (sequence 1) while the others receive the standard treatment (sequence 2). After the washout period, which is defined to be the adequate length to erode the effect of treatments in the first period, the subjects under study will get the other treatment, they have not received yet, in the second period. In Figure 2.2, an illustration of cross-over design is given.

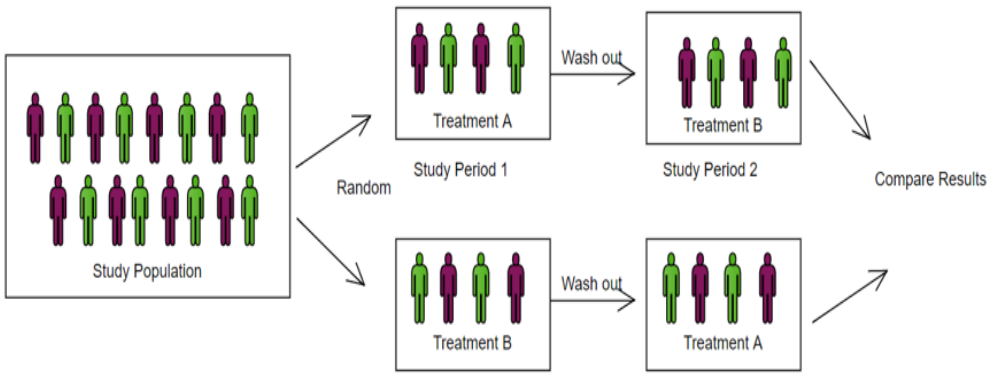

Figure 2.2. Cross-over Design

2.2 Trial Objective

For calculating the necessary sample size, deciding on the objective of the trial, that is related to the statistical hypothesis tests conducted in clinical trials, is another crucial point. The null and the alternative hypotheses are determined according to the trial objective. In this thesis, we focus on equality, superiority, non-inferiority and equivalence trial objectives.

2.2.1 Equality Trials

In equality trials, researchers investigate whether there is a difference between two different treatments or not (3). In this thesis, to illustrate two different treatments we used the terms "new treatment" and "standard treatment". Assume μ_N and μ_S represents the mean values of the new treatment and the standard treatment, respectively. The null and alternative hypotheses can be written:

$$H_0 : \mu_N = \mu_S \quad \text{versus} \quad H_1 : \mu_N \neq \mu_S. \tag{2.1}$$

The rejection of the null hypothesis implies that the new treatment is not equal to the standard one.

2.2.2 Superiority Trials

In clinical trials, we might point out that the effect of the new treatment is better than the standard one (3). The null and alternative hypotheses is shown as:

$$H_0 : \mu_N - \mu_S \leq \delta \quad \text{versus} \quad H_1 : \mu_N - \mu_S > \delta. \tag{2.2}$$

where δ represents the clinically important difference. If null hypothesis is rejected, then new treatment is superior to standard treatment. If δ is equal to zero, then this hypothesis named as statistical superiority.

2.2.3 Non-Inferiority Trials

In non-inferiority, trials we would like to investigate if the new treatment is not worse than the standard treatment (3). The null and alternative hypotheses can be written as:

$$H_0 : \mu_S - \mu_N \leq \delta \quad \text{versus} \quad H_1 : \mu_S - \mu_N > \delta. \tag{2.3}$$

If the null hypothesis is rejected, one should conclude that the effect of the new treatment is non-inferior compared to the standard one. To clarify this statement, difference between new treatment and standard treatment is less than δ. As a result, new treatment is not worse than the standard treatment.

2.2.4 Equivalence Trials

In some clinical trials, the aim is to show that there is no clinically meaningful difference between the two treatments which means that the two treatments under investigation are equivalent (3). Therefore, for testing equivalence, the hypotheses are written as:

$$H_0 : |\mu_N - \mu_S| \geq \delta \quad \text{versus} \quad H_1 : |\mu_N - \mu_S| < \delta \tag{2.4}$$

If the null hypothesis is rejected, there is no clinically important difference between the new treatment and the standard one which implies that the effects of these treatments are equivalent.

2.3 Primary End Point

In a clinical trial, the primary objective of the trial is assessed to test the hypothesis. An end point is set to indicate the efficacy and reliability of the trial (3). The sample size calculations rely on whether the primary endpoint variable is continuous, ordinal, binary or survival (time or probability). In this thesis, only continuous and binary variables are taken into account.

2.4 Effect Size

Deciding the effect size is another crucial step when determining the sample size. The main aim of calculation of the sample size is to satisfy an adequate power to reject the null hypothesis when the alternative hypothesis is true. For clinical trials, effect size is also known as clinically important difference. If this difference increases, then sample size would decrease.

2.5 Variability in Population

Variance estimation is one of the important steps of sample size calculation. The concerned variance is the variance of the variable of the primary end point. For variance estimation, retrospective data and previous studies are used (3). Variance is one of the components of sample size equations.

2.6 Type I and Type II Errors

In Table 2.1, decision table of hypothesis testing is given.

Table 2.1. Decision Table for Hypothesis Testing

Decision	Truth	
	H_0 is true	H_0 is not true
Fail to reject H_0	Correct decision	Type II Error (β)
Reject H_0	Type I Error (α)	Correct decision

Type I error (α) is the probability of rejecting H_0 accidentally which means that rejecting H_0 when H_0 is true. On the other hand, Type II error (β) is the probability of failing to reject H_0 accidentally which implies failing to reject H_0 when H_0 is not true. Power is the probability of finding the effect in case of there exist an effect in a study. Power (1-β) which is identified for an alternative hypothesis is rejecting H_0 when H_0 is not true. Power is inversely related with

Type II error. In hypothesis testing, we want to decrease the probability of Type I and Type II errors to increase the probability of correct decisions.

The type I error, type II error and power are represented as:

- α =P(type I error)=P(Reject H_0 | H_0 is true)
- β=P(type II error)=P(Fail to reject H_0 | H_0 is not true)
- Power=1-β=P(Reject H_0 | H_0 is not true)

3 MATERIAL and METHODS

In this chapter, the proof and calculations of sample size under different trial objectives and study designs are given with respect to whether population variance is known or not.

3.1 Parallel Group Design

Sample size calculations are given for continuous and binary primary end points with respect to different trial objectives with parallel group design.

3.1.1 Equality Trials

In equality trials, we investigate whether or not the effect of two treatments are same.

Continuous Primary End Point

Suppose x_{ij} represents the observed value of the j^{th} subject in the i^{th} treatment group where $j = 1, 2, 3, \ldots, n$ and $i = 1, 2$. Assume x_{ij} is normally independently distributed with μ_i and σ_i in the i^{th} group. $\bar{x}$ and S^2 represents the sample mean and the pooled sample variance, respectively.

$$\bar{x}_{i.} = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij} \quad and \quad S^2 = \frac{1}{n_1 + n_2 - 2}\sum_{i=1}^{2}\sum_{j=1}^{n_j}(x_{ij} - \bar{x}_{i.})^2 \tag{3.1}$$

Suppose $\epsilon = \mu_1 - \mu_2$ represents the true mean difference between new treatment(μ_1) and standard treatment (μ_2). As it is stated in Chapter 2.2.1, for equality trials ϵ is equal to 0. Therefore, hypotheses are constructed for equality trials as

$$H_0 : \epsilon = 0 \quad \text{and} \quad H_1 : \epsilon \neq 0. \tag{3.2}$$

Assume that the distributions of the response of the two treatments are as follows: $X_1 \sim \mathrm{N}(\mu_1, \sigma^2)$ and $X_2 \sim \mathrm{N}(\mu_2, \sigma^2)$. The sample means of two treatments are distributed normally based on Central Limit Theorem. Assume that sample sizes of the treatment groups are n_1 and n_2. Population variances are known and homogeneous for the both of the treatment groups presented with σ^2. Therefore, the distribution of the sample means are $\bar{X}_1 \sim N(\mu_1, \sigma^2/n_1)$ and $\bar{X}_2 \sim N(\mu_2, \sigma^2/n_2)$. As a result, the distribution of the difference between sample means can be writ-

ten as the following:

$$\bar{X}_1 - \bar{X}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}) \tag{3.3}$$

Under the null hypothesis (H_0:ϵ=0), the test statistic would be:

$$z_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.}}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.4}$$

According to the test statistic, if $|z_0| > z_{\alpha/2}$ reject H_0. Under the alternative hypothesis (H_1:$\epsilon \neq 0$), the test statistic would be represented as:

$$z_1 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.5}$$

If one can try to write the relationship between the test statistics under the null hypothesis and the alternative hypothesis, the following equation is obtained:

$$z_0 = z_1 + \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.6}$$

The type II error can be written with respect to test statistics z_0 and z_1.

$$\begin{aligned}
\beta &= P(|z_0| \leq z_{\alpha/2}) = P(-z_{\alpha/2} \leq z_0 \leq z_{\alpha/2}) \\
&= P(-z_{\alpha/2} \leq z_1 + \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \leq z_{\alpha/2}) \\
&= P(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \leq z_1 \leq z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) \\
&= \Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) - \Phi(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}})
\end{aligned} \tag{3.7}$$

ϕ represents the cumulative density function of normal distribution. As a result of Equation 3.7, the corresponding power (1-β) would be written as:

$$\begin{aligned}
1 - \beta &= 1 - [\Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) - \Phi(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}})] \\
&= 1 - \Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) + \Phi(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) \\
&= \Phi(\frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_{\alpha/2}) + \Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}})
\end{aligned} \tag{3.8}$$

The second term in Equation 3.8 is neglected as it's too small (6). Power would be found in Equation 3.9.

$$1-\beta \approx \Phi(\frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} - z_{\alpha/2}) \tag{3.9}$$

Therefore, the corresponding power can be written in Equation 3.10.

$$z_\beta = \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} - z_{\alpha/2} \tag{3.10}$$

Allocation ratio of two groups is demonstrated with r. Let say $n_1 = rn_2$. To find the sample size with 1-β, Equation 3.10 should be solved with respect to n_2.

$$n_2 = \frac{(z_\beta + z_{\alpha/2})^2(1+r)\sigma^2}{r\epsilon^2} \tag{3.11}$$

Therefore, sample size of clinical trial for one group is found as in Equation 3.11 (6).

If the population variance (σ^2) is unknown, one can use sample variance (S^2). S^2 leads to use of t distribution. Under the null hypothesis the test statistic will be:

$$t_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{3.12}$$

H_0 is rejected when $|t_0| > t_{\alpha/2,(n_1+n_2-2)}$. Under the alternative hypothesis, since ϵ does not equal to 0, t distribution turns out non-central t distribution. Non-central t distribution is general form of student t distribution. In this case, standard normal distribution (Z) and Chi-square distribution with degrees of freedom f (V) are written as $Z = \frac{\sqrt{(n_1+n_2-2)}(\bar{x}_{1.}-\bar{x}_{2.}-\mu_1-\mu_2)}{\sigma}$ and $V = \frac{(n_1+n_2-2)S^2}{\sigma^2}$. Moreover, Z and V are statistically independent. The non-central t distribution can be written in terms of Z and V distributions as:

$$T_{f,\theta} = \frac{Z+\theta}{\sqrt{V/f}} \tag{3.13}$$

θ and f symbolize the non-centrality parameter and the degrees of freedom of non-central t distribution, respectively. For equality trials, $\theta = |\epsilon|/\sigma$. If ϵ is added and dropped to the numerator and σ is divided by both the numerator and the denominator of the test statistic $\frac{\bar{x}_{1.}-\bar{x}_{2.}-(\mu_1-\mu_2)}{S/\sqrt{(n_1+n_2-2)}}$, the corresponding test statistic

would be similar to non-central t distribution.

$$\frac{\sqrt{(n_1+n_2-2)}(\bar{x}_{1.}-\bar{x}_{.2}-\mu_1-\mu_2-\epsilon)/\sigma+\sqrt{(n_1+n_2-2)}\epsilon/\sigma}{S/\sigma}$$
$$=\frac{Z+\theta}{\sqrt{V/(n_1+n_2-2)}}=T_{(n_1+n_2-2),\theta} \tag{3.14}$$

In Equation 3.14, we show obtaining non-central t distribution via standard normal distribution and chi square distribution (6). Under H_1,this statistic would follow a non-central t distribution with the non-centrality parameter of θ and the degrees of freedom f which are equal to $\frac{|\epsilon|}{\sigma}$ and n_1+n_2-2, respectively. If the cumulative distribution function of non-central t distribution is presented with $G_{f,\theta}(t)$ which is equal to $P(T_{f,\theta}\leq t)$. Therefore, the power for equality trials can be found in Equation 3.15.

$$\begin{aligned}1-\beta &= 1-P(|T_{f,\theta}|\leq t_{\alpha/2,(n_1+n_2-2)})\\ &= 1-P(-t_{\alpha/2,(n_1+n_2-2)}\leq T_{f,\theta}\leq t_{\alpha/2,(n_1+n_2-2)})\\ &= 1-T_{n_1+n_2-2}[t_{\alpha/2,(n_1+n_2-2)}|\theta]+T_{n_1+n_2-2}[-t_{\alpha/2,(n_1+n_2-2)}|\theta]\end{aligned} \tag{3.15}$$

Second term is neglected again. Power would be equal to in Equation 3.16 (6).

$$1-\beta=1-T_{n_1+n_2-2}[t_{\alpha/2,(n_1+n_2-2)}|\theta] \tag{3.16}$$

Chow and his colleagues (6) provide a table as for solving Equation 3.16. That table is given in Table 3.1.

Table 3.1. Smallest n with $T_{2r-2}[t_{\alpha/2,(2r-2)}|\sqrt{n}\theta/\sqrt{(1+1/r)}]$

	r=1				r=2			
	α=2.5%		α=5%		α=2.5%		α=5%	
	1-β		1-β		1-β		1-β	
θ	80%	90%	80%	90%	80%	90%	80%	90%
0.30	176	235	139	191	132	176	104	144
0.32	155	207	122	168	116	155	92	126
0.34	137	183	108	149	103	137	81	112
0.36	123	164	97	133	92	123	73	100
0.38	110	147	87	120	83	110	65	900
0.40	100	133	78	108	75	100	59	810
0.42	90	121	71	98	68	90	54	740
0.44	83	110	65	90	62	83	49	670
0.46	76	101	60	82	57	76	45	620
0.48	70	93	55	76	52	70	41	570
0.50	64	86	51	70	48	64	38	520
0.52	60	79	47	65	45	59	35	480
0.54	55	74	44	60	42	55	33	450
0.56	52	68	41	56	39	51	31	420
0.58	48	64	38	52	36	48	29	390
0.60	45	60	36	49	34	45	27	370
0.65	39	51	30	42	29	38	23	310
0.70	34	44	26	36	25	33	20	270
0.75	29	39	23	32	22	29	17	240
0.80	26	34	21	28	20	26	15	210
0.85	23	31	18	25	17	23	14	190
0.90	21	27	16	22	16	21	12	170
0.95	19	25	15	20	14	19	11	150
1.00	17	23	14	18	13	17	10	140
1.05	16	21	12	17	12	15	9	130
1.10	15	19	11	15	11	14	9	120
1.15	13	17	11	14	10	13	8	110
1.20	12	16	10	13	9	12	7	100
1.25	12	15	9	12	9	11	7	90
1.30	11	14	9	11	8	11	6	90
1.35	10	13	8	11	8	10	6	80
1.40	10	12	8	10	7	9	6	80
1.45	9	12	7	9	7	9	5	70
1.50	9	11	7	9	6	8	5	70

Binary Primary End Point

Suppose x_{ij} (3) represents the binary response j^{th} subject in the i^{th} treatment group where j=1,2,3,...,n and i=1,2. For a fixed i, x_{ij} are independent identically distributed with Bernoulli $P(x_{ij} = 1) = p_i$ represented as $X_{ij} \sim Ber(p_i)$ (3). Moreover, p_i is unknown, it is estimated by the observed proportion of each treatment.

$$\hat{p}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \tag{3.17}$$

The difference between the true proportion responses is represented with $\epsilon = p_1 - p_2$. For equality trials, hypotheses are written as $H_0 : \epsilon = 0$ and $H_1 : \epsilon \neq 0$. In this study, we only consider calculations based on normal approximation to binomial.Therefore, under the null hypothesis, the test statistic would be:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)/n_2}} \tag{3.18}$$

Reject H_0 if $|z_0| > z_{\alpha/2}$. Moreover, under the alternative hypothesis which indicates $\epsilon \neq 0$, power can be found from Equation 3.19

$$1 - \beta \approx \Phi(\frac{|\epsilon|}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} - z_{\alpha/2}) \tag{3.19}$$

To reach a power of 1-β, the following equation should be solved (6).

$$z_\beta = \frac{|\epsilon|}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} \tag{3.20}$$

If we consider n_1=rn_2, n_2 is found in Equation 3.21 (6):

$$n_2 = \frac{(z_\beta + z_{\alpha/2})^2}{\epsilon^2}[\frac{p_1(1-p_1)}{r} + p_2(1-p_2)] \tag{3.21}$$

3.1.2 Superiority Trials

In superiority trials, the aim is to show that the effect of the one of the treatment is superior to other one.

Continuous Primary End Point

Hypotheses of superiority trials are:

$$H_0 : \epsilon \leq \delta \quad \text{versus} \quad H_1 : \epsilon > \delta. \tag{3.22}$$

where $\delta > 0$. The hypothesis test becomes one-sided as seen in Equation 3.22. Under the null hypothesis, the test statistic would be:

$$z_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.23}$$

According to this test statistic, if $z_0 > z_\alpha$ reject H_0. Under the alternative hypothesis the test statistic is ($\epsilon > \delta$):

$$z_1 = \frac{(\bar{x}_{1.} - \bar{x}_{2.}) - (\epsilon - \delta)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.24}$$

If z_0 is written with respect to z_1, the following equation will appear:

$$z_0 = z_1 + \frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.25}$$

The probability of type II error and the power function can be written as by using Equation 3.25:

$$\begin{aligned} \beta = P(z_0 \leq z_\alpha) &= P(z_1 + \frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \leq z_\alpha) \\ &= P(z_1 \leq z_\alpha - \frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) \\ &= \Phi(z_\alpha - \frac{\epsilon}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) \end{aligned} \tag{3.26}$$

1-β will be equal to:

$$\begin{aligned} 1 - \beta &= 1 - \Phi(z_\alpha - \frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}}) \\ &= \Phi(\frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_\alpha) \end{aligned} \tag{3.27}$$

Suppose $n_1 = rn_2$, then the minimum sample size is written in Equation 3.28 (6).

$$n_2 = \frac{(z_\alpha + z_\beta)^2(1+r)\sigma^2}{r(\epsilon-\delta)^2} \tag{3.28}$$

If σ^2 is unknown, then non-central t distribution should be used for sample size calculation.

$$t_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \delta}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{3.29}$$

Reject H_0 if $t_0 > t_{\alpha,(n_1+n_2-2)}$. Under H_1, this statistic would follow a non-central t distribution with the non-centrality parameter of θ and degrees of freedom f which are equal to $\frac{(\epsilon-\delta)}{\sigma}$ and n_1+n_2-2, respectively. Therefore, power can be found for equality trials in Equation 3.30.

$$\begin{aligned} 1-\beta &= 1 - P(T_{f,\theta} \leq t_{\alpha,(n_1+n_2-2)}) \\ &= 1 - P(T_{f,\theta} \leq t_{\alpha,(n_1+n_2-2)}) \\ &= 1 - T_{n_1+n_2-2}[t_{\alpha,(n_1+n_2-2)}|\theta] \end{aligned} \tag{3.30}$$

Table 3.1 can be used for this situation (6).

Binary Primary End Point

For superiority trials, hypotheses are written as $H_0 : \epsilon \leq \delta$ and $H_1 : \epsilon > \delta$. Therefore, under the null hypothesis, the test statistic would be:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2 - (\epsilon-\delta)}{\sqrt{\hat{p}_1(1-\hat{p}_1)/n_1 + \hat{p}_2(1-\hat{p}_2)}/n_2} \tag{3.31}$$

Reject H_0 if $z_0 > z_\alpha$. Moreover, the under the alternative hypothesis which indicates $\epsilon > \delta$, power can be found from Equation 3.32 as:

$$1-\beta \approx \Phi\left(\frac{\epsilon-\delta}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} - z_\alpha\right) \tag{3.32}$$

To reach a power of 1-β, the following equation should be solved with respect to n_2 (6).

$$z_\beta = \frac{\epsilon-\delta}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} \tag{3.33}$$

If we consider n_1=rn_2, n_2 is found by Equation 3.34 (6):

$$n_2 = \frac{(z_\beta + z_\alpha)^2}{(\epsilon - \delta)^2}[\frac{p_1 \times (1 - p_1)}{r} + p_2 \times (1 - p_2)] \tag{3.34}$$

3.1.3 Non-inferiority Trials

In non-inferiority trials, as it is stated in Chapter 2.2.3, that aim is to show the effect of the one of the treatment is non-inferior to the other one. The proof of sample size calculation is same with the superiority trials. $\epsilon = \mu_1 - \mu_2$ in non-inferiority trials.

Continuous Primary End Point

Hypotheses of non-inferiority trials are:

$$H_0 : \epsilon \leq \delta \quad \text{versus} \quad H_1 : \epsilon > \delta. \tag{3.35}$$

Suppose $n_1 = rn_2$, then the minimum sample size is written in Equation 3.36 (6):

$$n_2 = \frac{(z_\alpha + z_\beta)^2(1 + r)\sigma^2}{r(\epsilon - \delta)^2} \tag{3.36}$$

If σ^2 is unknown, then non-central t distribution should be used for sample size calculation. Power can be found for equality trials by Equation 3.37.

$$1 - \beta = 1 - T_{n_1+n_2-2}[t_{\alpha,(n_1+n_2-2)}|\theta] \tag{3.37}$$

Table 3.1 can be used for this situation (6).

Binary Primary End Point

For non-inferiority trials, hypotheses are written as $H_0 : \epsilon \leq \delta$ and $H_1 : \epsilon > \delta$. The sample size is found by in Equation 3.38.

$$n_2 = \frac{(z_\beta + z_\alpha)^2}{(\epsilon - \delta)^2}[\frac{p_1 \times (1 - p_1)}{r} + p_2 \times (1 - p_2)] \tag{3.38}$$

3.1.4 Equivalence Trials

In equivalence trials, it is investigated that two treatments neither superior nor non-inferior to to each other. Both treatments have equivalent effects.

Continuous Primary End Point

The hypotheses for equivalence can be written as:

$$H_0 : |\epsilon| \geq \delta \quad \text{vs.} \quad H_1 : |\epsilon| < \delta. \tag{3.39}$$

The null hypothesis of test for equivalence leads to two the null hypotheses that are

$$\begin{aligned} &H_{01} : \epsilon \leq -\delta \quad \text{vs.} \quad H_{11} : \epsilon > -\delta \quad \text{and} \\ &H_{02} : \epsilon \geq \delta \quad \text{vs.} \quad H_{12} : \epsilon < \delta. \end{aligned} \tag{3.40}$$

According to the two the null hypotheses, the test statistics under the null hypotheses are indicated below:

$$z_{01} = \frac{\bar{x}_{1.} - \bar{x}_{2.} + \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \quad \text{and} \quad z_{02} = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.41}$$

H_{01} and H_{02} are rejected if $z_{01} > z_\alpha$ and $z_{02} < -z_\alpha$, respectively. Under the alternative hypotheses, the test statistics are:

$$z_{11} = \frac{\bar{x}_{1.} - \bar{x}_{2.} - (\epsilon + \delta)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \quad \text{and} \quad z_{12} = \frac{\bar{x}_{1.} - \bar{x}_{2.} - (\epsilon - \delta)}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} \tag{3.42}$$

Power of hypotheses testings are found based on two different the alternative hypothesis. Therefore, to reach the power with 1-β, β divided by β_1 and β_2 ($\beta_1 + \beta_2 = \beta$).

$$1 - \beta_1 \approx \Phi(\frac{\epsilon + \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_\alpha) 1 - \beta_2 \approx \Phi(\frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_\alpha) \tag{3.43}$$

By summing up type II errors, power of hypothesis testing is found as in Equation 3.44.

$$\begin{aligned} 2 - \beta_1 - \beta 2 &= \Phi(\frac{\epsilon + \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_\alpha) + \Phi(\frac{\epsilon - \delta}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_\alpha) \\ 2 - \beta_1 - \beta_2 &= 2\Phi(\frac{\delta - |\epsilon|}{\sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}} - z_\alpha) \end{aligned}$$

$$2-\beta = 2\Phi(\frac{\delta-|\epsilon|}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} - z_\alpha)$$
$$1-\frac{\beta}{2} = \Phi(\frac{\delta-|\epsilon|}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} - z_\alpha) \tag{3.44}$$

To find the minimum required sample size, the following equation should be solved with respect to sample size.

$$z_{\beta/2} = \frac{\delta-|\epsilon|}{\sqrt{\frac{\sigma^2}{n_1}+\frac{\sigma^2}{n_2}}} - z_\alpha \tag{3.45}$$

Lets say $n_1 = rn_2$. To find sample size with 1-β, Equation 3.45 should be solved with respect to n_2.

$$n_2 = \frac{(z_{\beta/2}+z_\alpha)^2(1+r)\sigma^2}{r(\delta-|\epsilon|)^2} \tag{3.46}$$

Similarly, if the population variance is unknown, then using non-central t distribution is applicable based on the sample variance.

$$t_{01} = \frac{\bar{x}_{1.}-\bar{x}_{2.}+\delta}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \quad \text{and} \quad t_{02} = \frac{\bar{x}_{1.}-\bar{x}_{2.}-\delta}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \tag{3.47}$$

Reject H_0 if $t_{01} > t_{\alpha,n_1+n_2-2}$ or $t_{02} < -t_{\alpha,n_1+n_2-2}$. Under the alternative hypotheses;

$$1-\beta = 1 - T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta-\epsilon}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}) - T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta+\epsilon}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}})$$
$$1-\beta = 2T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta-|\epsilon|}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}) - 1$$
$$2-\beta = 2T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta-|\epsilon|}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}})$$
$$1-\beta/2 = T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta-|\epsilon|}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}) \tag{3.48}$$

Sample size can be found by using Table 3.1 (6).

Binary End Point

The hypotheses for equivalence can be written as:

$$H_0 : |\epsilon| \geq \delta \quad \text{vs.} \quad H_1 : |\epsilon| < \delta. \tag{3.49}$$

The null hypothesis of test for equivalence leads to two the null hypotheses that are

$$\begin{aligned} &H_{01} : \epsilon \leq -\delta \quad \text{vs.} \quad H_{11} : \epsilon > -\delta \quad \text{and} \\ &H_{02} : \epsilon \geq \delta \quad \text{vs.} \quad H_{12} : \epsilon < \delta \end{aligned} \tag{3.50}$$

Then the two the null hypotheses, the test statistics under the null hypotheses are indicated below.

$$z_{01} = \frac{\hat{p}_1 - \hat{p}_2 - \delta}{\sqrt{\frac{\hat{p}_1 \times (1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2 \times (1-\hat{p}_2)}{n_2}}} \quad \text{and} \quad z_{02} = \frac{\hat{p}_1 \times \hat{p}_2 + \delta}{\sqrt{\frac{\hat{p}_1 \times (1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2 \times (1-\hat{p}_2)}{n_2}}} \tag{3.51}$$

H_0 is rejected if $z_{01} > z_\alpha$ or $z_{02} < -z_\alpha$. Under the alternative hypotheses the test statistics are:

$$z_{11} = \frac{p_1 - p_2 - (\epsilon + \delta)}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} \quad \text{and} \quad z_{12} = \frac{p_1 - p_2 - (\epsilon - \delta)}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} \tag{3.52}$$

Power of these hypotheses presented separately as $1 - \beta_1$ and $1 - \beta_2$ where $\beta = \beta_1 + \beta 2$

$$\begin{aligned} 1 - \beta_1 &\approx \Phi\left(\frac{\epsilon + \delta}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha\right) \quad \text{and} \\ 1 - \beta_2 &\approx \Phi\left(\frac{\epsilon - \delta}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha\right) \\ 2 - \beta_1 - \beta 2 &= \Phi\left(\frac{\epsilon + \delta}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha\right) + \Phi\left(\frac{\epsilon - \delta}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha\right) \\ 2 - \beta &= 2\Phi\left(\frac{\delta - |\epsilon|}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha\right) \\ 1 - \beta/2 &= \Phi\left(\frac{\delta - |\epsilon|}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha\right) \end{aligned} \tag{3.53}$$

As a result, to find the minimum sample size, the following equation is needed to solve:

$$z_{\beta/2} = \frac{\delta - |\epsilon|}{\sqrt{\frac{p_1 \times (1-p_1)}{n_1} + \frac{p_2 \times (1-p_2)}{n_2}}} - z_\alpha \tag{3.54}$$

If the allocation ratio is taken as r, the sample size is found as following (6):

$$n_2 = \frac{(z_{\beta/2} + z_\alpha)^2}{(\delta - |\epsilon|)^2}[\frac{p_1 \times (1-p_1)}{r} + p_2 \times (1-p_2)] \tag{3.55}$$

3.2 Cross-Over Trials

In this section, we give sample size calculations 2x2 m replicated cross-over trials for continuous and binary primary end points.

3.2.1 Equality Trials

In equality trials, we investigate whether or not the effect of the two treatments are same when all subjects take both of the treatments.

Continuous Primary End Point

Suppose y_{ijkl} is continuous response of j^{th} subject (j=1,2,...,n) in i^{th} sequence (i=1,2) with k^{th} treatment (k=1,2) on l^{th} replicate (l=1,2,...,m).

$$y_{ijkl} = \mu_k + \gamma_{ik} + s_{ijk} + e_{ijkl} \tag{3.56}$$

μ_k represents the effect of kth treatment(1,2), γ_{ik} represents the fixed effect of the ith sequence in kth treatment, s_{ijk} represents the random effect of the jth subject in the ith sequence on lth replicate. Assume there is no interaction effect between sequence and treatment. Joint distribution of (s_{ij1}, s_{ij2}) is assumed as a bivariate normal random variable with mean 0 and covariance matrix:

$$\sum = \begin{bmatrix} \sigma_{BT}^2 & \rho\sigma_{BT}\sigma_{BR} \\ \rho\sigma_{BT}\sigma_{BR} & \sigma_{BR}^2 \end{bmatrix} \tag{3.57}$$

where σ_{BT}^2 represents variance between subjects for "new treatment" group while σ_{BR}^2 represents the variance between subjects for "standard treatment" group. Moreover, ρ represents the correlation between subjects in "new treatment" and "standard treatment" groups. If σ_D^2 is defined as the variation based on the effect between subject and treatment interaction, which represents the heteroscedastic-

ity of subject random effect between new treatment and standard treatment.

$$\sigma_D^2 = \sigma_{BT}^2 + \sigma_{BR}^2 - 2\rho\sigma_{BT}\sigma_{BR} \tag{3.58}$$

Under the assumption of e_{ij1l} and e_{ij2l} are independent normally distributed with mean 0 and variances σ_{WT}^2 and σ_{WR}^2, respectively. σ_{WT}^2 and σ_{WR}^2 are reflects the variability within two different treatments.

$$\bar{y}_{ijk.} = \frac{1}{m}(y_{ijk1} + y_{ijk2} + ... + y_{ijkm}) \quad \text{and} \quad d_{ij} = \bar{y}_{ij1.} - \bar{y}_{ij2.} \tag{3.59}$$

Suppose ϵ =μ_2-μ_1 (new treatment-standard treatment). An unbiased estimates of $\hat{\epsilon}$ is written as:

$$\hat{\epsilon} = \frac{1}{2n}\sum_{i=1}^{2}\sum_{j=1}^{n} d_{ij} \tag{3.60}$$

Under the model 3.56, $\hat{\epsilon} \sim N(\epsilon, \sigma_m^2/2n)$. σ_m^2 represents the within subject standard deviation. The formulation of σ_m^2 is:

$$\sigma_m^2 = \sigma_D^2 + \sigma_{WT}^2 + \sigma_{WR}^2 \tag{3.61}$$

Since σ_m^2 is usually unknown, we need to use unbiased estimate of σ_m^2. An unbiased estimate of $\hat{\sigma_m^2}$ is:

$$\begin{aligned} \hat{\sigma_m^2} &= \frac{1}{2n-2}\sum_{i=1}^{2}\sum_{j=1}^{n}(d_{ij} - \bar{d_{i.}})^2 \\ &\text{where,} \quad \bar{d_{i.}} = \frac{1}{n}\sum_{j=1}^{n} d_{ij} \end{aligned} \tag{3.62}$$

For equality trials, we figure out $H_0 : \epsilon = 0$ or $H_0 : \epsilon \neq 0$. Under the null hypothesis, test statistic can be written as:

$$z_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.}}{\sqrt{\sigma_m^2/2n}} \tag{3.63}$$

If $|z_0| > z_{\alpha/2}$, H_0 is rejected. Under the alternative hypothesis the test statistic is:

$$z_1 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \epsilon}{\sqrt{\sigma_m^2/2n}} \tag{3.64}$$

If one can try to write the relationship between the test statistics under the null hypothesis and the alternative hypothesis, the following equation is obtained:

$$z_0 = z_1 + \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} \tag{3.65}$$

The type II error can be written with respect to test statistics z_0 and z_1:

$$\begin{aligned}
\beta &= P(|z_0| \leq z_{\alpha/2}) = P(-z_{\alpha/2} \leq z_0 \leq z_{\alpha/2}) \\
&= P(-z_{\alpha/2} \leq z_1 + \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} \leq z_{\alpha/2}) \\
&= P(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} \leq z_1 \leq z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) \\
&= \Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) - \Phi(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}})
\end{aligned} \tag{3.66}$$

As a result of Equation 3.66, power (1-β) would be written as:

$$\begin{aligned}
1 - \beta &= 1 - [\Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) - \Phi(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}})] \\
&= 1 - \Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) + \Phi(-z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) \\
&= \Phi(\frac{\epsilon}{\sqrt{\sigma_m^2/2n}} - z_{\alpha/2}) + \Phi(z_{\alpha/2} - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}})
\end{aligned} \tag{3.67}$$

Second term in Equation 3.67 is neglected as it's too small (6). Therefore, power would be found in Equation 3.68.

$$1 - \beta \approx \Phi(\frac{\epsilon}{\sqrt{\sigma_m^2/2n}} - z_{\alpha/2}) \tag{3.68}$$

The sample size can be found by solving Equation 3.69 (6) with respect to n.

$$z_\beta = \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} - z_{\alpha/2} \tag{3.69}$$

The sample size is written in Equation 3.70.

$$n = \frac{(z_{\alpha/2} + z_\beta)^2 \sigma_m^2}{2\epsilon^2} \tag{3.70}$$

If σ_m^2 is unknown, the test statistic under H_0 is calculated in Equation 3.71,

$$t_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.}}{\sqrt{\hat{\sigma_m}^2/2n}} \tag{3.71}$$

If $t_0 > t_{(\alpha/2,2n-2)}$, H_0 is rejected. Then the power of hypothesis test would be found with non-centrality parameter θ ($\theta = \frac{|\epsilon|}{\sqrt{\sigma_m^2/2n}}$):

$$1 - \beta = 1 - T_{2n-2}(t_{\alpha/2,2n-2}|\frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) \tag{3.72}$$

The sample size for non-central distribution approach can be found from Table 3.1 (6).

Binary Primary End Point

Suppose y_{ijkl} is binary response of j^{th} subject (j=1,2,...,n) in i^{th} sequence (i=1,2) with k^{th} treatment (k=1,2) on l^{th} replicate (l=1,2,...,m). Under the assumption of no period, sequence or carry-over effect, $P(y_{ijkl} = 1) = p_k$. Suppose ϵ =p_2-p_1 (new treatment-standard treatment).

$$\bar{y}_{ijk.} = \frac{1}{m}(y_{ijk1} + y_{ijk2} + ... + y_{ijkm}) \quad \text{and} \quad d_{ij} = \bar{y}_{ij1.} - \bar{y}_{ij2.} \tag{3.73}$$

An unbiased estimate of $\hat{\epsilon}$ is

$$\hat{\epsilon} = \frac{1}{2n}\sum_{i=1}^{2}\sum_{j=1}^{n} d_{ij}, \tag{3.74}$$

Based on Central Limit Theorem $\hat{\epsilon}$ asymptotically normally distributed as $\mathrm{N}(0, \sigma_d^2)$. σ_d^2 is unbiased estimate of the variance of $\hat{\epsilon}$. $\hat{\sigma}_d^2$ is estimated as in Equation 3.75:

$$\hat{\sigma}_d^2 = \frac{1}{2n-2}\sum_{i=1}^{2}\sum_{j=1}^{n}(d_{ij} - \bar{d_{i.}})^2 \tag{3.75}$$

For equality trials, hypotheses are written as $H_0 : \epsilon = 0$ and $H_1 : \epsilon \neq 0$. In this study, we only take calculations based on normal approximation to binomial into account. Therefore, under the null hypothesis, the test statistic would be:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{\sigma}_d^2/2n}} \tag{3.76}$$

Reject H_0 if $|z_0| > z_{\alpha/2}$. Moreover, under the alternative hypothesis, which indicates $\epsilon \neq 0$ power can be found from Equation 3.77:

$$1 - \beta \approx \Phi(\frac{|\epsilon|}{\sqrt{\hat{\sigma}_d^2/2n}} - z_{\alpha/2}) \tag{3.77}$$

To reach a power of 1-β, following equation should be solved (6).

$$z_\beta = \frac{|\epsilon|}{\sqrt{\hat{\sigma}_d^2/2n}} - z_{\alpha/2} \tag{3.78}$$

Minimum sample size required for cross-over equality design is found in Equation 3.79 (6):

$$n = \frac{(z_\beta + z_{\alpha/2})^2 \hat{\sigma}_d^2}{2\epsilon^2} \tag{3.79}$$

3.2.2 Superiority Trials

For superiority trials, we figure out that new treatment is superior to standard treatment.

Continuous Primary End Point

$H_0 : \epsilon \leq \delta$ or $H_1 : \epsilon > \delta$. Under the null hypothesis, test statistic can be written as:

$$z_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \delta}{\sqrt{\sigma_m^2/2n}} \tag{3.80}$$

If $z_0 > z_\alpha$, H_0 is rejected. Under the alternative hypothesis the test statistic is:

$$z_1 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - (\epsilon - \delta)}{\sqrt{\sigma_m^2/2n}} \tag{3.81}$$

If one can try to write the relationship between the test statistics under the null hypothesis and the alternative hypothesis, the following equation is obtained.

$$z_0 = z_1 + \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} \tag{3.82}$$

The type II error can be written with respect to test statistics z_0 and z_1.

$$\begin{aligned} \beta &= P(z_0 \leq z_\alpha) = P(z_0 \leq z_\alpha) \\ &= P(z_1 + \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} \leq z_\alpha) \\ &= P(z_1 \leq z_\alpha - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) \\ &= \Phi(z_\alpha - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) \end{aligned} \tag{3.83}$$

As a result of Equation 3.83, power (1-β) would be written as:

$$\begin{aligned} 1-\beta &= 1-\Phi(z_\alpha - \frac{\epsilon}{\sqrt{\sigma_m^2/2n}}) \\ &= \Phi(\frac{\epsilon}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \end{aligned} \tag{3.84}$$

Sample size can be found by solving Equation 3.85 regarding to n (6).

$$z_\beta = \frac{\epsilon}{\sqrt{\sigma_m^2/2n}} - z_\alpha \tag{3.85}$$

The minimum required sample size is written in Equation 3.86.

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma_m^2}{2(\epsilon - \delta)^2} \tag{3.86}$$

If σ_m^2 is unknown, the test statistic under H_0 is calculated in Equation 3.87,

$$t_0 = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \delta}{\sqrt{\hat{\sigma_m}^2/2n}} \tag{3.87}$$

If $t_0 > t_{(\alpha/2,2n-2)}$, H_0 is rejected. Then power of hypothesis test would be found with non centrality parameter θ ($\theta = \frac{\epsilon-\delta}{\sqrt{\sigma_m^2/2n}}$).

$$1-\beta = 1 - T_{2n-2}(t_{\alpha,2n-2}|\frac{\epsilon-\delta}{\sqrt{\sigma_m^2/2n}}) \tag{3.88}$$

Sample size for non-central distribution approach can be found from Table 3.1 (6).

Binary Primary End Point

For superiority trials, hypotheses are written as $H_0 : \epsilon \leq \delta$ and $H_1 : \epsilon > \delta$. In this study, we only take into account calculations based on normal approximation to binomial. Therefore, under the null hypothesis the test statistic would be:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{\sigma}_d^2/2n}}. \tag{3.89}$$

Reject H_0 if $z_0 > z_\alpha$. Moreover, under the alternative hypothesis which indicates

$\epsilon \neq 0$ power can be found from Equation 3.90.

$$1 - \beta \approx \Phi(\frac{\epsilon - \delta}{\sqrt{\hat{\sigma}_d^2/2n}} - z_\alpha) \tag{3.90}$$

To reach a power of 1-β, minimum sample size required for cross-over superiority design is found in Equation 3.91 (6):

$$n = \frac{(z_\beta + z_\alpha)^2 \hat{\sigma}_d^2}{2(\epsilon - \delta)^2} \tag{3.91}$$

3.2.3 Non-Inferiority Trials

In non-inferiority trials, it is investigated that new treatment is not worse than standard one.

Continuous Primary End Point

For non-inferiority trials, we figure out $H_0 : \epsilon \leq \delta$ or $H_1 : \epsilon > \delta$. Obtaining sample size calculations are same with superiority trials.

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma_m^2}{2(\epsilon - \delta)^2} \tag{3.92}$$

If σ_m^2 is unknown, the non-central t distribution approach is used. Power can be calculated as follows:

$$1 - \beta = 1 - T_{2n-2}(t_{\alpha,2n-2}|\frac{\epsilon - \delta}{\sqrt{\sigma_m^2/2n}}) \tag{3.93}$$

Sample size for non-central distribution approach can be found from Table 3.1 (6).

Binary Primary End Point

For non-inferiority trials, hypotheses are written as $H_0 : \epsilon \leq \delta$ and $H_1 : \epsilon > \delta$.

$$n = \frac{(z_\beta + z_\alpha)^2 \hat{\sigma}_d^2}{2(\epsilon - \delta)^2} \tag{3.94}$$

3.2.4 Equivalence Trials

In equivalence trials, it is investigated that two treatments neither superior nor non-inferior to to each other. Both treatments have equivalent effects.

Continuous Primary End Point

The hypotheses for equivalence can be written as:

$$H_0 : |\epsilon| \geq \delta \quad \text{vs.} \quad H_1 : |\epsilon| < \delta. \tag{3.95}$$

The null hypothesis of test for equivalence leads to two the null hypotheses that are:

$$\begin{aligned} &H_{01} : \epsilon \leq -\delta \quad \text{vs.} \quad H_{11} : \epsilon > -\delta \quad \text{and} \\ &H_{02} : \epsilon \geq \delta \quad \text{vs.} \quad H_{12} : \epsilon < \delta. \end{aligned} \tag{3.96}$$

According to the two the null hypotheses, the test statistics under the null hypotheses are indicated below.

$$z_{01} = \frac{\bar{x}_{1.} - \bar{x}_{2.} + \delta}{\sqrt{\sigma_m^2/2n}} \quad \text{and} \quad z_{02} = \frac{\bar{x}_{1.} - \bar{x}_{2.} - \delta}{\sqrt{\sigma_m^2/2n}} \tag{3.97}$$

H_{01} and H_{02} are rejected if $z_{01} > z_\alpha$ and $z_{02} < -z_\alpha$, respectively. Under the alternative hypotheses the test statistics are:

$$z_{11} = \frac{\bar{x}_{1.} - \bar{x}_{2.} - (\epsilon + \delta)}{\sqrt{\sigma_m^2/2n}} \quad \text{and} \quad z_{12} = \frac{\bar{x}_{1.} - \bar{x}_{2.} - (\epsilon - \delta)}{\sqrt{\sigma_m^2/2n}} \tag{3.98}$$

Power of hypotheses testings are found based on two different the alternative hypothesis. Therefore, to reach power with 1-β, β divided as β_1 and β_2 ($\beta_1 + \beta_2 = \beta$).

$$1 - \beta_1 \approx \Phi(\frac{\epsilon + \delta}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \quad \text{and} \quad 1 - \beta_2 \approx \Phi(\frac{\epsilon - \delta}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \tag{3.99}$$

$$\begin{aligned} 2 - \beta_1 - \beta 2 &= \Phi(\frac{\epsilon + \delta}{\sqrt{\sigma_m^2/2n}} - z_\alpha) + \Phi(\frac{\epsilon - \delta}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \\ 2 - \beta_1 - \beta_2 &= 2\Phi(\frac{\delta - |\epsilon|}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \\ 2 - \beta &= 2\Phi(\frac{\delta - |\epsilon|}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \\ 1 - \frac{\beta}{2} &= \Phi(\frac{\delta - |\epsilon|}{\sqrt{\sigma_m^2/2n}} - z_\alpha) \end{aligned} \tag{3.100}$$

To find minimum required sample size should be solved with respect to sample size.

$$z_{\beta/2} = \frac{\delta - |\epsilon|}{\sqrt{\sigma_m^2/2n}} - z_\alpha \tag{3.101}$$

To find sample size with 1-β, Equation 3.101 should be solved with respect to n.

$$n = \frac{(z_\alpha + z_{\beta/2})^2 \sigma_m^2}{2(\delta - |\epsilon|)^2} \tag{3.102}$$

Similarly, if population variance is unknown, then using non-central t distribution is used based on sample variance. Power functions of equivalent trial is written as following:

$$\begin{aligned}
1 - \beta &= 1 - T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta - \epsilon}{\sqrt{\sigma_m^2/2n}}) - T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta + \epsilon}{\sqrt{\sigma_m^2/2n}}) \\
2 - \beta &= 2T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta - |\epsilon|}{\sqrt{\sigma_m^2/2n}}) \\
1 - \beta/2 &= T_{n_1+n_2-2}(t_{\alpha,n_1+n_2-2} \mid \frac{\delta - |\epsilon|}{\sqrt{\sigma_m^2/2n}})
\end{aligned} \tag{3.103}$$

Sample size can be found by using Table 3.1 (6).

Binary End Point

The hypotheses for equivalence can be written as:

$$H_0 : |\epsilon| \geq \delta \quad \text{vs.} \quad H_1 : |\epsilon| < \delta. \tag{3.104}$$

The null hypothesis of test for equivalence leads to two the null hypotheses that are

$$\begin{aligned}
&H_{01} : \epsilon \leq -\delta \quad \text{vs.} \quad H_{11} : \epsilon > -\delta \quad \text{and} \\
&H_{02} : \epsilon \geq \delta \quad \text{vs.} \quad H_{12} : \epsilon < \delta.
\end{aligned} \tag{3.105}$$

According to the two the null hypotheses, the test statistics under the null hypotheses are indicated below.

$$z_{01} = \frac{\hat{p}_1 - \hat{p}_2 - \delta}{\sqrt{\hat{\sigma}_d^2/2n}} \quad \text{and} \quad z_{02} = \frac{\hat{p}_1 - \hat{p}_2 + \delta}{\sqrt{\hat{\sigma}_d^2/2n}} \tag{3.106}$$

H_0 is rejected if $z_{01} > z_\alpha$ or $z_{02} < -z_\alpha$. Under the alternative hypotheses the test statistics are:

$$z_{11} = \frac{p_1 - p_2 - (\epsilon + \delta)}{\sqrt{\hat{\sigma}_d^2/2n}} \quad \text{and} \quad z_{12} = \frac{p_1 - p_2 - (\epsilon - \delta)}{\sqrt{\hat{\sigma}_d^2/2n}} \tag{3.107}$$

Power of these hypothesis presented separately as $1 - \beta_1$ and $1 - \beta_2$ where $\beta = \beta_1 + \beta 2$

$$\begin{aligned} 1 - \beta_1 &\approx \Phi(\frac{\epsilon + \delta}{\sqrt{\hat{\sigma}_d^2/2n}} - z_\alpha) \quad \text{and} \quad 1 - \beta_2 = \Phi(\frac{\epsilon - \delta}{\sqrt{\hat{\sigma}_d^2/2n}} - z_\alpha) \\ 2 - \beta &= 2\Phi(\frac{\delta - |\epsilon|}{\sqrt{\hat{\sigma}_d^2/2n}} - z_\alpha) \\ 1 - \beta/2 &= \Phi(\frac{\delta - |\epsilon|}{\sqrt{\hat{\sigma}_d^2/2n}} - z_\alpha) \end{aligned} \tag{3.108}$$

As a result to find minimum sample size the following equation is needed to solve:

$$\frac{\delta - |\epsilon|}{\sqrt{\hat{\sigma}_d^2/2n}} - z_\alpha = z_{\beta/2} \tag{3.109}$$

Sample size is found as following (6):

$$n = \frac{(z_\alpha + z_{\beta/2})^2 \sigma_d^2}{2(\delta - |\epsilon|)^2} \tag{3.110}$$

3.3 Numeric Examples

Numeric examples of sample size calculations under different trial objectives and study designs are given with respect to whether population variance is known or not. For those numeric examples, solution is also given by using "TrialSize" package (7) in for RStudio (8) version 1.0.136.

3.3.1 Parallel Group Design

Numeric examples are given for continuous and binary primary end points with respect to different trial objectives with parallel group design.

Numeric Examples for Equality Trials in Parallel Group Design

Continuous Primary End Point

A researcher wants to show that the effect of the new drug to decrease cholesterol level is similar to standard drug. To achieve this, mean responses of low density lipoprotein (LDL) of two treatments are compared. It is ex-

pected (from previous studies) that the difference between two treatment is 0.05mmol/dL. Standard deviation of the population is assumed as 0.1 mmol/dL. Type I and II errors are taken as 0.05 and 0.20. Allocation ratio is one (6). The minimum required sample size to test equality of these treatments can be found by applying Equation 3.11 in below.

$$n_2 = \frac{(z_\beta + z_{\alpha/2})^2(1+r)\sigma^2}{r\epsilon^2} = \frac{(1.96+0.84)^2\ 2\ 0.1^2}{0.05^2} = 62.72 \approx 63$$

Minimum sample size is 63 for each group based on calculation. If the population variance is unknown, Table 3.1 can be used. Let keep all parameters same in a numeric example with assuming the standard deviation for sample is also 0.1 mmol/dL. From Table 3.1, by taking non-centrality parameter $\theta = \frac{0.05}{0.1} = 0.5$, $\alpha = 0.025$ (as hypothesis testing for equality trials is two sided, $\alpha = 0.05/2 = 0.025$) and 1-β=0.80, minimum sample size found as 64 for each group. From "TrialSize" package, the required minimum sample size is found as 62.79≈ 63 for each group based on Equation 3.11.

R Codes for Equality Trials for Parallel Group Design with Continuous End Point

```
>twoeq=TwoSampleMean.Equality(alpha=0.05, beta=0.20,
+ sigma=0.1, k=1,margin=0.05)
>twoeq
[1]62.79104
```

According to R codes, alpha, beta, sigma, k and margin represent Type I error, Type II error, population standard deviation, allocation rate and expected difference between means of the two groups.

Binary Primary End Point

A pharmaceutical company wants to investigate the efficacy, safety, and tolerability of two antibacterial agents. The objective of this study is to search whether there is skin infection or not.To test the equality of two agents, pilot study is conducted. According to the pilot study, the proportions of the two agents of two antibacterial agents are found as p_1=0.65 and p_2=0.85 (ϵ=0.85-0.65=0.2). Type I error and Type II error are taken as 0.05 and 0.20, respectively (6). For equal sample size allocation, the minimum sample size is found by using Equation

3.21.

$$n_2 = \frac{(z_\beta + z_{\alpha/2})^2}{\epsilon^2}[\frac{p_1(1-p_1)}{r} + p_2(1-p_2)]$$
$$= \frac{(0.84+1.96)^2}{0.2^2}[\frac{0.65(1-0.65)}{1} + 0.85(1-0.85)] = 69.58 \approx 70$$

Minimum sample size is 70 for each group based on calculation. From "TrialSize" package, the required minimum sample size is found as 69.65≈ 70 for each group as seen in below.

R Codes for Equality Trials for Parallel Group Design with Binary End Point

```
>TwoSampleProportion.Equality(alpha=0.05, beta=0.20,
+ p1=0.85,p2=0.65,k=1,delta=0.2)
[1]69.65881
```

Alpha, beta, p_1, p_2, k and delta represent Type I error, Type II error, proportions, allocation ratio and difference between proportions of two groups in R codes.

Numeric Examples of Superiority Trials in Parallel Group Design

Continuous Primary End Point

The same numeric example with equality trial is investigated under same scenario. Additionally, superiority margin that is clinically meaningful difference to show the superiority, is accepted 0.01 mmol/dL. For equal sample size allocation, the minimum sample size is found by using Equation 3.28.

$$n_2 = \frac{(z_\alpha + z_\beta)^2(1+r)\sigma^2}{r(\epsilon-\delta)^2}$$
$$= \frac{(1.64+0.845)^2\ 2\ 0.1^2}{(0.05-0.01)^2} = 77.19 \approx 78$$

The minimum sample size is 78 for each group based on calculation. We consider standard deviation for sample is also 0.1 mmol/dL. From the Table 3.1, by taking non-centrality parameter as $\theta = \frac{0.05-0.01}{0.1} = 0.4$, $\alpha = 0.05$ (as hypothesis testing for superiority trials is one sided) and 1-β=0.80, the minimum sample size found as 78 for each group. From "TrialSize" package, the required minimum sample

size is found as 77.28≈ 78.

R Codes for Superiority Trials for Parallel Group Design with Continuous End Point

```
>twosup=TwoSampleMeans.NIS(alpha=0.05, beta=0.20,
+ sigma=0.1, k=1, delta=0.01, margin=0.05)
>twosup
[1]77.28197
```

In superiority trials,alpha, beta, sigma, k, delta and margin represent Type I error, Type II error, the population standard deviation, the allocation rate, the superiority margin (clinically importance difference) and the expected difference between means of two groups.

Binary Primary End Point

The same numeric example with equality trial is investigated. For superiority trial, 0.1 is accepted as clinically meaningful difference to show superiority of one of the agent ($\delta = 0.1$). If sample size allocation is considered as one, by using Equation 3.34 the minimum sample size is found as below.

$$n_2 = \frac{(z_\beta + z_\alpha)^2}{(\epsilon - \delta)^2}[\frac{p_1 \times (1-p_1)}{r} + p_2 \times (1-p_2)]$$
$$= \frac{(0.84+1.64)^2}{(0.2-0.1)^2}[\frac{0.65(1-0.65)}{1} + 0.85(1-0.85)] = 218.34 \approx 219$$

The minimum sample size is 219 for each group based on calculation. From "TrialSize" package, the required minimum sample size is found as 219.48≈220 as stated in the below.

R Codes for Superiority Trials for Parallel Group Design with Binary End Point

```
>TwoSampleProportion.NIS(alpha=0.05, beta=0.20,
+ p1=0.85,p2=0.65, k=1, delta=0.2, margin=0.1)
[1]219.4808
```

Alpha, beta, p_1, p_2, k, delta and margin represent Type I error, Type

II error, the population proportions, the allocation ratio, the difference of two proportions and the clinically meaningful difference to show superiority.

Numeric Examples of Non-inferiority Trials in Parallel Group Design

Continuous Primary End Point

The same numeric example with equality trial is investigated. 0.01 mmol/dL is accepted non-inferiority margin ($\delta = -0.01$). For equal sample size allocation, the minimum sample size is found by using Equation 3.36.

$$n_2 = \frac{(z_\alpha + z_\beta)^2(1+r)\sigma^2}{r(\epsilon - \delta)^2}$$
$$= \frac{(1.64 + 0.845)^2\ 2\ 0.1^2}{(0.05 - (-0.01))^2} = 34.31 \approx 35$$

The minimum sample size is 35 for each group based on calculation. For unknown population variance, considering standard deviation for sample as 0.1 mmol/dL, non-centrality parameter is calculated as $\theta = \frac{0.05-(-0.01)}{0.1} = 0.6$. From Table 3.1, with θ=0.5, α=0.5 and β=0.2, minimum sample size found as 36. From "TrialSize" package, the required minimum sample size is found 34.35≈ 35

R Codes for Non-Inferiority Trials for Parallel Group Design with Continuous End Point

```
>twonon=TwoSampleMeans.NIS(alpha=0.05, beta=0.20,
+ sigma=0.1, k=1, delta=-0.01, margin=0.05)
>twonon
[1]34.34754
```

Arguments of non-inferiority trials are same with arguments of superiority trials in R.

Binary Primary End Point

The same numeric example with equality trial is investigated. 0.1 is accepted as clinically meaningful difference to show non-inferiority of new treatment(δ=-0.1). With equal sample size allocation, the minimum sample size to show non-

inferiority is found as 3.38.

$$n_2 = \frac{(z_\beta + z_\alpha)^2}{(\epsilon - \delta)^2}[\frac{p_1 \times (1-p_1)}{r} + p_2 \times (1-p_2)]$$
$$= \frac{(0.84+1.64)^2}{(0.2-(-0.1))^2}[\frac{0.65(1-0.65)}{1} + 0.85(1-0.85)] = 24.26 \approx 25$$

The minimum sample size is 25 for each group based on calculation. From "TrialSize" package, the required minimum sample size is found as 24.38≈25

R Codes for Non-Inferiority Trials for Parallel Group Design with Binary End Point

```
>TwoSampleProportion.NIS(alpha=0.05, beta=0.20,
+ p1=0.85,p2=0.65, k=1, delta=0.2, margin=-0.1)
[1]24.38675
```

Arguments of non-inferiority trials are also same with arguments of superiority trials in R for binary end point.

Numeric Examples of Equivalence Trials in Parallel Group Design

Continuous Primary End Point

Same research topic is investigated in equality trial. It is expected (from previous studies) that the difference between two treatment is 0.01 mmol/dL. Standard deviation of the population is assumed as 0.1 mmol/dL. 0.05 mmol/dL is accepted as to show the equivalence limit. By taking equal sample size allocation in Equation 3.46, the minimum sample size can be calculated as below:

$$n_2 = \frac{(z_{\beta/2} + z_\alpha)^2(1+r)\sigma^2}{r(\delta - |\epsilon|)^2}$$
$$= \frac{(1.28+1.64)^2\ 2\ 0.1^2}{(0.05-|0.01|)^2} = 106.94 \approx 107$$

The minimum sample size is 107 for each group based on calculation. We assume sample standard deviation as 0.1 mmol/dL for unknown population variance. Therefore, $\theta = \frac{0.05-|0.01|}{0.1} = 0.4$, by taking $\beta = 0.10$ (As β is divided by two due to two the alternative hypotheses) and $\alpha = 0.05$, 108 subjects are needed for each group. From "TrialSize" package, the required minimum sample size is found as

107.05≈ 108.

R Codes for Equivalence Trials for Parallel Group Design with Continuous End Point

```
>twopareq=TwoSampleMeans.Equivalence(alpha=0.05,
+ beta=0.20,sigma=0.1, k=1, delta=0.05, margin=0.01)
>twopareq
[1]107.0481
```

Arguments in equivalent trials are same with in superiority trials.

Binary End Point

The same illustration is investigated in equality trials. According to the pilot study, the proportions of the two agents of two antibacterial agents are found as p_1=0.65 and p_2=0.75 (ϵ=0.75-0.65=0.1). Moreover, 0.2 is accepted as the clinically meaningful difference to say these two treatments are equivalent (δ=0.2). Considering equal sample size allocation, minimum required sample size can be found by using Equation 3.55 as in below.

$$\begin{aligned} n_2 &= \frac{(z_{\beta/2}+z_{\alpha})^2}{(\delta-|\epsilon|)^2}[\frac{p_1(1-p_1)}{r}+p_2(1-p_2)] \\ &= \frac{(1.28+1.64)^2}{(0.2-|0.1|)^2}(0.75\times 0.25+0.65\times 0.35) = 353.84 \approx 354 \end{aligned}$$

The minimum sample size is 356 for each group based on calculation. From "TrialSize" package, the required minimum sample size is found as 355.40≈ 356.

R Codes for Equivalence Trials for Parallel Group Design with Binary End Point

```
>TwoSampleProportion.Equivalence(alpha=0.05, beta=0.20,
+ p1=0.85,p2=0.65, k=1, delta=0.2, margin=-0.1)
[1]24.38675
```

3.3.2 Cross-Over Trials

We discuss numeric examples of sample size calculations 2x2 m replicated cross-over trials for continuous and binary primary end points in this section.

Numeric Examples of Equality Trials in Cross-Over Study Design

Continuous Primary End Point

Suppose a 2x2 (m=1) cross-over design to compare effect of new treatment and standard treatment to cholesterol level. A researcher wants to show that the effect of the new drug to decrease cholesterol level is similar to standard drug. To achieve this, mean responses of low density lipoprotein (LDL) of two treatments are compared. It is expected (from previous studies) that the difference between two treatment is 0.05 mmol/dL. Within subject standard deviation of the population is assumed as 0.1 mmol/dL. Type I and II errors are taken as 0.05 and 0.20 (6). The minimum required sample size to test equality of these treatments can be found by applying Equation 3.70 in below:

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2 \sigma_m^2}{2\epsilon^2} = \frac{(1.96 + 0.84)^2 0.1^2}{2 \times 0.05^2} = 15.68 \approx 16$$

The minimum sample size is 16 based on calculation. For the unknown variance case, assume unbiased estimation of within subject standard deviation of the sample to be 0.1 mmol/dL. From Table 3.1, by taking non-centrality parameter as $\theta = \frac{2|0.05|}{0.1} = 1.0$ with α=0.025 (α=0.05 is divided by 2 as for two-sided hypothesis test) and β=0.20, 16 subjects are needed. By using "TrialSize" package, 15.70≈ minimum 16 subjects are required as stated in below.

R Codes for Equality Trials for Cross-Over Group Design with Continuous End Point

```
>twocseq=TwoSampleCrossOver.Equality(alpha=0.05,
+ beta=0.20,sigma=0.1, margin=0.05)
>twocsup
[1]19.32049
```

Alpha, beta, sigma and margin represent Type I error, Type II error, the population within subject standard deviation and the expected difference between

means of two treatments as for R codes.

Binary Primary End Point

A pharmaceutical company wants to investigate the efficacy, safety, and tolerability of two antibacterial agents. The objective of this study is that whether there is skin infection or not. 2x2 cross-over design is constructed. According to pilot study, proportion of two antibacterial agents are found as p_1=0.65 and p_2=0.85 (ϵ=0.85-0.65=0.2). Type I error, Type II error and σ_d are taken as 0.05, 0.20 and 0.7 respectively (6). To test equality the minimum sample size is found by using Equation 3.79.

$$n = \frac{(z_\beta + z_{\alpha/2})^2 \hat{\sigma}_d^2}{2\epsilon^2} = \frac{(0.84 + 1.96)^2 0.7^2}{2 \times 0.2^2} = 48.02 \approx 49$$

The minimum sample size is 49 based on calculation. From "TrialSize" package, the required minimum sample size is found as 48.07≈49.

R Codes for Equality Trials for Cross-Over Design with Binary End Point

```
>TwoSampleSeqCrossOver.Equality(alpha=0.05, beta=0.20,
+sigma=0.49, sequence=2, delta=0.2)
[1]48.07439
```

Alpha, beta, sigma, sequence and delta represent Type I error, Type II error, variance, number of treatment sequence and difference between proportions from pilot study.

Numeric Examples of Superiority Trials in Cross-Over Study Design

The same numeric example with equality trial is investigated. 0.01 is accepted superiority margin ($\delta = 0.01$). The minimum sample size is found by using Equation 3.86 for testing superiority.

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma_m^2}{2(\epsilon - \delta)^2} = \frac{(1.64 + 0.84)^2 0.1^2}{2 \times (0.05 - 0.01)^2} = 19.22 \approx 20$$

The minimum sample size is 35 based on calculation. If population variance is unknown, assume unbiased estimation of within subject standard deviation for sample is 0.01 mmol/dL. From Table 3.1, $\theta = \frac{2(0.05-0.01)}{0.1} = 0.8$, 21 subjects are needed. By using "TrialSize" package, 19.32≈ minimum 20 subjects are required.

R Codes for Superiority Trials for Cross-Over Group Design with Continuous End Point

```
>twocssup=TwoSampleCrossOver.NIS(alpha=0.05,
+ beta=0.20,sigma=0.1, delta=0.01, margin=0.05)
>twocssup
[1]19.32049
```

According to R codes, alpha, beta, sigma, delta and margin represent Type I error, Type II error, population within subject standard deviation, superiority margin and expected difference between means of two treatments.

Binary Primary End Point

The same numeric example with equality trial is investigated. 0.1 is accepted as clinically meaningful difference to show superiority of new treatment(δ=0.1). Minimum sample size to show superiority is found as in Equation 3.91.

$$n = \frac{(z_\beta + z_\alpha)^2 \hat{\sigma}_d^2}{2(\epsilon - \delta)^2}$$
$$= \frac{(0.84 + 1.64)^2 0.7^2}{2 \times (0.2 - 0.1)^2} = 150.69 \approx 151$$

Minimum sample size is 152 based on calculation.From "TrialSize" package, the required minimum sample size is found as 151.47≈152.

R Codes for Superiority Trials for Cross-Over Design with Binary End Point

```
>TwoSampleSeqCrossOver.NIS(alpha=0.05, beta=0.20,
+sigma=0.49, sequence=2, delta=0.2,margin=0.1)
[1]151.4727
```

Alpha, beta, sigma, sequence, delta and margin represent Type I error,

Type II error, variance, treatment sequence, difference between proportions from pilot study and superiority margin.

Numeric Examples of Non-Inferiority Trials in Cross-over Study Design

The same numeric example with equality trial is investigated. 0.01 mmol/dL change is accepted non-inferiority margin ($\delta = -0.01$). The minimum sample size is found by using Equation 3.92 for testing non-inferiority:

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma_m^2}{2(\epsilon - \delta)^2}$$
$$= \frac{(1.64 + 0.84)^2 0.1^2}{2 \times (0.05 - (-0.01)^2} = 8.54 \approx 9$$

Minimum sample size is 9 based on calculation. For unknown variance, if 0.1 is taken as unbiased estimation of within subject standard deviation non-centrality parameter is calculated as $\theta = \frac{2(0.05-(-0.01))}{0.1} = 1.2$ From Table 3.1, 10 subjects are needed. By using "TrialSize" package, 8.59 $\approx$ minimum 9 subjects are required.

R Codes for Non-Inferiority Trials for Cross-Over Group Design with Continuous End Point

```
>twocsno=TwoSampleCrossOver.NIS(alpha=0.05,
+ beta=0.20,sigma=0.1, delta=-0.01, margin=0.05)
>twocsno
[1]8.586685
```

Arguments of R codes are same with superiority trials.

Binary Primary End Point

The same numeric example with equality trial is investigated. 0.1 is accepted as clinically meaningful difference to show non-inferiority of new treatment(δ=-0.1). Minimum sample size to show non-inferiority is found as in Equation 3.94.

$$n = \frac{(z_\beta + z_\alpha)^2 \hat{\sigma}_d^2}{2(\epsilon - \delta)^2}$$
$$= \frac{(0.84 + 1.64)^2 0.7^2}{2 \times (0.2 - (-0.1))^2} = 16.74 \approx 17$$

Minimum sample size is 17 based on calculation. From "TrialSize" package, the required minimum sample size is found as 16.83≈17.

R Codes for Non-Inferiority Trials for Cross-Over Design with Binary End Point

```
>TwoSampleSeqCrossOver.NIS(alpha=0.05, beta=0.20,
+sigma=0.49, sequence=2, delta=0.2,margin=-0.1)
[1]16.83029
```

Arguments of R codes in non-inferiority trials are same with superiority trials.

Numeric Examples of Equivalence Trials in Cross-over Study Design

Continuous Primary End Point

Same research topic is investigated in equality trial. It is expected (from previous studies) that the difference between two treatment is 0.01 mmol/dL. Within subject standard deviation of the population is assumed as 0.1 mmol/dL. 0.05 mmol/dL is accepted as to show the equivalence limit. From Equation 3.102, the minimum sample size can be calculated as below:

$$\begin{aligned} n &= \frac{(z_\alpha + z_{\beta/2})^2 \sigma_m^2}{2(\delta - |\epsilon|)^2} \\ &= \frac{(1.64 + 1.28)^2 0.1^2}{2 \times (0.05 - |0.01|)^2} = 26.65 \approx 27 \end{aligned}$$

Minimum sample size is 27 based on calculation. For unknown population variance cases, if we find unbiased estimation of within subject standard deviation as 0.1, from Table 3.1, $\theta = \frac{2(0.05-|0.01|)}{0.1} = 0.8$, with α=0.05 and β=0.10 (β=0.20 is divided by two due to two the null hypothesis), 28 subjects are needed. By using "TrialSize" package, 26.76≈ minimum 27 subjects are required as stated in below.

R Codes for Equivalence Trials for Cross-Over Group Design with Continuous End Point

```
>twocsequ=TwoSampleCrossOver.Equivalence(alpha=0.05,
+ beta=0.20,sigma=0.1, delta=0.05, margin=0.01)
>twocsequ
[1]26.76202
```

Arguments in equivalence trials are same with superiority trials.

Binary End Point

The same numeric example with equality trial is investigated. According to pilot study, proportion of two antibacterial agents are found as p_1=0.75 and p_2=0.85 (ϵ=0.85-0.75=0.1). 0.20 is accepted as clinically meaningful difference to show the effect of those of treatments are equivalent (δ=0.2). The minimum sample size is found by using Equation 3.110 for testing equivalence.

$$n = \frac{(z_\alpha + z_{\beta/2})^2 \sigma_d^2}{2(\delta - |\epsilon|)^2} = \\ = \frac{(1.28 + 1.64)^2 0.7^2}{2 \times (0.2 - |0.1|)^2} = 208.90 \approx 209$$

Minimum sample size is 209 based on calculation.From "TrialSize" package, the required minimum sample size is found as 209.81≈210.

R Codes for Equivalence Trials for Cross-Over Design with Binary End Point

```
>TwoSampleSeqCrossOver.Equivalence(alpha=0.05,
+ beta=0.20,sigma=0.49, sequence=2,delta=0.2,margin=0.1)
[1]209.8143
```

Arguments of R codes for equivalence studies are same as in superiority trials.

Table 3.2. Results of Numeric Example for Continuous Primary End Point

	Parallel Group Design				2x2 Cross-over Design			
	Equality	Superiority	Non-inferiority	Equivalence	Equality	Superiority	Non-inferiority	Equivalence
Formulation	63×2	78×2	35×2	107×2	16	20	9	27
Non-central t table	64×2	78×2	36×2	108×2	16	21	10	28
R	63×2	78×2	35×2	108×2	16	20	9	27

Table 3.3. Results of Numeric Example for Binary Primary End Point

	Parallel Group Design				2x2 Cross-over Design			
	Equality	Superiority	Non-inferiority	Equivalence	Equality	Superiority	Non-inferiority	Equivalence
Formulation	70×2	219×2	25×2	304×2	49	151	17	209
R	70×2	219×2	25×2	305×2	49	152	17	210

In Table 3.2 and 3.3, the sample sizes calculated with different tools are represented based on numeric examples. With regarding to standard normal distribution approach, non-central t distribution approach (by using Table 3.1) and using R program, sample sizes are almost same with each other in each trial objective for different primary end points.

3.4 Simulation Scenarios

3.4.1 Fixed Clinical Importance Effect

In this part, we create simulation scenarios with respect to different sample sizes, distributions, trial objectives and study designs. Suppose a scenario such that new drug and standard drug are compared in parallel group and cross-over designs whether effect of decrease headache pain is same or not. The means of new drug and standard drug obtained from pilot study are 2.5 and 2, respectively (ϵ=0.5). By taking known population variance, β and α 4, 0.2 and 0.05. The minimum required sample sizes are stated in Table 3.4 and 3.5 for parallel group(with taking allocation ratio as 1) and cross-over designs.

Table 3.4. Minimum Required Sample Size for Parallel Group Design

Trial objective	δ (Superiority or Non-inferiority margin)	Minimum sample size
Equality	-	252
Superiority	0.1	310
Non-inferiority	0.1	138

Table 3.5. Minimum Required Sample Size for Cross-over Design

Trial objective	δ (Superiority or Non-inferiority margin)	Minimum sample size
Equality	-	63
Superiority	0.1	78
Non-inferiority	0.1	35

Simulation scenarios are formed based on following assumptions:

- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from Normal distributions X∼N(2.5,4) and Y∼N(2,4).
- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250,300 from X is distributed as lognormal with mean 2.5 and variance 4 and Y∼ N(2,4).
- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from lognormal distribution with means 2.5, and 2 and variance 4.
- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from X is distributed as lognormal with mean 2.5 and variance 4 and Y∼Exp(0.5).

To reach intended mean and variance of lognormal distributions, we consider two-parameter lognormal distribution. Suppose a random variable Y has a lognormal distribution with where μ location parameter and σ shape parameter if ln(Y)∼N(μ,σ^2). The distribution of lognormal distribution is as following (9).

$$f(y) = \frac{1}{\sqrt{2\pi} \times \sigma \times y} exp\left(-\frac{(ln(y) - \mu)^2}{2 \times \sigma^2}\right), \quad y > 0 \tag{3.111}$$

Based on Equation 3.111, mean and variance of lognormal distribution can be found by using location and scale parameters.

$$\begin{aligned} E(Y) &= exp(\mu + \frac{1}{2} \times \sigma^2) \\ Var(Y) &= exp(2 \times (\mu + \sigma^2)) - exp(2 \times \mu + \sigma^2) \end{aligned} \tag{3.112}$$

For specified mean and variance of lognormal distribution, if Equation 3.112 is solved with respect to μ and σ, one can get parameters of lognormal distribution. Then, lognormal distribution could be generated easily.

3.4.2 Fixed Effect Sizes

We apply simulation scenarios with respect to different sample sizes, distributions, trial objectives and study designs. For all conditions, we specify effect sizes as small, medium and large which are equal to 0.2, 0.5 and 0.8, respectively (10).

- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from Normal distributions by taking into consideration small, medium and large effect sizes.
- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from lognormal and normal distribution considering small, medium and large effect sizes.
- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from lognormal distributions based on small, medium and large effect sizes.
- Generate two independent distributions with n=10, 30, 50, 100, 150, 200, 250, 300 from lognormal and exponential distribution considering small, medium and large effect sizes.

4 RESULTS

In this chapter, simulation results are given. According to simulation scenarios that are given in Chapter3, we try to detect in how many trials we get reach at least 80% of power over 10000 replications for each of the simulation scenarios of equality, superiority and non-inferiority cases. We used "pwr" package (11).

4.1 Simulation Results

4.1.1 Fixed clinical Importance Effect

In Tables 4.1, 4.2 and 4.3, we present the rate of number of samples whose power is more than 80% with parallel group under different trial objectives, sample sizes and different distributions. To get observed power, two sample t test approach is applied.

Table 4.1. Simulation Results in Equality Trials for Parallel Group Design

		Equality Trials							
Response of New Drug	Response of Standard Drug	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
X∼N(2.5,4)	Y∼N(2,4)	0.0074	0.0287	0.0566	0.1519	0.2619	0.3839	0.4900	0.6023
X∼Lognormal(μ=0.67,σ=0.70)	Y∼N(2,4)	0.0002	0.0053	0.0179	0.1013	0.2266	0.3584	0.4827	0.6033
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Lognormal(μ=0.35,σ=0.83)	0.0132	0.0299	0.0551	0.1460	0.2605	0.3854	0.4932	0.6056
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Exp(0.5)	0.0134	0.0315	0.0590	0.1465	0.2568	0.3828	0.4881	0.6084

Table 4.2. Simulation Results in Superiority Trials for Parallel Group Design

		Superiority Trials							
Response of New Drug	Response of Standard Drug	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
X∼N(2.5,2)	Y∼N(2,2)	0.0156	0.0417	0.0647	0.1375	0.2317	0.3143	0.3969	0.4843
X∼Lognormal(μ=0.67,σ=0.70)	Y∼N(2,2)	0.0238	0.0469	0.0731	0.1421	0.2142	0.3030	0.3790	0.4834
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Lognormal(μ=0.35,σ=0.83)	0.0170	0.0355	0.0638	0.1389	0.2190	0.3055	0.3971	0.4867
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Exp(0.5)	0.0210	0.0437	0.0680	0.1380	0.2265	0.3139	0.3993	0.4829

Table 4.3. Simulation Results in Non-inferiority Trials for Parallel Group Design

		Non-Inferiority Trials							
Response of New Drug	Response of Standard Drug	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
X∼N(2.5,2)	Y∼N(2,2)	0.0268	0.0901	0.1500	0.3579	0.5398	0.7019	0.8034	0.8777
X∼Lognormal(μ=0.67,σ=0.70)	Y∼N(2,2)	0.0363	0.0909	0.1588	0.3483	0.5358	0.6957	0.8086	0.8852
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Lognormal(μ=0.35,σ=0.83)	0.0287	0.0848	0.1552	0.3609	0.5421	0.7054	0.8060	0.8832
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Exp(0.5)	0.0277	0.0882	0.1577	0.3528	0.5449	0.6868	0.8031	0.8818

According to Tables 4.1, 4.2 and 4.3, when the sample size increases, the likelihood that we get 80% power increases. In Table 3.4, we give the minimum required sample sizes in each group to reach 80% power for the scenario we created. Therefore, if we compare the results of the observed power in Tables 4.5, 4.6 and 4.7 with Table 3.4, as to show superiority we need larger sample than equality and non-inferiority, the fewest observed power is obtained from for superiority trials. On the other hand, since we need smaller sample in non-

inferiority trials, the highest rate of reaching 80% of observed power in non-inferiority trials.

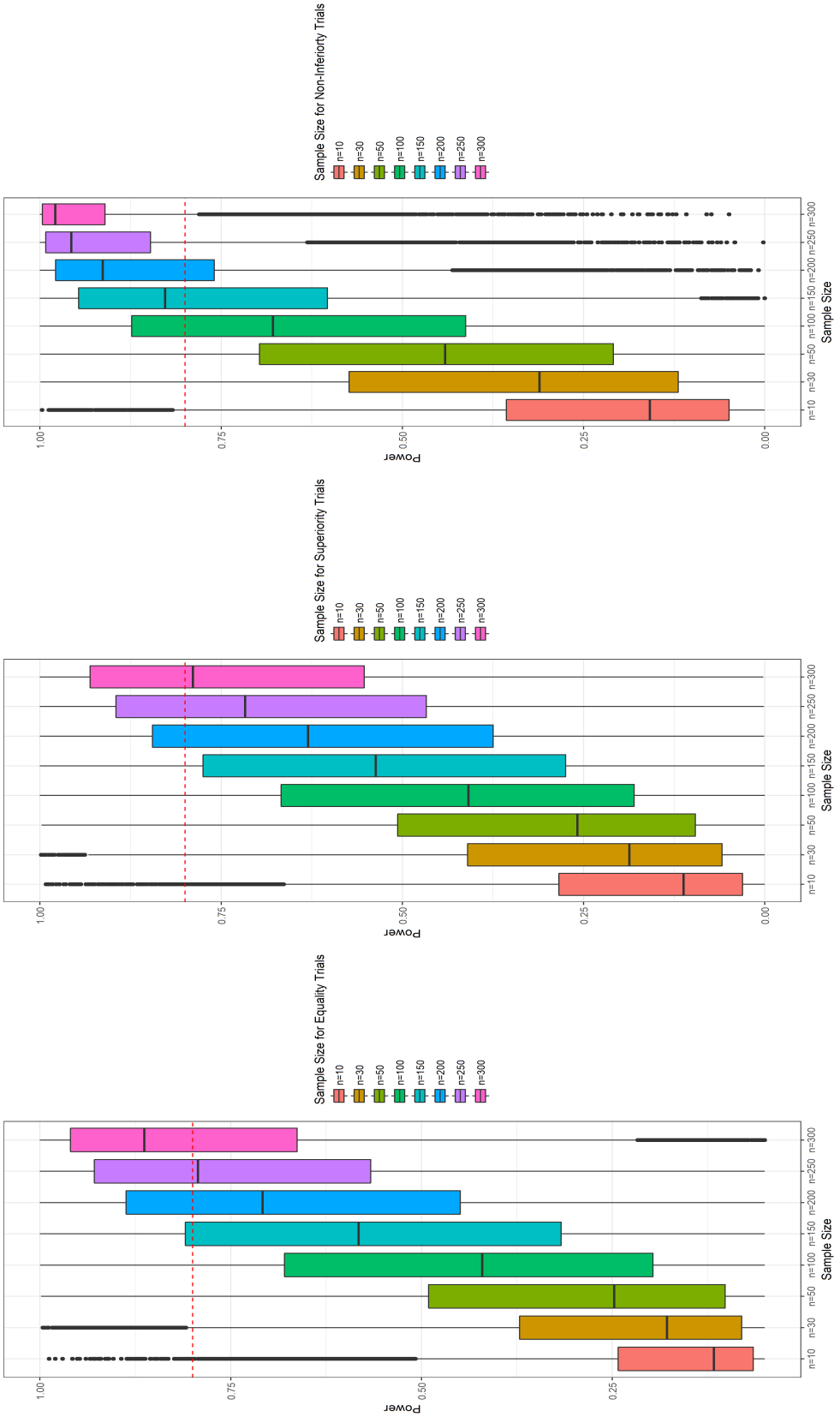

Figure 4.1. Simulation Results for Responses of Both Groups are Normally Distributed for Parallel Group Design

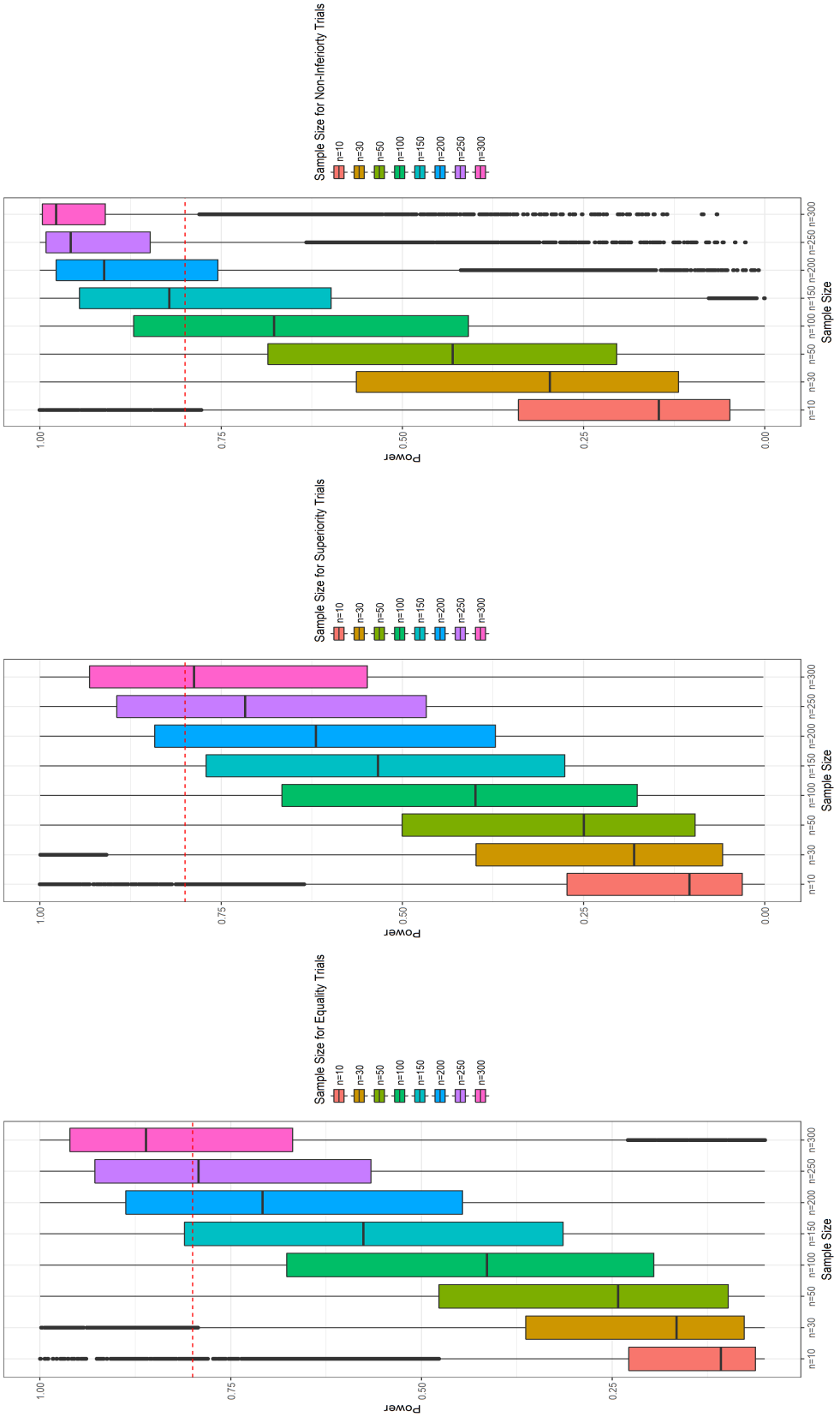

Figure 4.2. Simulation Results for Responses of Groups Have Lognormal and Normal Distributions for Parallel Group Design

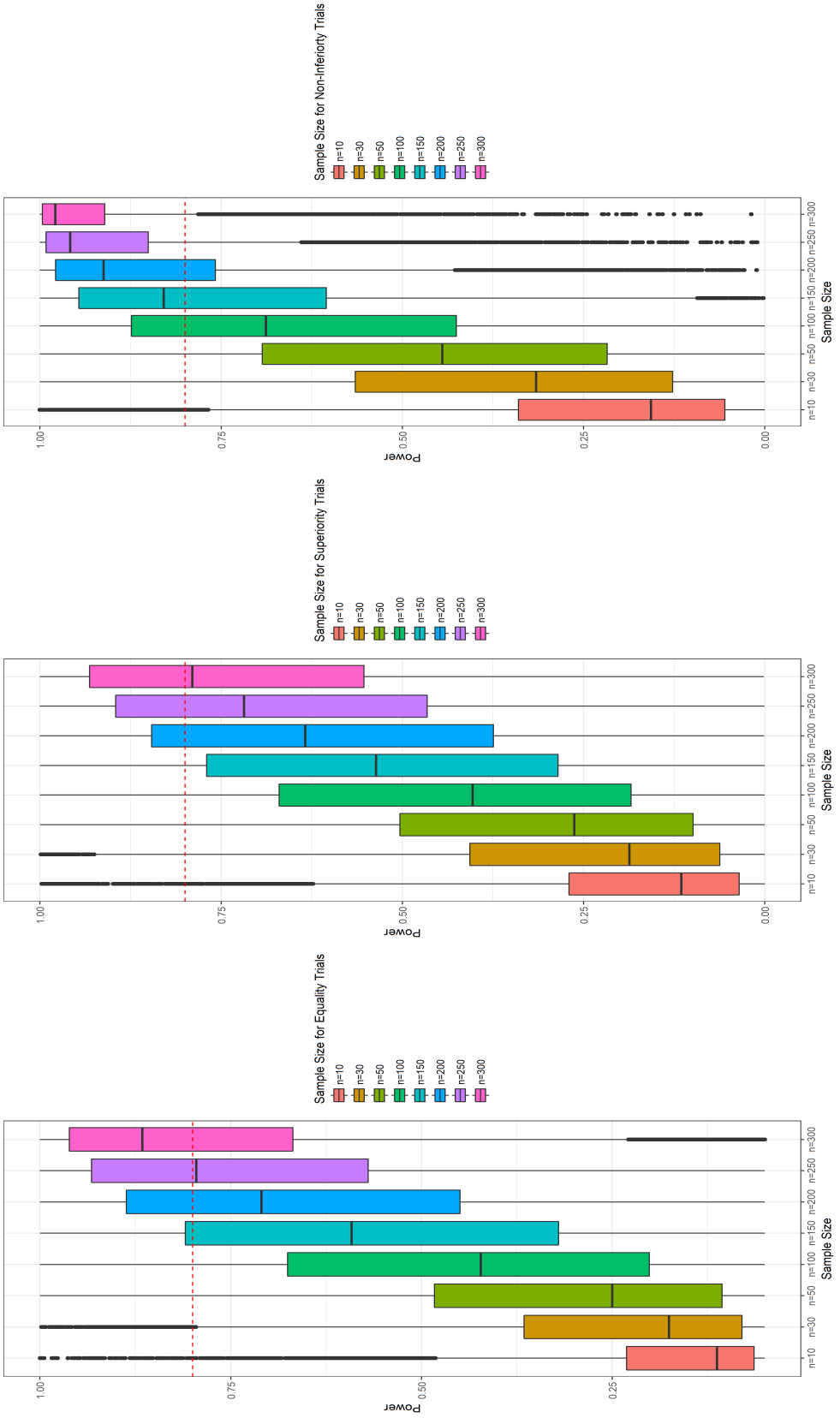

Figure 4.3. Simulation Results for Responses of Both Groups Have Lognormal Distribution for Parallel Group Design

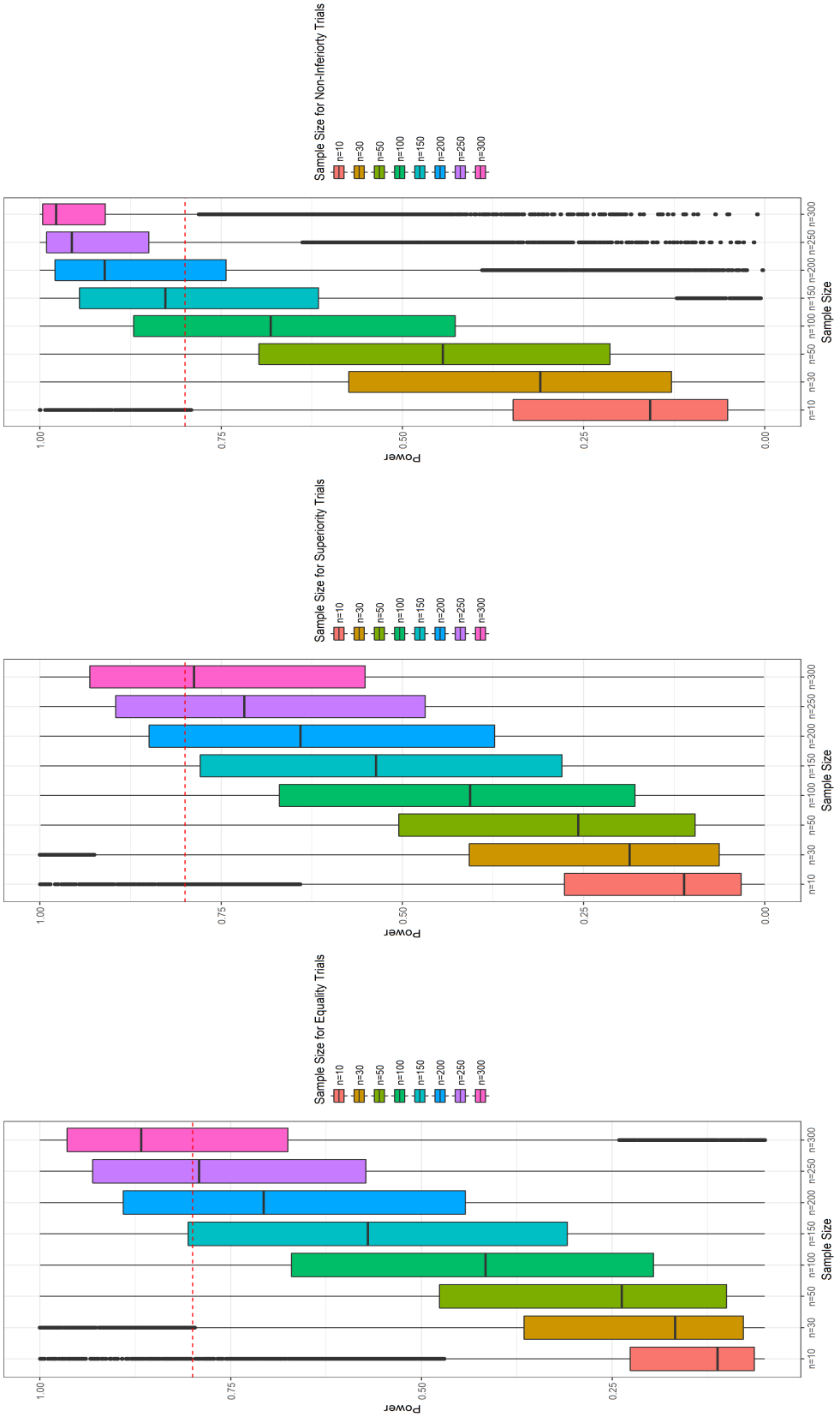

Figure 4.4. Simulation Results for Responses of Groups Have Lognormal and Exponential Distributions for Parallel Group Design

Figures 4.1, 4.2, 4.3 and 4.4 are given to show same results by using box plots. We draw the plots in Figures 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7 and 4.8 by using "ggplot2" package (12) in RStudio. In these figures, dashed red line represents the place of 80% power. Moreover, to highlight the effect of different trial objectives, under same scenario, the results of observed power are plotted in same figure for different trial objectives. Moreover, from these plots, same interpretations from Tables 4.1,4.2 and 4.2 are easily obtained.

For cross-over design, we also follow same path of parallel group design. We assume that there is no carry-over effect and interactions between subjects, treatments and periods. The reason of considering these assumptions is using the paired t test approach. In Tables 4.4, 4.5 and 4.6, we give the rate of number of samples whose power is more than 80% with cross-over design under different trial objectives, sample sizes and different distributions.

Table 4.4. Simulation Results in Equality Trials for Cross-over Design

		Equality Trials							
Response of New Drug	Response of Standard Drug	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
X∼N(2.5,2)	Y∼N(2,2)	0.3368	0.5104	0.6153	0.7835	0.8808	0.9391	0.9646	0.9803
X∼Lognormal(μ=0.67,σ=0.70)	Y∼N(2,2)	0.3163	0.4875	0.6057	0.7821	0.8766	0.9399	0.9803	0.9811
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Lognormal(μ=0.35,σ=0.83)	0.3239	0.5071	0.6224	0.7909	0.8805	0.9385	0.9655	0.9811
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Exp(0.5)	0.3183	0.4897	0.6020	0.7810	0.8734	0.9365	0.9682	0.9807

Table 4.5. Simulation Results in Superiority Trials for Cross-over Design

		Superiority Trials							
Response of New Drug	Response of Standard Drug	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
X∼N(2.5,2)	Y∼N(2,2)	0.3108	0.4531	0.5387	0.6985	0.8127	0.8697	0.9108	0.9407
X∼Lognormal(μ=0.67,σ=0.70)	Y∼N(2,2)	0.2903	0.4334	0.5378	0.6998	0.8021	0.8723	0.9132	0.9440
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Lognormal(μ=0.35,σ=0.83)	0.2918	0.4489	0.5484	0.7020	0.8068	0.8700	0.9133	0.9432
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Exp(0.5)	0.2998	0.4427	0.5452	0.7012	0.8036	0.8679	0.9062	0.9415

Table 4.6. Simulation Results in Non-inferiority Trials for Cross-over Design

		Non-Inferiority Trials							
Response o New Drug	Response of Standard Drug	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
X∼N(2.5,2)	Y∼N(2,2)	0.3847	0.6106	0.7164	0.8964	0.9548	0.9815	0.9935	0.9969
X∼Lognormal(μ=0.67,σ=0.70)	Y∼N(2,2)	0.3609	0.5887	0.7247	0.8913	0.9579	0.9847	0.9936	0.9977
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Lognormal(μ=0.35,σ=0.83)	0.3733	0.6093	0.7431	0.8909	0.9554	0.9821	0.9935	0.9977
X∼Lognormal(μ=0.67,σ=0.70)	Y∼Exp(0.5)	0.3751	0.6100	0.7364	0.8951	0.9590	0.9813	0.9936	0.9966

In Tables 4.4, 4.5 and 4.6, same pattern in parallel group is observed. Therefore, when the sample size increases, the likelihood that we get 80% power increases. Similarly, in Table 3.5, we give the minimum required sample sizes in each group to reach 80% power for the scenario we created. Likewise in parallel group design, the minimum required subjects was calculated highest to fewest respectively, in superiority, equality and non-inferiority trials. As a result, in non-inferiority trials, we get 80% power with less number of samples compared to equality and superiority trials. Similarly, we give box plots of simulation results.

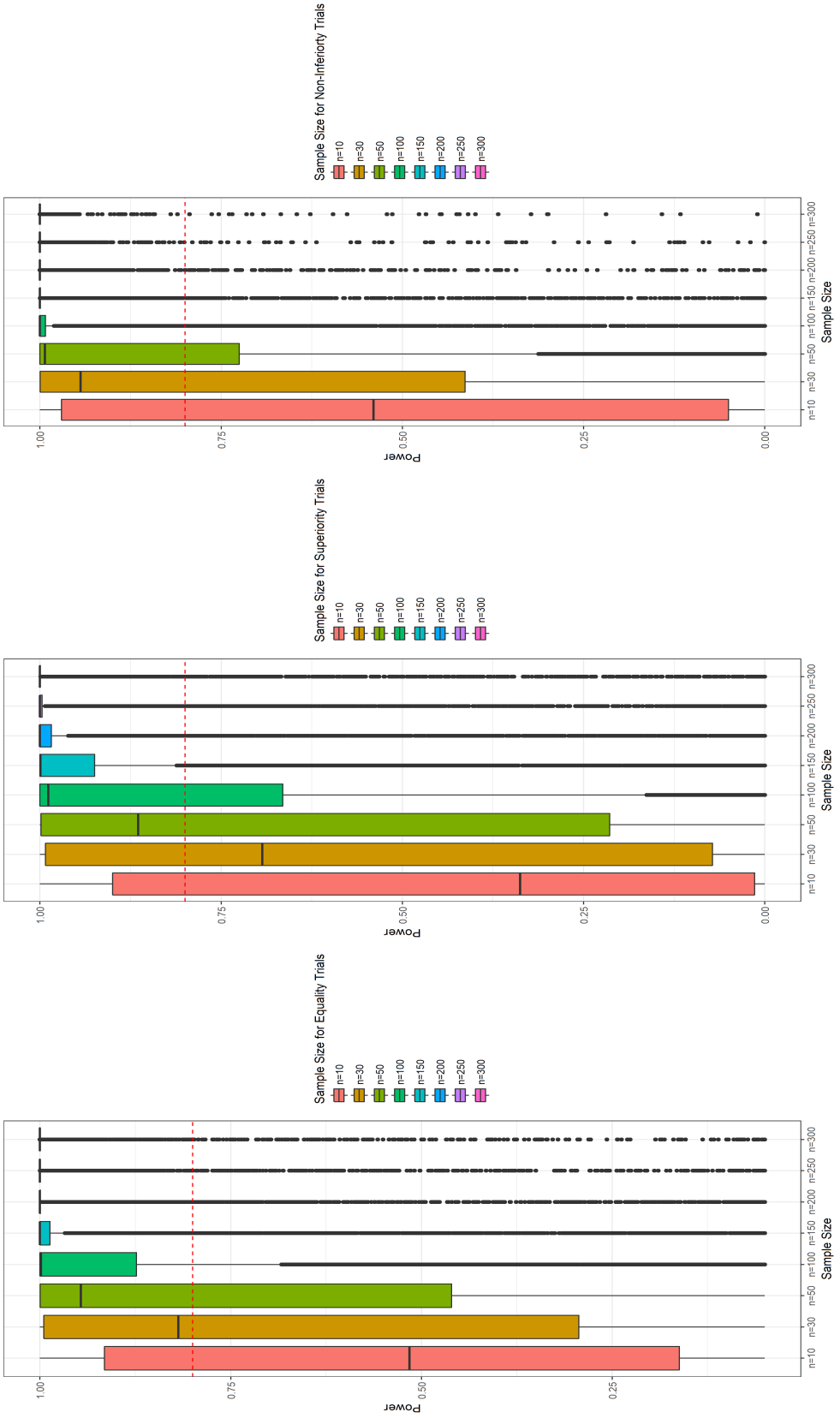

Figure 4.5. Simulation Results for Responses of Both Groups are Normally Distributed for Cross-over Design

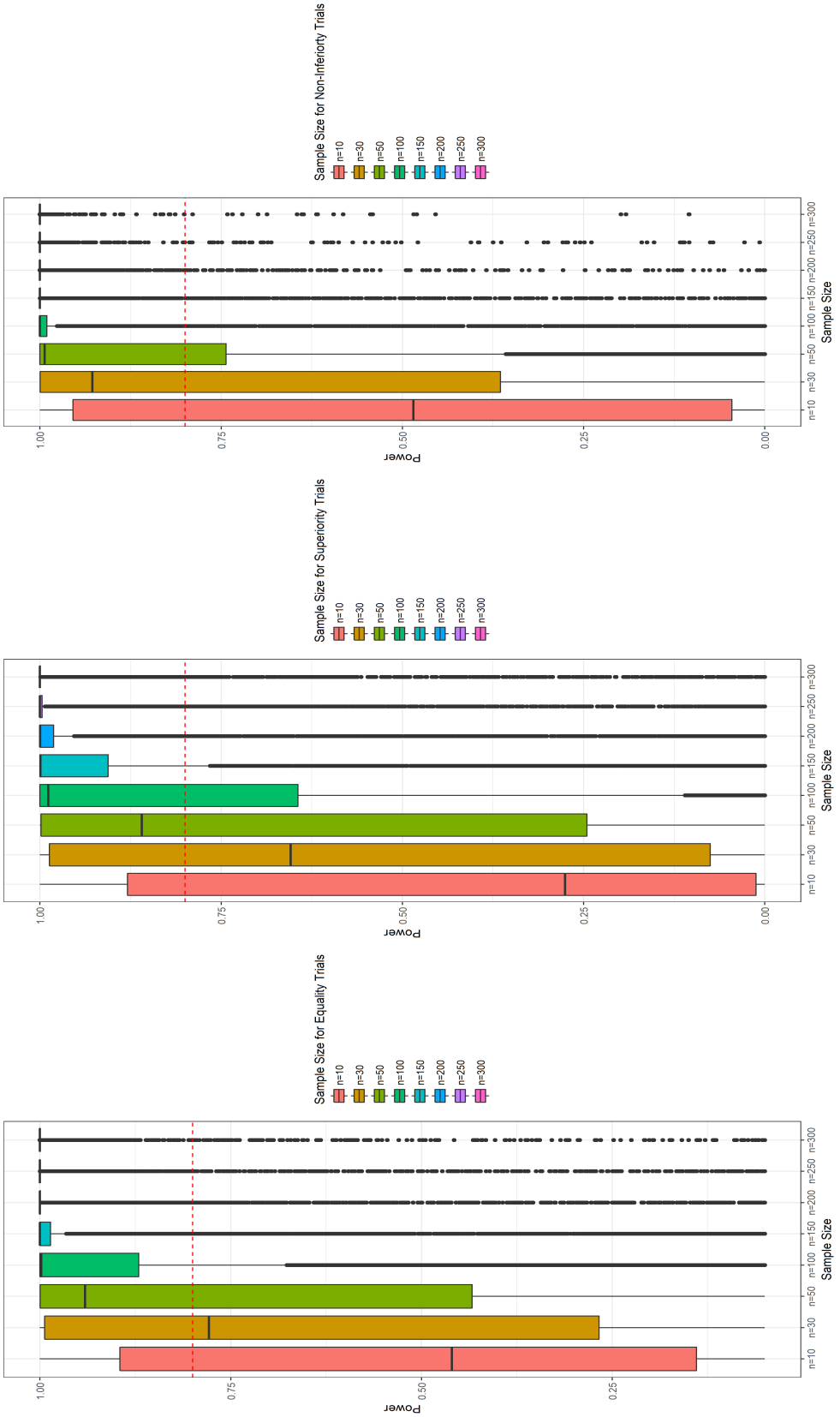

Figure 4.6. Simulation Results for Responses of Groups Have Lognormal and Normal Distributions for Cross-over Design

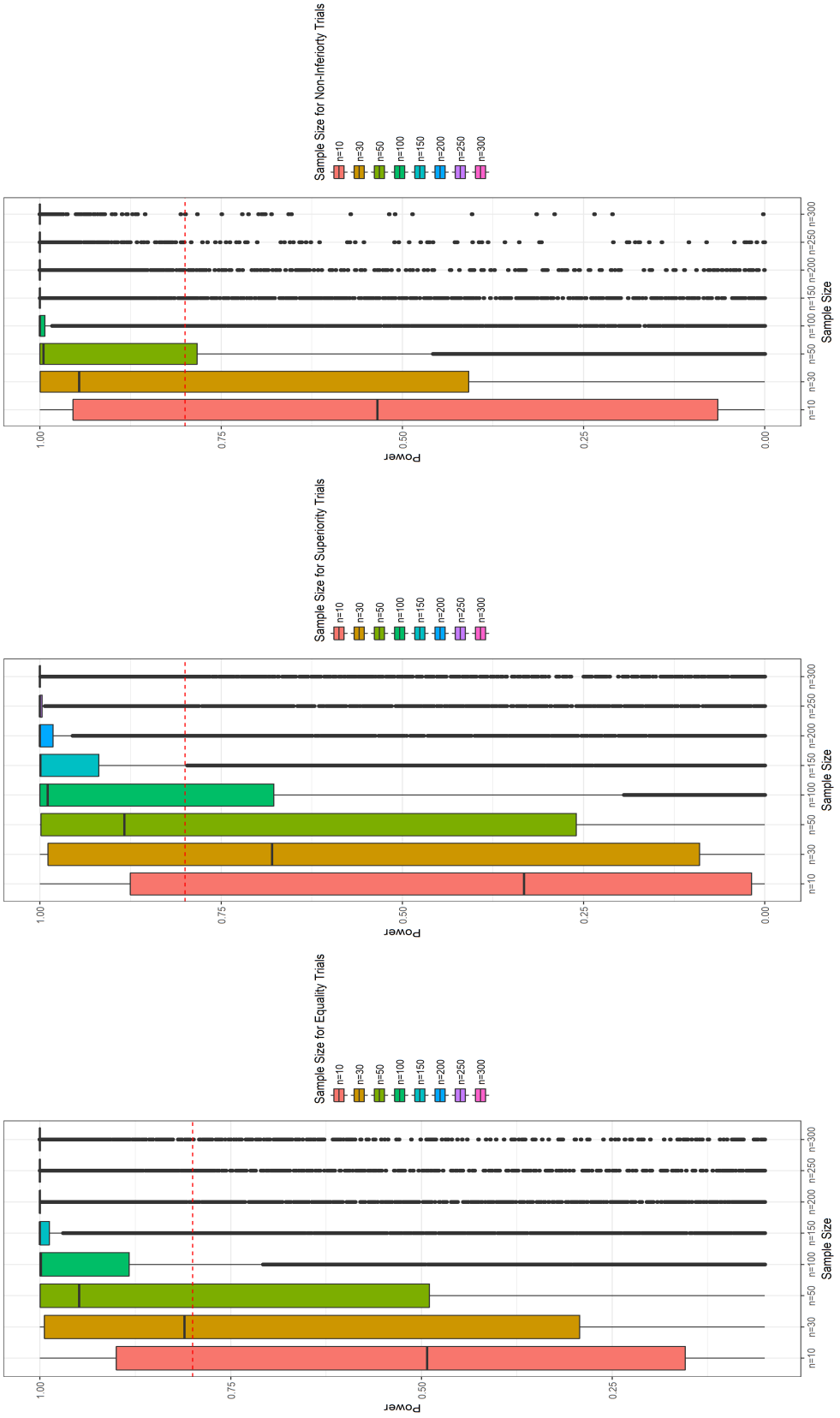

Figure 4.7. Simulation Results for Responses of Both Groups Have Lognormal Distribution for Cross-over Design

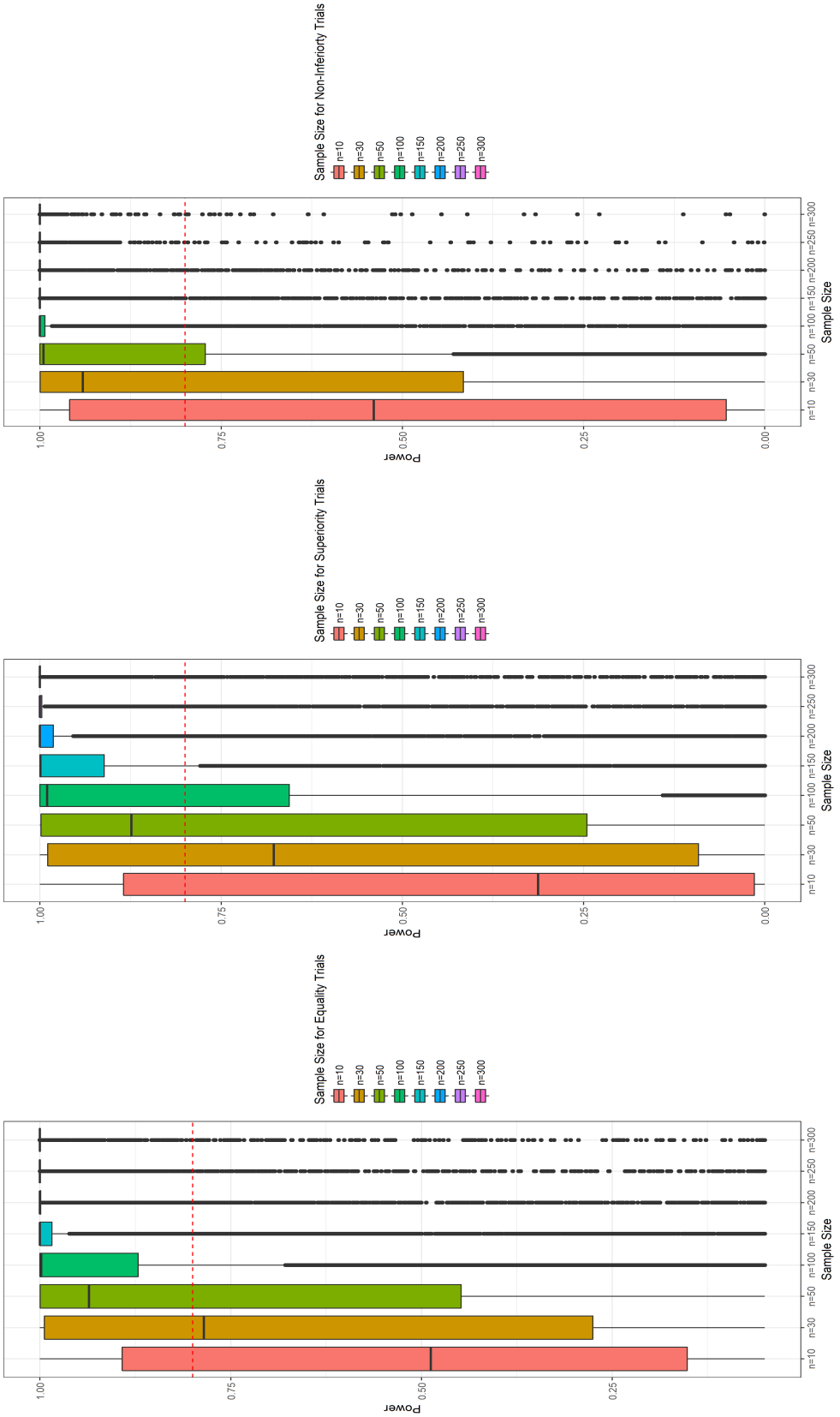

Figure 4.8. Simulation Results for Responses of Groups Have Lognormal and Exponential Distributions for Cross-over Design

From Figures 4.5, 4.6, 4.7 and 4.8, the number of samples reaching over 80% of power is obtained from highest to fewest in non-inferiority, equality and superiority trials, respectively. We can conclude that:

- The observed power is higher in non-inferiority trials compared to equality and superiority trials considering that same clinical important difference, study design, Type I error and sample size.
- The observed power is higher in cross-over design compared to parallel group design with same clinical important difference, Type I error, sample size and trial objectives.
- There is no remarkable difference between distributions of groups. The observed powers are similar in different distribution under same with same clinical important difference, Type I error, sample size, trial objectives and study designs.

4.1.2 Fixed Effect Sizes

In Tables 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 4.13 and 4.14, we show the ratio of number of samples whose power is more than 80% with parallel group and cross-over designs under different trial objectives, sample sizes and different distributions based on specified effect sizes.To get observed power, two sample t test and paired t test approaches are used.

Table 4.7. Simulation Results when Both Groups Normally Distributed in Parallel Group Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.0058	0.0185	0.0322	0.0842	0.1443	0.2116	0.2748	0.3570
	Superiority Trials	0.0142	0.0394	0.0658	0.1379	0.2212	0.3074	0.3923	0.4861
	Non-Inferiority Trials	0.0168	0.0414	0.0687	0.1393	0.2196	0.3112	0.4060	0.4849
d=0.5	Equality Trials	0.0315	0.1818	0.3812	0.7660	0.9365	0.9872	0.9971	0.9994
	Superiority Trials	0.0723	0.2774	0.5001	0.8482	0.9661	0.9948	0.9990	0.9999
	Non-Inferiority Trials	0.0712	0.2861	0.5011	0.8478	0.9681	0.9940	0.9990	0.9999
d=0.8	Equality Trials	0.1160	0.6046	0.8851	0.9975	1.0000	1.0000	1.0000	1.0000
	Superiority Trials	0.2121	0.7156	0.9347	0.9995	0.9999	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.2055	0.7148	0.9339	0.9993	1.0000	1.0000	1.0000	1.0000

Table 4.8. Simulation Results for Responses of Groups Have Lognormal and Normal Distribution in Parallel Group Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.0054	0.0154	0.0281	0.0810	0.1383	0.2139	0.2822	0.3638
	Superiority Trials	0.0079	0.0314	0.0602	0.1383	0.2235	0.3124	0.3957	0.4933
	Non-Inferiority Trials	0.0098	0.0340	0.0636	0.1383	0.2215	0.3166	0.4117	0.4877
d=0.5	Equality Trials	0.0248	0.1785	0.3855	0.7667	0.9342	0.9861	0.9967	0.9991
	Superiority Trials	0.0615	0.2807	0.5079	0.8496	0.9660	0.9940	0.9988	0.9999
	Non-Inferiority Trials	0.0631	0.2912	0.5108	0.8479	0.9658	0.9933	0.9984	1.0000
d=0.8	Equality Trials	0.1113	0.6090	0.8853	0.9968	1.0000	1.0000	1.0000	1.0000
	Superiority Trials	0.2138	0.7233	0.9316	0.9991	0.9999	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.2053	0.7195	0.9336	0.9988	1.0000	1.0000	1.0000	1.0000

Table 4.9. Simulation Results of Responses of both Groups Have Lognormal Distribution in Parallel Group Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.0112	0.0196	0.0333	0.0812	0.1397	0.2105	0.2803	0.3567
	Superiority Trials	0.0167	0.0355	0.0636	0.1388	0.2191	0.3060	0.3967	0.4877
	Non-Inferiority Trials	0.0181	0.0388	0.0648	0.1424	0.2173	0.3126	0.4046	0.4854
d=0.5	Equality Trials	0.0299	0.1774	0.3811	0.7720	0.9346	0.9872	0.9971	0.9992
	Superiority Trials	0.0950	0.3453	0.5800	0.8931	0.9814	0.9978	0.9998	1.0000
	Non-Inferiority Trials	0.0677	0.2865	0.5036	0.8501	0.9657	0.9933	0.9979	1.0000
d=0.8	Equality Trials	0.1139	0.6045	0.8855	0.9975	1.0000	1.0000	1.0000	1.0000
	Superiority Trials	0.2073	0.7199	0.9324	0.9992	1.0000	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.2056	0.7190	0.9347	0.9987	1.0000	1.0000	1.0000	1.0000

Table 4.10. Simulation Results for Responses of Groups Have Lognormal and Exponential Distribution in Parallel Group Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.0113	0.0209	0.0376	0.0835	0.1422	0.2157	0.2783	0.3701
	Superiority Trials	0.0211	0.0434	0.0682	0.1378	0.2267	0.3140	0.4005	0.4835
	Non-Inferiority Trials	0.0181	0.0429	0.0650	0.1374	0.2172	0.3084	0.3954	0.4755
d=0.5	Equality Trials	0.0335	0.1778	0.3682	0.7664	0.9330	0.9871	0.9980	0.9999
	Superiority Trials	0.1138	0.3761	0.6164	0.9180	0.9874	0.9989	0.9998	1.0000
	Non-Inferiority Trials	0.0656	0.2868	0.5037	0.8533	0.9675	0.9946	0.9992	0.9997
d=0.8	Equality Trials	0.1141	0.6003	0.8813	0.9982	1.0000	1.0000	1.0000	1.0000
	Superiority Trials	0.2090	0.7266	0.9351	0.9988	1.0000	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.2073	0.7316	0.9338	0.9996	1.0000	1.0000	1.0000	1.0000

Table 4.11. Simulation Results when Both Groups Normally Distributed in Cross-Over Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.1315	0.2106	0.2585	0.3445	0.4285	0.5043	0.5629	0.6276
	Superiority Trials	0.1503	0.2258	0.2869	0.3975	0.4892	0.5535	0.6200	0.6866
	Non-Inferiority Trials	0.1488	0.2363	0.2868	0.4028	0.4860	0.5644	0.6253	0.6766
d=0.5	Equality Trials	0.2255	0.4806	0.6451	0.8637	0.9533	0.9857	0.9944	0.9983
	Superiority Trials	0.2875	0.5322	0.7034	0.8956	0.9648	0.9901	0.9975	0.9994
	Non-Inferiority Trials	0.2821	0.5375	0.7010	0.8942	0.9668	0.9895	0.9964	0.9992
d=0.8	Equality Trials	0.3810	0.7749	0.9220	0.9945	0.9996	1.0000	1.0000	1.0000
	Superiority Trials	0.4666	0.8120	0.9436	0.9977	0.9998	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.4649	0.8178	0.9422	0.9975	0.9999	1.0000	1.0000	1.0000

Table 4.12. Simulation Results for Responses of Groups Have Lognormal and Normal Distribution in Cross-Over Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.1291	0.2109	0.2631	0.3547	0.4378	0.5077	0.5715	0.6304
	Superiority Trials	0.1397	0.2312	0.2933	0.4009	0.4941	0.5620	0.6226	0.6880
	Non-Inferiority Trials	0.1441	0.2389	0.2973	0.4100	0.4920	0.5678	0.6310	0.6808
d=0.5	Equality Trials	0.2304	0.4882	0.6539	0.8618	0.9481	0.9836	0.9935	0.9978
	Superiority Trials	0.2960	0.5402	0.7064	0.8953	0.9646	0.9898	0.9962	0.9991
	Non-Inferiority Trials	0.2897	0.5536	0.7079	0.8910	0.9653	0.9886	0.9956	0.9993
d=0.8	Equality Trials	0.3955	0.7756	0.9212	0.9933	0.9997	1.0000	1.0000	1.0000
	Superiority Trials	0.4859	0.8190	0.9395	0.9969	0.9998	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.4724	0.8180	0.9419	0.9962	0.9999	1.0000	1.0000	1.0000

Table 4.13. Simulation Results of Responses of both Groups Have Lognormal Distribution in Cross-Over Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.1236	0.2004	0.2513	0.3426	0.4327	0.5066	0.5613	0.6276
	Superiority Trials	0.1349	0.2226	0.2863	0.3890	0.4892	0.5550	0.6229	0.6821
	Non-Inferiority Trials	0.1332	0.2260	0.2801	0.4008	0.4835	0.5625	0.6303	0.6771
d=0.5	Equality Trials	0.2238	0.4780	0.6477	0.8662	0.9502	0.9854	0.9939	0.9982
	Superiority Trials	0.3306	0.6008	0.7636	0.9315	0.9823	0.9963	0.9991	0.9999
	Non-Inferiority Trials	0.2719	0.5450	0.7050	0.8936	0.9653	0.9894	0.9965	0.9992
d=0.8	Equality Trials	0.3796	0.7712	0.9249	0.9945	0.9997	1.0000	1.0000	1.0000
	Superiority Trials	0.4693	0.8191	0.9407	0.9969	0.9997	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.4610	0.8188	0.9434	0.9967	0.9998	1.0000	1.0000	1.0000

Table 4.14. Simulation Results for Responses of Groups Have Lognormal and Exponential Distribution in Cross-Over Design

Effect Size	Trial Objective	n=10	n=30	n=50	n=100	n=150	n=200	n=250	n=300
d=0.2	Equality Trials	0.1242	0.1973	0.2429	0.3359	0.4227	0.5031	0.5612	0.6286
	Superiority Trials	0.1431	0.2252	0.2865	0.3962	0.4891	0.5616	0.6207	0.6823
	Non-Inferiority Trials	0.1329	0.2342	0.2860	0.3952	0.4795	0.5556	0.6244	0.6759
d=0.5	Equality Trials	0.2111	0.4663	0.6337	0.8630	0.9498	0.9856	0.9959	0.9988
	Superiority Trials	0.3539	0.6383	0.7946	0.9515	0.9887	0.9982	0.9995	0.9998
	Non-Inferiority Trials	0.2740	0.5435	0.7055	0.8995	0.9668	0.9892	0.9962	0.9989
d=0.8	Equality Trials	0.3779	0.7682	0.9228	0.9957	0.9996	1.0000	1.0000	1.0000
	Superiority Trials	0.4661	0.8251	0.9420	0.9956	0.9998	1.0000	1.0000	1.0000
	Non-Inferiority Trials	0.8513	0.9972	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

According to fixed effect sizes (Table 4.7, 4.8, 4.9, 4.10, 4.11, 4.12, 4.13 and 4.14):

- There is also no significant differences with different distributions.
- It is not surprising to get that the the lowest observed power is related to small effect size.
- When effect size increases, the likelihood of obtaining 80% of observed power is also increases.
- The observed power is also higher in cross-over design compared to parallel group design.

5 DISCUSSION

According to GCP (4), clinical trial is "*any investigation in human subjects intended to discover or verify the clinical, pharmacological and/or other pharmacodynamic effects of an investigational product(s), and/or to identify any adverse reactions to an investigational product(s), and/or to study absorption, distribution, metabolism, and excretion of an investigational product(s) with the object of ascertaining its safety and/or efficacy. The terms clinical trial and clinical study are synonymous*".

As the volume of medical research increases, the number of conflicting findings and consequences increases. Although the observed differences might show true differences, the results differ from the sampling variability, as all studies are performed on a certain number of subjects. In order to control Type II error or to estimate the precision of trial, calculating sample size becomes inevitable (13).

According to guideline of statistical principals for clinical trials (14) in The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), sample size calculation should be written in detail in a study protocol.

The philosophy behind the sufficiency of sample size is to reduce the possibility that the value or the decision to be reached randomly and increase the likelihood of the real situation in the society (15). As clinical trials are very expensive, consuming a lot time and resources, sample size consideration for clinical trials is more significant. Mickenautsch (16) reported from the systematic review of clinical trials is that one of the reasons imprecision of results of clinical trials is based on small sample size.

Sample size in clinical trials is affected by study design, trial objective, effect size, Type I and Type II errors, variability in population and other factors such as drop-out rate (3).

As it is stated in Chapter 3 and 4, for the calculations of sample size, researchers need to be informed about variability in and clinical meaningful difference in the population. The information about the population can be obtained from a pilot study or investigators' knowledge (17). For pilot studies, as sample size is small, variability could change widely. To overcome instability of sample size, using Bayesian approach to calculate sample size for comparing means is suggested by Wang and his colleges (18). They state that sample size calculation can be considered as a decision problem and use a loss or utility function.

Julious and Owen (19) give another approach for sample size calculation.

This approach enables the estimation of imprecise sample variance for investigators. They indicate that with unknown population variance, sample variance S^2 should be estimated. S^2 with a few degrees of freedom (m<200) have an important effect on sample size. Therefore, if the degrees of freedom of estimated variance is few, then the formulations they suggest can be used.

We consider only one primary end point in this study. However, in some clinical trials, dealing with two or more co-primary end points might satisfy better evaluation. There has been increasing trend about comparing more than one primary end points in pharmaceutical drug research. There exist several methods to deal with co-primary end points. However, performing these methods are not easy due to complicated mathematics and programming (20). In the paper of Sugitima et al. (20), they give useful formulations and tables in parallel group design for two treatments with continuous co-primary end points.

In this thesis, we take known and unknown population variances cases into account. For known population variance cases we used standard normal distribution approach while for unknown population variance cases we used non-central t distribution approach. Moreover, we conduct simulation study to understand the

6 CONCLUSION

"*Statistical analysis allows us to put limits on our uncertainty, but not to prove anything*" (21). As this quote explains, we try to decrease uncertainty in statistics. Sample size calculation gives this chance to investigators at the beginning of study for trying to reduce Type II error.

Determination of sample size is especially crucial in clinical trials. Developing a new treatment or a new drug takes several years from the beginning of discovering a new agenda. Moreover, it is expensive. Economic, ethical and scientific problems might arise in clinical trials if investigators do not create a detailed plan for trials at the beginning of the study. Sample size calculation is one of the steps of clinical trials. There are several factors that play role in calculation of sample size as stated in Chapter 3.

In the light of the information provided, the aim of this thesis is to put an emphasis on significance of sample size calculation for clinical trials. In Chapter 3, we show how to obtain formulations of sample size with regarding to the trial objectives (equality, superiority, etc.) and study designs (parallel group and cross-over) with respect to known and unknown population variance. For known and unknown population variance cases, standard normal distribution approach and non-central t distribution approach are used, respectively. We give numerical examples to make clarify the sample size calculations and show that the sample sizes obtained from standard normal distribution and non-central t distribution approaches are close to each other. We also present how these calculations are calculated by using RStudio. In Chapter 4, we create simulation scenarios under different distributions, trial objectives, sample sizes and specified effect sizes to compare observed power. We present that the observed power is higher in non-inferiority trials than in superiority and equality trials based on same clinical important difference, Type I error, study design and sample size. The observed power is higher in cross-over design compared to parallel group design with same clinical important difference, Type I error, trial objective and sample size. Each subject "serves as her or his own control" and within-subject variability is generally lower than between-subject variability, the increased efficiency might be achieved in cross-over trials. However, cross-over design is not applicable in each clinical trials. For cross-over design, the disease should be chronic and stable and treatments should result alleviate the disease condition (22). Moreover, when the effect size increases, the observed power also increases. The different distributions of responses do not create considerable difference to take into account.

7 BIBLIOGRAPHY

[1] Altman DG. Statistics and ethics in medical research: III How large a sample? Bmj. 1980;281(6251):13368.

[2] Moher D. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA: The Journal of the American Medical Association. 1994;272(2):1224.

[3] Julious SA. Sample sizes for clinical trials. Boca Raton, Fla.: CRC Press; 2010.

[4] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH),*Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice [E6(R2)]*. 2016. Available from: `http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/-Guidelines/Efficacy/E6/E6_R2__Step_4_2016_1109.pdf`

[5] Chow S-C, Liu J-pei. Design and analysis of clinical trials: concepts and methodologies. Hoboken, NJ: Wiley; 2014.

[6] Chow S-C, Shao J, Wang H. Sample Size Calculations in Clinical Research. Boca Raton: Chapman & Hall/CRC; 2008.

[7] Ed Zhang ; Vicky Qian Wu ; Shein-Chung Chow ; Harry G.Zhang (2013). TrialSize: R functions in Chapter 3,4,6,7,9,10,11,12,14,15. R package version 1.3. https://CRAN.R-project.org/package=TrialSize.

[8] RStudio Team (2017). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. http://www.rstudio.com.

[9] Krishnamoorthy K. Handbook of statistical distributions with applications. Boca Raton: CRC Press -Taylor & Francis Group; 2016.

[10] Cohen JW. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1977.

[11] Stephane Champely (2017). pwr: Basic Functions for Power Analysis. R package version 1.2-1. https://CRAN.R-project.org/package=pwr.

[12] Wickham, H. (2009). ggplot2. *Elegant Graphics for Data Analysis*. New York, NY : Springer-Verlag New York, 2009.

[13] Biau DJ, Kernéis S, Porcher R. Statistics in Brief: The Importance of Sample Size in the Planning and Interpretation of Medical Research. Clinical Orthopaedics and Related Research. 2008;466(9):22828.

[14] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH),*Statistical Principles for Clinical Trials [E9]*. 1998. Available from: `http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.`

pdf.

[15] Sümbüloğlu K, Sümbüloğlu V, Güney Z. Klinik araştırmalar: Bilimsel Planlama ve Analiz Yöntemleri. Ankara: Hatiboğlu Yayınları; 2007.

[16] Mickenautsch S. Research gaps identified during systematic reviews of clinical trials: glass-ionomer cements. BMC Oral Health. 2012;12(1).

[17] Chow S-C. Controversial statistical issues in clinical trials. Boca Raton: CRC Press; 2011.

[18] Wang H, Chow S-C, Chen M. A Bayesian Approach on Sample Size Calculation for Comparing Means. Journal of Biopharmaceutical Statistics. 2005;15(5):799807.

[19] Julious SA, Owen RJ. Sample size calculations for clinical studies allowing for uncertainty about the variance. Pharmaceutical Statistics. 2006;5(1):2937.

[20] Sugimoto T, Sozu T, Hamasaki T. A convenient formula for sample size calculations in clinical trials with multiple co-primary continuous endpoints. Pharmaceutical Statistics. 2012;11(2):11828.

[21] Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1991.

[22] Piantadosi S. Clinical trials: a methodologic perspective. Hoboken: Wiley; 2017.

8 CURRICULUM VITAE

Ebru Öztürk

CONTACT INFORMATION
Hacettepe Üniversitesi Tıp Fakültesi
Biyoistatistik Ana Bilim Dalı, 06100 Sıhhiye Ankara ebru.ztrk3@gmail.com

RESEARCH INTERESTS
Biostatistics, fMRI data, clinical trials

EDUCATION

Middle East Technical University, Ankara

M.S., Statistics, September 2017

- Thesis Topic: *Statistical Disease Detection with Resting State Functional Magnetic Resonance Imaging*

B.S., Statistics, June 2014

RESEARCH EXPERIENCE

Research Assistant October 2016 to present
Division of Biostatistics,
Hacettepe University

Research Assistant November 2015 to October 2016
Department of Statistics,
Middle East Technical University

CERTIFICATES, COURSES AND INTERNSHIPS

Courses

- Structural Equation Modelling and It's Applications — Oct 2017
- Summer School on Optimization, Big data and Applications (OBA) — July 2017
- An Introduction to Joint Models for Longitudinal and Survival Data, with Applications in R — May 2017
- Bioinformatic with R — Oct 2016
- fMRI Basic Analysis and Connectivity Analysis Course with SPM8 — July 2016

Certificates and Internships

- Central Bank of the Republic of Turkey in Workers' Remittances Department — September 2013
- Ankara University in Biostatistics Department — July 2013
- Republic of Turkey Prime Ministry Undersecretariat of Treasury in Insurance Department — June 2013
- Participation of education seminar in Capital Markets Board of Turkey — June 2013
- Organization of "New Developments in Theory and Applications of Statistics"(NEDETAS) — May 2011
- Participation of NEDETAS — May 2011

PRESENTATIONS

Statistical Meetings

- Ozturk, E., Ilk, O. Classifying Resting State Functional Magnetic Resonance Imaging Data 9th EMR-IBS and Italian Region conference, Thessaloniki, Greece 8-12 May 2017.

HARDWARE AND SOFTWARE SKILLS

Computer Programming:

- C, R, SAS, MATLAB, Minitab, SPSS, LaTeX