國立臺灣大學生物暨農學院農藝學研究所

生物統計學組碩士論文

Division of Biometry Graduate Institute of Agronomy

College of Bioresources and Agriculture

National Taiwan University

Master thesis

平行檢定法於生物製劑學名藥

對等性評估之應用

On Study of Applications of the Parallel Line Assays

to Evaluation of Biosimilar Drug Products

林亞靚

Ya-Ching Lin

指導教授：劉仁沛 博士

Advisor: Jen-Pei Liu, PH.D.

中華民國 99 年 6 月

June, 2010

誌謝

　　轉眼間兩年的碩士生涯就到了尾聲。還記得當初從一個數學的領域轉到生物統計那種慌張的感覺還歷歷在目，要謝謝很多人幫助我熟悉這個環境。其中最感謝劉仁沛老師的教導，在寫論文的過程有很多不懂的地方，但老師總是很有耐心的一次又一次的講解，使我的論文可以順利完成。不只在學業上老師也常常關心學生的生活，提供我們良好的學習環境，能夠讓劉仁沛老師指導，真的是件很幸運的事情。另外也要感謝三位口試委員：季瑋珠老師、林志榮老師和謝宗成老師，三位老師在論文上給了很多寶貴的意見，使我的論文能更加完整。

　　感謝完師長也要謝謝研究室的同學們。感謝詩婷學姐、瑱芳學姐在程式上給了我很大的幫助。還有同窗的孟陽、證伊一起討論作業一起互相打氣。靖瑩、妍凌還有畢琳謝謝你們總是陪我分享所有的喜怒哀樂，在我不順心的時候替我加油，快樂的時候陪我開心，很謝謝有你們的陪伴!還有雅純學姐、西閔學長、姜杰學長、柏志學長、志熙、書弘有了你們讓我的研究生涯更加精采。

　　最後要感謝我的家人，謝謝爸爸媽媽讓我沒有後顧之憂的一路念到研究所，給我最大的支持。還有承叡，如果沒有你一路上的支持與鼓勵，我想我很難走到這一步吧!當我遇到問題時，謝謝你總是幫我想很多的方法陪著我一起解決問題，當我的精神支柱。

要感謝的人太多了，真的很謝謝大家。我愛你們!

English Abstract

In recent years, biologics market increases rapidly, but the development costs are also very high. Therefore, after the patent of biological products expires, many pharmaceutical companies have invested in the development of biosimilar drug products. But biological products are different from traditional small molecule drug products. Therefore the methods for assessment of biosimilar drug products are also different. Innovators ask for clinical trials to provide the efficacy and safety data to approve biosimialr drug products. However if approval of biosimilar products requires clinical trials, then development cost of biosimilar products will be the same as that of the innovators. As a result, it cannot achieve the goal of cost reduction. Therefore we try to apply the method of the parallel line assay to test whether the approval of the biosimilar should require clinical trials and to evaluate the similarity between the biosimilar products and innovator's biological products. The results of type I error and power from the simulation studies are presented. A numerical example is used to illustrate the application of the proposed method.

Keywords: Biological products, Biosimilar, Parallel line assay, Bioequivalence, Extrapolation

中文摘要

　　近年來生物製劑的市場逐漸增加，但所開發的成本太高。所以在生物製劑的專利期到期之後，許多藥廠紛紛投入生物製劑學名藥的開發。但生物製劑有別於一般傳統的小分子學名藥，所以評估生物製劑學名藥的方法可能不同。生技原廠認為生物製劑學名藥須執行臨床試驗才能被核准。但若生物製劑學名藥須執行大型臨床試驗，則生物製劑學名藥的研發則與發展新的生物製劑相同，而不能達到降低成本造福病患的目的。所以本論文以平行檢定的方法評估生物製劑學名藥是否必須進行大型臨床試驗及提出生物製劑學名藥與原廠生物製劑藥品相似的方法，執行模擬評估所提出方法之第一型錯誤機率和檢定力，並以數值例子來介紹提出方法之應用。

關鍵字: 生物製劑，生物製劑學名藥，平行檢定，生物相等性，外插

Contents

List of Tables

List of Figures

Chapter 1 Introduction

1.1 Background and Motivation

Biological drug products are therapeutic moiety manufactured by a living system or organism, such as human, plants, animals or microorganisms, and they are used for medical, prevention and cure of human diseases. It can be applied to cancer, diabetes mellitus, growth impairment, and many other diseases. Many biologic drug products are important life-saving drug products for patients with unmet medical needs. The worldwide sales of biological products in 2008 reached US $125 billion dollars which accounts about 20% of the pharmaceutical industry.

In addition, unlike the chemical drug products, biological drug products are proteins or peptides with much larger molecular weights. In other words, biological drug products are organic compounds made of amino acids, which possess primary, secondary, tertiary, and quaternary structures. Biologic products are sensitive to heat, light and shock, and it can easily be contaminated (Keith O, 2007). Therefore, the manufacture of biological agents has to be noted, such as types of bacteria, the cells to create the protein, and the quality of production equipments. Therefore the biological

drugs are very expensive and can reach up to US $ 100,000 per patient per year. This high cost denies the accessibility of most of patients to the most of life-saving biological products.

Many best-selling biological products in recent years will face patent expiration. Table 1.1 shows that several biological products have reached the end of their patent protection. Therefore, generic versions of the innovator's biological products can be made available. These generic versions of the innovator's biological products are called biosimalar products by the European Union and follow-on products in the United States. In 2008, the US Congressional Budget Office predicted a savings of US $25 billion dollars in the next 10 years. It is therefore hoped that due to competition and avoidance of unnecessary drug testing and clinical trials, the price of the biosimilar drug products can be reduced and become affordable for most needed patients.

Biosimilars or follow-on biologics are not like the more common small-molecule chemical drugs. Biologics generally exhibit high molecular complexity and may be quite sensitive to manufacturing process changes. Table 1.2 provides comparisons of the differences between small-molecule chemical generics and biosimilar products

Biosimilar products and small molecule generics are different. Therefore, assessment of biosimilar products may also be different. The innovators of the biological products suggest that biosimilar products should not be approved without

data from large clinical trials. However if approved of biosimilar products requires clinical trials, then development cost of biosimilar products will be the same as that of the innovators. As a result, it cannot achieve the goal of cost reduction to benefit the patients.

Therefore, we try to apply the method of parallel line assay to verify whether the clinical trials are required for approval of the biosimilar products and to assess the equivalence between the innovator and biosimilar products if validated characteristics of the drug products are reliable predictors of clinical responses.

1.2 Bioequivalence

For small molecular chemical compounds, the measure of bioavailability include the peak concentration (C_{max}) and the area under the concentration – time curve (AUC). The assessments of traditional chemical generic drugs are based on pharmacokinetics measures and the approval of traditional chemical generic drugs is based on the following Fundamental Bioequivalence Assumption (Chow and Liu, 2008)

When two drug products are equivalent in the rate and extent to which the active drug ingredient or therapeutic moiety is absorbed and becomes available at the site of drug action, it is assumed that they will be therapeutically equivalent and can be used interchangeably.

Next, we briefly review the regulations for approval of the traditional chemical generic

drug products of small molecules.

Europe

According to European regulations two medicinal products are bioequivalent if they are pharmaceutically equivalent or pharmaceutical alternatives and if their bioavailability after administration in the same molar dose is similar to such a degree that their effects, with respect to both efficacy and safety, will be essentially the same. This can be demonstrated if the 90% confidence intervals of the means ratios between the two preparations based on C_{max} and AUC lie in the range of (0.80, 1.25) or (-0.2231, 0.2231) for the log-transformed data.

United States

The FDA considers two products are bioequivalent if based on C_{max}, $AUC_{(0-t)}$ and $AUC_{(0-\infty)}$, the 90% CI of the ratio of the means of the test (e.g. generic formulation) to the reference (e.g. innovator brand formulation) is within 80.00% to 125.00%.

Taiwan

The Department of Health of Taiwan also has similar regulations for approval of generic products with respect to C_{max}, $AUC_{(0-t)}$ and $AUC_{(0-\infty)}$ based on the logarithm scale. If the 90% confidence interval of the mean ratio between generic and innovator of products is within (0.8, 1.25), then the generic drug product is claimed to the bioequivalent to the innovator's drug product.

Since approval of the traditional chemical generic drugs is not require conducting expensive and large clinical trials, the costs of chemical generic drugs are quite inexpensive, usually 1/5 to 1/2 of the innovator products. Therefore many chemical generic drugs become affordable to many needed patients. On the other hand, the biological products and traditional generic drug are fundamentally different. Currently, there is no regulation to approve the follow-on products in the United States. The European Union requires the efficacy and safety data on clinical trials to approve biosimilar products. Therefore it cannot reduce the cost of biosimailar products, and cannot achieve the goal of reducing the costs to the needed patients. The main reason is that even if characteristics of the biosimilar products are equivalent to those of innovator products, it does not guarantee that the efficacy and safety of the biosimilar products are similar to those of the innovator biological products.

Table1.1: Patent status of leading biopharmaceuticals

Product	Active Substance	Company	Patent Expiry (year)	Worldwide sales (2005 in $million)
Cerezyme	imiglucerase	Genzyme	2001	933
Humulin	human insulin	Eli Lily	2001	1005
Novolin	human insulin	Novo Nordisk	2001	1618
Intron-A	interferon alpha2b	Schering-Plough	2002	287
Avonex	interferon-beta	Biogen Idec	2003	1543
Humatrope	somatropin	Eli Lily	2003	414
Nutropin	somatropin	Genentech	2003	370
Procrit	erythropoietin	J&J	2004	3324
Epogen	erythropoietin	Amgen	2004	2455
Neupogen	figrastim	Amgen	2006	1216
			Total	**13,165**

Source: Generic Pharmaceutical Association and Company Reports(2005)

Table1.2: The differences between small-molecule generics and biosimilar products.

	Small molecule generics	Biosimilars
Product characteristics	• Small molecules • Often very stable • Mostly without a device	• Large, complex molecules • Stability requires special handling • Device is often a key differentiator
Production	• Produced by chemical synthesis	• Produced in living organisms • Highly sensitive to manufacturing changes • Often comparatively high costs
Regulation	• Abbreviated registration procedures in Europe and US • Usually enjoy "substitutability" status	• Regulatory pathway now defined by EMEA • "Comparability" status • No pathway yet in US under BLA
Marketing	• No or limited detailing to physicians • Key role of wholesalers and payors • Market substitution in pharmacies • High price discounts	• Detailing to (specialist) physicians required • Pharmacists may not substitute • Price discounts smaller; price sensitivity is product specific

http://www.sandoz.com/site/en/product_range/more_about_biosimilars/index.shtml

Chapter 2 Literature Review

In this chapter, we briefly review the method of assessment of the bioequivalence of traditional chemical generic drugs and the parallel line assay.

2.1 Interval Hypotheses

The assessment of average bioequivalence is based on the comparison of bioavailability profiles between formulations. However, in practice, it is recognized that no two formulations will have exactly the same bioavailability profiles. Therefore, if the profiles of the two formulations differ by less than a (clinically) meaningful limit, the profiles of the two formulations may be considered equivalent. Following this concept, Schuirmann (1981) first introduced the use of interval hypotheses for assessing average bioequivalence.

Let μ_T be the average bioavailability of the test (T) formulation and μ_R the average bioavailability of the reference (R) formulation. The interval hypotheses for average bioequivalence on the log-scale can be formulated as

$$
\begin{cases}
H_0\colon \mu_T - \mu_R \leq \theta_L \ \text{ or } \ \mu_T - \mu_R \geq \theta_U \\
\text{vs.} \\
H_a\colon \theta_L < \mu_T - \mu_R < \theta_U
\end{cases}
\tag{2.1}
$$

Where $\theta_L = -\theta_U = -0.2231$ are some clinically meaningful limits.

The concept of interval hypotheses Eq. (2.1) is to show average bioequivalence by rejecting the null hypothesis of inbioequivalence.

Schuirmann's Two One-Sided Tests Procedure

The interval hypotheses (2.1) can be decomposed into two sets of one-sided hypotheses

$$H_{01} : \mu_T - \mu_R \leq \theta_L \text{ vs. } H_{a1} : \mu_T - \mu_R > \theta_L$$
$$\text{and} \tag{2.2}$$
$$H_{02} : \mu_T - \mu_R \geq \theta_U \text{ vs. } H_{a2} : \mu_T - \mu_R > \theta_U$$

The first set of hypotheses is to verify that the average bioavailability of the test formulation is not too low, whereas the second set of hypotheses is to verify that the average bioavailability of the test formulation is not too high. A relatively low (or high) average bioavailability may refer to the concern of efficacy (or safety) of the test formulation. If one concludes that $\theta_L < \mu_T - \mu_R$ (i.e., reject H_{01}) and $\mu_T - \mu_R < \theta_U$ (i.e., reject H_{02}), then it has been concluded that

$$\theta_L < \mu_T - \mu_R < \theta_U$$

μ_T and μ_R, thus, are equivalent. The rejection of H_{01} and H_{02}, which leads to the conclusion of average bioequivalence, is equivalent to rejecting H_0 in Equation 2.1. Schuirmann (1981, 1987) first introduced the two one-sided tests procedure based on Equation 2.2 for assessing average bioequivalence between formulations. The proposed two one-sided tests procedure suggests the conclusion of equivalence of μ_T and μ_R at

the a level of significance if, and only if, H_{01} and H_{02} in Equation 2.2 are rejected at a predetermined a level of significance. Under the normality assumptions, the two sets of one-sided hypotheses can be tested with ordinary one-sided t tests. We conclude that μ_T and μ_R are average equivalent if

$$T_L = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_L}{\hat{\sigma}_d \sqrt{\dfrac{1}{n_R} + \dfrac{1}{n_T}}} > t(\alpha, n_R + n_T - 2)$$

and

$$T_U = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_U}{\hat{\sigma}_d \sqrt{\dfrac{1}{n_R} + \dfrac{1}{n_T}}} < -t(\alpha, n_R + n_T - 2)$$

where $\bar{Y}_T, \bar{Y}_R$ and $\hat{\sigma}_d$ are the least squares estimator of the test and reference means and square root of error mean square.

The two one-sided t tests procedure is operationally equivalent to the classic (shortest) confidence interval approach; that is, if the classic $(1-2\alpha)100\%$ confidence interval for μ_T and μ_R is within (θ_L, θ_U), then both H_{01} and H_{02} are also rejected at the α level by the two one-sided t tests procedure (Chow and Liu, 2008).

2.2 Two-by-Two Crossover Design

Several useful designs for assessing bioequivalence in variety of situations were discussed in literature (Chow and Liu, 2008). The most commonly used statistical design for comparing average bioavailability between two formulations of a drug

probably is a two-sequence, two-period, crossover design, we shall refer to this design as the standard 2×2 crossover design. For the standard 2×2 crossover design, each subject is randomly assigned to either sequence RT or sequence TR at two dosing periods. In other words, subjects within RT (TR) receive formulation R(T) at the first dosing period and formulation T(R) at the second dosing period. The dosing periods are separated by a washout period of sufficient length for the drug received in the first period to be completely metabolized or excreted from the body. An example of a 2×2 crossover design is illustrated in Figure 2.1.

Because the biological products and traditional small molecular chemical generic drugs are fundamentally different, therefore the methods for evaluation of bioequivalence based on pharmacokinetics measures are not applicable to assess the biosimilar products. As a result, we propose to employ the parallel line assay to evaluate biosimilar products. In the next two sections, biological assay and parallel line assay will briefly reviewed.

2.3 Biological assay

Bioassay is a preparation (such as drugs, hormones, etc.) added to the subject. It is the method to determine the nature, constitution, and potency of the preparation by the reaction of subjects. The preparation added to the subject can have a single dose, or

several different doses. Therefore the subjects produce a response value or multiple

response values, we can test the potency of the preparation by these values. The purpose

of testing is usually to compare the relative potency of two or more preparations.

So we always take a common nature of the preparation known as standard

preparation, while the other is the unknown nature of the newly developed preparation,

referred to as test preparation. The two preparations produce the difference between

response values in the same dose, i.e., for inference of the potency of the test

preparation. Generally we use the indirect assay. Indirect assay consists of quantitative

responses and quantal responses. Indirect assay with quantitative responses also

involves parallel line assay and slope ratio assay.

For our current research, we propose to use the parallel line assay to evaluation of

the similarity of biosimilar products and its innovator's biological products. In next

section we briefly reviewed parallel line assay.

2.4 Parallel Line Assay

Liu and Chow (2010) refer to group means of a well-defined product characteristic

can be computed for each dose level for the biosimilar and innovator's product

respectively. Using the group means of the well-defined characteristic as the

independent variable, a simple linear regression equation can be fitted to primary

efficacy endpoint for the biosimilar and innovator's biological products, respectively. It follows that the concept of relative potency in the parallel-line bioassay can be then employed to investigate extrapolability of equivalence in product characteristic to equivalence in efficacy.

Parallel line assay is to compare potency of two preparations under a linear regression model. The first step is to test for a linear regression model. If the relationship between the product characteristic and clinical responses are linear, then the second step is to test whether regression lines are parallel. In the setting, the innovative biological product is treated as the standard preparation, and the biosimilar product is referred to as the test preparation.

Let the mean characteristics at different doses be the independent variable (x), and clinical efficacy response be the dependent variable (Y). Then the regression model of the reference preparation is given as

$$E(Y_R) = \alpha + \beta z^{\lambda} = \alpha + \beta x \ . \tag{2.3}$$

Where α, β, λ are unknown constants, and $x = z^{\lambda}$

When λ=0, x = logz ,

$$E(Y_R) = \alpha + \beta \log z = \alpha + \beta x. \tag{2.4}$$

Under the same reaction, the ratio of the two preparation doses is called the relative potency (ρ)

$$\rho = \frac{Z_S}{Z_T} \ , \ Z_S = \rho Z_T \ .$$

It follows that

$$\begin{aligned} E(Y_T) &= \alpha + \beta(\rho z)^\lambda = \alpha + \beta \rho^\lambda z^\lambda \\ &= \alpha + \beta \rho^\lambda x \ . \end{aligned}$$

(2.5)

When λ=0

$$\begin{aligned} E(Y_T) &= \alpha + \beta(\log \rho + \log z) = \alpha + \beta \log \rho + \beta \log z \\ &= \alpha' + \beta x \ . \end{aligned}$$

(2.6)

It is also a linear function. Therefore the regression line for the response preparation in (2.4) and that for the test preparation in (2.6) have same slope (β) but different intercepts.

2.4.1 Testing linearity of a regression model

For lack of fit test for a linear regression model, multiple efficacy responses must observed for each mean product characteristic. The data structure for testing linearity is given Table 2.1.

The corresponding ANOVA table for lack of fit test is given in Table 2.2.

where $n = \sum n_i$, $y_{..} = \sum\sum y_{ij}$, $T_i = \sum_{j=1}^{n_i} y_{ij}$, $x_. = \sum n_i x_i$

$$S_{xx} = \sum_{i}^{k} n_i x_i^2 - \frac{x_.^2}{n}$$

$$S_{xy} = \sum\sum x_i y_{ij} - \frac{x_. y_{..}}{n}$$

$$S_{yy} = \sum\sum y_{ij}^2 - \frac{y_{..}^2}{n}$$

The F value of lack of fit test is $F_0 = \frac{MSL}{MSP}$

14

If $F_0 < F_{\alpha,k-2,n-k}$, then we fail to reject the null hypothesis of no lack of fit.

2.4.2 Testing for Parallel Lines

Next, we want to know whether two linear regression lines are parallel. In other words we need to check whether the slopes are equal between the two regression lines. The linear regression models for the test and reference preparation are given as respectively.

$$Y_R = \alpha_R + \beta_R x_R + \varepsilon_R \ ,$$

and

$$Y_T = \alpha_T + \beta_T x_T + \varepsilon_T \ .$$

The corresponding estimated linear regression equations are given as respectively.

$$\hat{Y}_R = a_R + b_R x_R \quad and \quad \hat{Y}_T = a_T + b_T x_T \ ,$$

$$where \quad b_R = \frac{_R S_{xy}}{_R S_{xx}} \ , \quad a_R = \overline{y}_R - b_R \overline{x}_R \ .$$

$$b_T = \frac{_T S_{xy}}{_T S_{xx}} \ , \quad a_T = \overline{y}_T - b_T \overline{x}_T \ .$$

$_R S_{xy}$ and $_T S_{xy}$ are the connected sums of cross products between mean characteristic sand clinical responses for reference and test preparations, similarly $_R S_{xx}$ and $_T S_{xx}$ are the sums of squares of the mean characteristics of the reference and test preparations, $\overline{y}_R$ and $\overline{y}_T$ are the average clinical responses for the reference and test preparations, and $\overline{x}_R$ and $\overline{x}_T$ are the average of the mean characteristics for the reference and test preparations, respectively.

The hypothesis of equal slopes is given as

$$H_0 : \beta_R = \beta_T \ v.s. \ H_a : \beta_R \neq \beta_T .$$

Let $S_R{}^2$ and $S_T{}^2$ be the residue mean squares for the reference and test preparations respectively as follows

$$S_R^2 = \frac{{}_R Syy - b_R \cdot {}_R Sxy}{n_R - 2} , \ and$$

$$S_T^2 = \frac{{}_T Syy - b_T \cdot {}_T Sxy}{n_T - 2} .$$

If $F = \frac{\max\{S_R^2, S_T^2\}}{\min\{S_R^2, S_T^2\}} < F(\alpha, n_R - 2, n_T - 2)$

We fail to reject the null hypothesis of homogeneity of the error variances between the reference and test formulations.

If the residual variances of the reference and test preparation are homogeneous, the test statistic for equality of the two slopes is given as

$$t = \frac{b_R - b_T}{\sqrt{V(b_R - b_T)}} , \tag{2.7}$$

$$where \ \hat{V}(b_R - b_T) = \hat{V}(b_R) + \hat{V}(b_T) = S_p^2 (\frac{1}{{}_R Sxx} + \frac{1}{{}_T Sxx}) ,$$

$$and \ S_p^2 = \frac{(n_R - 2)S_R^2 + (n_T - 2)S_T^2}{n_R + n_T - 4}$$

If $t > t(\alpha/2, n_R + n_T - 4)$, the null hypothesis of equal slope is rejected at the α significance level.

If the null hypothesis of equal slopes is not rejected at the α significance level, then the common slope is estimated as:

$$b_c = \frac{{}_R Sxy + {}_T Sxy}{{}_R Sxx + {}_T Sxx}$$

It follow that estimated linear regression lines for the reference and test preparation are given as respectively,

$$\hat{y}_R = a_R + b_c x_R$$
$$\hat{y}_T = a_T + b_c x_T$$

2.4.3 Relative Potency

Assume that the test preparation is more potent than the reference preparation, then the regression line of the test preparation is above that of the clinical reference preparation. When responses of two drugs are the same, let the dose of the test preparation be x_0.

Denote Δ as the horizontal distance between the two regression lines. Then x_0 dose of the test preparation effect is the same with the $x_0 + \Delta$ dose of the reference preparation effect under the $\log_{10}$ scale. Therefore the 10^Δ dose of the test preparation effect is the same with the $10^{x_0+\Delta} = 10^{x_0 + \log R}$ dose of the reference preparation effect under the original scale. $10^{x_0 + \log R} = 10^{x_0} \cdot 10^{\log_{10} R} = R \cdot 10^{x_0}$, $where\ \Delta = \log_{10} R$. The left hand side of the above equation is the dose of test preparation where the right hand side of the above equation is the dose of reference preparation multiplied by R. In other words, the potency of the test preparation is R times of the reference preparation.

An estimate of Δ is given as

$$\hat{\Delta} = \frac{a_R - a_T}{b_c} = \overline{x}_T - \overline{x}_R + \frac{\overline{y}_R - \overline{y}_T}{b_c}$$

$$\hat{R} = anti\log_{10}\hat{\Delta} = 10^{\hat{\Delta}}$$

The confidence interval of Relative Potency

Let (x_1, x_2) is bivariate normal distribution. Its expectation is (μ_1, μ_2) with variance $(\sigma^2 v_{11}, \sigma^2 v_{22})$, and covariance $\sigma^2 v_{12}$（沈明來, 2000）.

If x_1, x_2 are the sample mean, then $v_{11} = \frac{1}{n_1}, v_{22} = \frac{1}{n_2}$ 。

If x_1, x_2 are the regression slopes, then $v_{11} = \frac{1}{\sum(x_1 - \overline{x}_1)^2}, v_{22} = \frac{1}{\sum(x_2 - \overline{x}_2)^2}$

Let $\theta = \frac{\mu_1}{\mu_2}$，its estimator is $\hat{\theta} = \frac{x_1}{x_2}$，and S^2 is an estimate of σ^2

Let $x_1 - \hat{\theta}x_2$. The variance of $x_1 - \hat{\theta}x_2$ is given as

$$Var(x_1 - \hat{\theta}x_2) = \sigma^2(v_{11} - 2\theta v_{12} + \theta^2 v_{22})$$

It follows that

$$x_1 - \hat{\theta}x_2 \sim N\left(0, \sigma^2(v_{11} - 2\theta v_{12} + \theta^2 v_{22})\right)$$

Therefore $\dfrac{x_1 - \hat{\theta}x_2}{\sqrt{S^2(v_{11} - 2\theta v_{12} + \theta^2 v_{22})}}$, follows a central t-distribution with $n_R + n_T = 4$

degrees of freedom.

The probability statement of constructing an 100(1-α)% confidence interval for θ is

given as

$$\Pr\left\{\frac{\left(x_1 - \hat{\theta} x_2\right)^2}{S^2(v_{11} - 2\theta v_{12} + \theta^2 v_{22})} \leq t_{\alpha/2}^2\right\} = 1-\alpha$$

Where $t_{\alpha/2}$ is the upper α/2 quantifile of a central distribution with $n_R + n_T - 4$

degrees of freedom.

It turns out that it is a quadratic equation of $\hat{\theta}$ as expressed in the following

$$\hat{\theta}^2(x_2^2 - t_{\alpha/2}^2 S^2 v_{22}) - 2\hat{\theta}(x_1 x_2 - t_{\alpha/2}^2 S^2 v_{12}) + (x_1^2 - t_{\alpha/2}^2 S^2 v_{11}) \leq 0$$

The two roots of the above quadratic equation constitute the lower and upper limits of

the 100(1-α)% confidence interval for θ as

$$\hat{\theta} = \frac{1}{x_2^2 - t^2 S^2 v_{22}}\left\{(x_1 x_2 - t^2 S^2 v_{12}) \pm \sqrt{(x_1 x_2 - t^2 S^2 v_{12})^2 - (x_2^2 - t^2 S^2 v_{22})(x_1^2 - t^2 S^2 v_{11})}\right\}$$

It can also be rearranged as

$$\hat{\theta} = \frac{1}{1 - \dfrac{t^2 S^2 v_{22}}{x_2^2}}\left\{\left(\frac{x_1}{x_2} - \frac{t^2 S^2 v_{12}}{x_2^2}\right) \pm \sqrt{\left(\frac{x_1}{x_2} - \frac{t^2 S^2 v_{12}}{x_2^2}\right)^2 - \left(1 - \frac{t^2 S^2 v_{22}}{x_2^2}\right)\left(\frac{x_1^2}{x_2^2} - \frac{t^2 S^2 v_{11}}{x_2^2}\right)}\right\}$$

Let $\quad g = \dfrac{t^2 S^2 v_{22}}{x_2^2}, \hat{\theta} = \dfrac{x_1}{x_2}$

To get $\quad \dfrac{1}{1-g}\left\{\hat{\theta} - g\dfrac{v_{12}}{v_{22}} \pm \dfrac{tS}{x_2}\sqrt{(1-g)v_{11} - 2\hat{\theta}v_{12} + \hat{\theta}^2 v_{22} + g\dfrac{v_{12}^2}{v_{22}}}\right\}$

If x_1 and x_2 are independent, $v_{12} = 0$

It follows that $\dfrac{1}{1-g}\left\{\hat{\theta} \pm \dfrac{tS}{x_2}\sqrt{(1-g)v_{11} + \hat{\theta}^2 v_{22}}\right\}$.

Let $x_1 = \bar{y}_T - \bar{y}_R$, $x_2 = b_c$

$$v_{11} = \frac{1}{n_T} + \frac{1}{n_R}$$

$$v_{22} = \frac{1}{\sum (x_T - \bar{x}_T)^2 + \sum (x_R - \bar{x}_R)^2}$$

and $\hat{\Delta} - (\bar{x}_R - \bar{x}_T) = \dfrac{\bar{y}_T - \bar{y}_R}{b_c} = \hat{\theta}$

A 100(1-α)% confidence interval of Δ is given as

$$\left[(\bar{x}_R - \bar{x}_T) + \frac{1}{1-g} \left\{ (\hat{\Delta} - \bar{x}_R + \bar{x}_T) \pm \frac{t_{\alpha/2} S}{b_c} \sqrt{(1-g)(\frac{1}{n_R} + \frac{1}{n_T}) + \frac{(\hat{\Delta} - \bar{x}_R + \bar{x}_T)^2}{\sum (x_T - \bar{x}_T)^2 + \sum (x_R - \bar{x}_R)^2}} \right\} \right] \quad (2.8)$$

If g < 0.1, then it can be ignored in the equation (2.8), an 100(1-α)% confidence interval

of Δ which is given as

$$\Delta_U, \Delta_L = (\bar{x}_R - \bar{x}_T) + \left\{ (\hat{\Delta} - \bar{x}_R + \bar{x}_T) \pm \frac{t_{\alpha/2} S}{|b_c|} \left[\left(\frac{1}{n_R} + \frac{1}{n_T} \right) + \frac{(\hat{\Delta} - \bar{x}_R + \bar{x}_T)^2}{\sum (x_R - \bar{x}_R)^2 + \sum (x_T - \bar{x}_T)^2} \right]^{\frac{1}{2}} \right\} \quad (2,9)$$

If $\bar{x}_R = \bar{x}_T$ then equation (2.9) can be reduced to

$$\Delta_U, \Delta_L = \hat{\Delta} \pm \frac{t_{\alpha/2} S}{|b_c|} \left[\left(\frac{1}{n_R} + \frac{1}{n_T} \right) + \frac{\hat{\Delta}^2}{\sum (x_R - \bar{x}_R)^2 + \sum (x_T - \bar{x}_T)^2} \right]^{\frac{1}{2}}$$

Table 2.1: The data structure for testing linearity.

Independent variable	x_1	x_2	...	x_k
Dependent variable	y_{11}	y_{21}	...	y_{k1}
	y_{12}	y_{22}	...	y_{k2}
	.	.		.
	.	.		.
	y_{1n_1}	y_{2n_2}	...	y_{kn_k}
Sum	T_1	T_2	...	T_k
Replications	n_1	n_2	...	n_k

Table2.2: The ANOVA table of lack of fit test

Source of Variation	df	Sum of Squares	Mean Squares	F-statistics
Regression(R)	1	$SSR = S_{xy}^2 \big/ S_{xx}$	MSR	
Residual error(E)	n-2	$SSE = S_{yy} - SSR$	MSE	
Lack of fit (L)	k-2	$SSL = SSE - SSP$	MSL	$F_0 = MSL / MSP$
Pure error(P)	n-k	$SSP = \sum\sum y_{ij}^2 - \sum T_i^2 \big/ n_i$	MSP	
Total(T)	n-1	$SST = S_{yy}$		

	Period			
	I			II
Seq1	Test	W A S H O U T		Reference
Subjects → (R)				
Seq2	Reference			Test
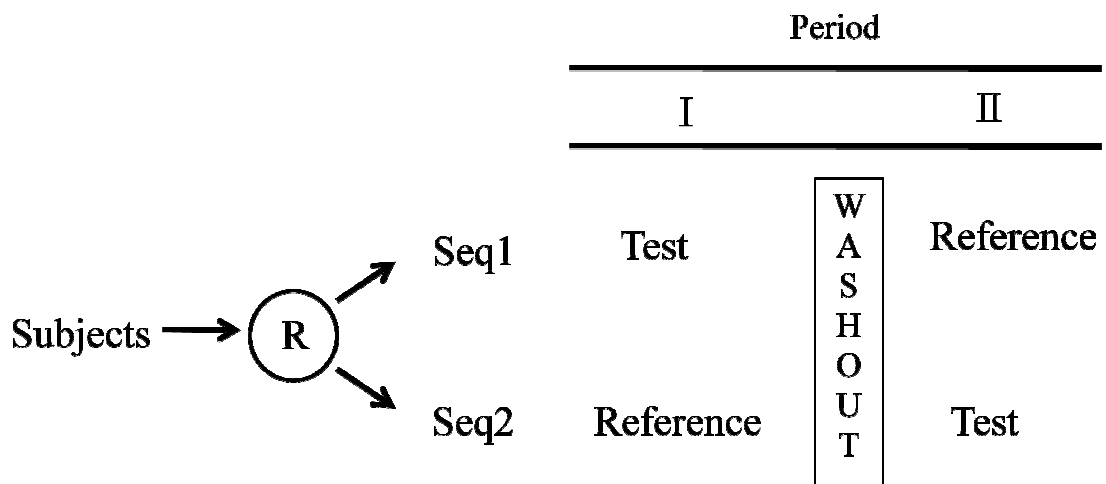
Figure 2.1: The standard two-by two crossover design

Chapter 3 Proposed Methods

In the previous chapter we reviewed the statistical methods for parallel line assay.

In this chapter we propose the applications of the parallel line assay to evaluation of

biosimilar products.

3.1 Design

Design (a) in Figure 3.1 proposed by Liu and Chow (2010) consists of two

dose-response trials, one for the biosimilar product and one for the innovator's

biological product, each with at least three dose levels with a placebo group. Eligible

patients are first randomized into biosimilar or innovator groups. Within each group,

patients are randomized again to receive one of the doses for the respective products.

Well-defined product characteristics and primary efficacy endpoints are evaluated for all

patients at their respective doses. Suppose that after a statistically significant

relationship as represented by a simple linear regression equation can be established

between the well-defined product characteristics and the primary efficacy endpoint

through dose levels for the innovator's product, maybe after a suitable transformation. If

a similar linear relationship can be also obtained for the biosimilar product and its

corresponding linear regression equation is very close to the one for the innovator's

product, then equivalence in efficacy based on the primary efficacy endpoint may be claimed. Because the innovator's product has been approved by the regulatory agencies due to its confirmed efficacy, therefore the objective of Design (a) is not to establish the efficacy of either biological products but to establish the similar patterns of the relationship between the well-defined product characteristics and primary efficacy endpoint for the two products. As a result, the sample size of Design (a) can be reduced significantly.

3.1.1. Fundamental Biosimilarity Assumption

Similar to the chemical generic drug products, approval for the biosimilar drug products can be treated as the evaluation of post-approval changes. (ICH Q5E, 2005) and this post-approval change is the change of drug manufacturers. Therefore, as indicated by Liu and Chow (2010) if the biosimilar drug products and the corresponding innovator's biological products appear highly similar, and if in addition, based on the accumulated experience relevant information and data, minute differences observed in the product characteristics are expected to have no clinically meaningful adverse effect of safety and efficacy profiles. Under this circumstance, biosimilar drug products and innovator's products can be considered similar. Therefore, except for the traditional pivotal bioequivalence study, no further data from pivotal phase III trials should be

requested. However, the above statement is based on a crucial assumption that at least one of the product characteristics is a validated and reliable predictor of the safety and efficacy profiles of the biological products. As a result, the Fundamental Biosimilarity Assumption (FBA) is

When a biosimilar product is claimed to be biosimilar to an innovator's product based on some well-defined product characteristics and is therapeutically equivalent provided that the well-defined product characteristics are validated and reliable predictors of safety and efficacy of the products.

For the chemical generic products, the well-defined product characteristics are the exposure measures for early, peak, and total portions of the concentration-time curve. The Fundamental Bioequivalence Assumption is assumed that the equivalence in the exposure measures implies therapeutically equivalent. However, due to the complexity of the biosimilar drug products, one has to verify that some validated product characteristics are indeed reliable predictors of the safety and efficacy. It follows that the design and analysis for evaluation of equivalence between the biosimilar drug product and innovator product are substantially different from those of the chemical generic products.

For design (a) Liu and Chow (2010) suggest that, the standard statistical method for analysis of parallel line assays can be used to construct the $(1-2\alpha)100\%$ confidence

interval for the relative potency in the following steps:

Step 1: Fit a linear regression equation to the primary efficacy endpoint with the group mean of the product characteristics at each dose level as the independent variable separately for the biosimilar and innovator's biological products. This may be done after a suitable transformation.

Step 2: If the estimate of the slope of any one product is not significant at the pre-defined level, then conclude that no simple relationship can be established between the product characteristic and primary endpoint and hence, a full clinical evaluation of the biosimilar product is required. Otherwise, go to Step 3.

Step 3: Test whether the two estimated simple linear regressions are parallel at the predefined significance level. If the two estimated linear regressions are not parallel, then further clinical evaluation of the biosimilar product is warranted. Otherwise, proceed to Step 4.

Step 4: Compute estimated relative potency and its corresponding $(1-2\alpha)100\%$ confidence interval. If the $(1-2\alpha)100\%$ confidence interval for the relative potency is within the predefined margins (δ_L, δ_U)**,** then equivalence in the product characteristic can be extrapolated to equivalence in the primary efficacy endpoint at the α significance level. Otherwise, further clinical investigation of the biosimilar product is needed.

The objective of application of Design (a) is not to establish the efficacy of

biosimilar product but rather to test whether the relative potency is within some pre-defined limits. We will use this design to assess the biosimilar products. Let Y be clinical responses, and X be the sample means of biomarkers or PK measures under some doses to establish the regression model.

3.2 The procedure and methods based on relative potency of product characteristics

Figure 3.2 provides a graphic depiction of the application of the parallel line assay to evaluation of the extrapolability of equivalence in product characteristic to equivalence in efficacy.

The interval hypothesis based on the relative potency is given as

$$\begin{cases} H_0 : \Delta \leq \delta_L \quad or \quad \Delta \geq \delta_U \\ v.s. \\ H_a : \delta_L \leq \Delta \leq \delta_U \end{cases} ,$$

(3.1)

where δ_L and δ_U are some pre-defined lower and upper margins.

If the $100(1-2\alpha)\%$ confidence interval for Δ given in (2.8) is within (δ_L, δ_U) then null hypothesis (3.1) is rejected and the biosimilar product is similar to the innovator product at the α significance level.

3.3 The procedure and methods based on clinical responses

The method for the relative potency mention above is based on the horizontal distance for a fixed clinical response. However, on the other hand, sometimes it is of interest to see whether the clinical responses of the biosimilar and innovator's products are similar at a fixed level of mean characteristic. This can be accomplished using the vertical distance. A graphical representation of the concept is given in Figure 3.3.

Let θ be the vertical distance between the two regression lines.

$$\theta = E(Y_R) - E(Y_T) = \alpha_R - \alpha_T \ .$$

The interval hypothesis based on θ is given as:

$$\begin{cases} H_0 : \theta \le \theta_L \quad or \quad \theta \ge \theta_U \\ v.s. \\ H_a : \theta_L \le \theta \le \theta_U \end{cases} .$$

(3.2)

An estimate of θ is given as $\hat{\theta} = \hat{Y}_R - \hat{Y}_T$.

For a fixed mean characteristic x_0 , the expected value of $\hat{\theta}$ is the difference of the intercepts of the two regression line between the innovator's and biosimilar products.

$$E(\hat{\theta} \mid x_0) = E(\hat{Y}_R \mid x_0) - E(\hat{Y}_T \mid x_0) = (\alpha_R + \beta_c x_0) - (\alpha_T + \beta_c x_0) = \alpha_R - \alpha_T \ .$$

$\hat{\theta}$ is an unbiased estimation of θ.

Since $\hat{Y}_T$ and $\hat{Y}_R$ are independent, it follows that the variance of $\hat{\theta}$ is given as

$$Var(\hat{\theta} \mid x_0) = Var(\hat{Y}_R - \hat{Y}_T \mid x_0) = Var(\hat{Y}_R \mid x_0) + Var(\hat{Y}_T \mid x_0)$$
$$= Var(a_R + b_c x_0) + Var(a_T + b_c x_0) \ .$$

On the other hand

$$Var(\hat{Y}_T \mid x_0) = Var(a_T) + 2x_0 Cov(a_T, b_c) + x_0{}^2 Var(b_c)$$

$$= \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2}\right)\sigma_T{}^2 + 2x_0 \frac{-\bar{x}\sigma_T{}^2}{\sum(x_i - \bar{x})^2} + \frac{x_0{}^2\sigma_T{}^2}{\sum(x_i - \bar{x})^2}$$

$$= \frac{\sigma_T{}^2}{n} + \frac{(x_0 - \bar{x})^2 \sigma_T{}^2}{\sum(x_i - \bar{x})^2} = \sigma_T{}^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right].$$

Similarly

$$Var(\hat{Y}_R \mid x_0) = \sigma_R{}^2\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right].$$

It follows that

$$Var(\hat{\theta} \mid x_0) = \left(\sigma_R{}^2 + \sigma_T{}^2\right)\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right].$$

Then an estimate of variance of $\hat{\theta}$ is given as

$$\hat{V}ar(\hat{\theta} \mid x_0) = \left(S_R{}^2 + S_T{}^2\right)\left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right] \text{ and}$$

$$\hat{\theta} \sim N\left((\alpha_R - \alpha_T), \left(\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum(x_i - \bar{x})^2}\right)\left(\sigma_R{}^2 + \sigma_T{}^2\right)\right).$$

As a result,

a $(1 - 2\alpha)100\%$ C.I. of $\hat{\theta}$ is given as $\left[(a_R - a_T) \pm t_{2(n-2),\alpha}\sqrt{\hat{V}ar(\hat{\theta})}\right]$. (3.3)

Let $S_R{}^2 = \dfrac{\sum(Y_i - a_R - b_c x_i)^2}{n - 2}$, and $S_T{}^2 = \dfrac{\sum(Y_i - a_T - b_c x_i)^2}{n - 2}$ be given in section 2.4.2.

If the $100(1-2\alpha)\%$ confidence interval for $\hat{\theta}$ is within (θ_L, θ_U), then the biosimilar

product is similar to the innovator product at the α significance level.

The rejection region for the proposed procedure for hypothesis (3.2) is given as

$$R\left\{(\theta,s)\,\big|\,\hat{\theta}-ts>\theta_L \text{ and } \hat{\theta}+ts<\theta_U\right\}$$

$$=R\left\{(\theta,s)\,\big|\,\theta_L+ts<\hat{\theta}<\theta_U-ts\right\}, \text{ where } s=\sqrt{\hat{Var}(\hat{\theta})}.$$

It follows that the power function is given as

$$\Phi(\theta)=P\left(\theta_L+ts<\hat{\theta}<\theta_U-ts\,|\,\theta\in\Theta\right),$$

where $\Theta=(\theta_L,\theta_U)$.

For simplicity, we only consider the situation of symmetric equivalence margin where

$$\theta_U=-\theta_L=\theta_0.$$

Next we determine the require sample size for at least 1-β power, when θ=0 and

$$n_T=n_R=n.$$

Since $\Phi(0)=P\left[-\theta_0+ts<\hat{\theta}<\theta_0-ts\,|\,\theta=0\right]\geq1-\beta$.

It follows that

$$\frac{\theta_0}{\sqrt{\left(\dfrac{1}{n}+\dfrac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)\left(S_R^{\,2}+S_T^{\,2}\right)}}-t(\alpha,2n-2)\leq t(\frac{\beta}{2},2n-2),$$

$$\Rightarrow \frac{\theta_0}{\sqrt{\left(\dfrac{1}{n}+\dfrac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right)\left(S_R^{\,2}+S_T^{\,2}\right)}}\leq t(\frac{\beta}{2},2n-2)+t(\alpha,2n-2).$$

The required sample size per drug at least 1-β power is given as

$$n=\left\{\left[\left(\frac{\theta_0}{t(\frac{\beta}{2},2n-2)+t(\alpha,2n-2)}\right)^2\frac{1}{\left(S_R^{\,2}+S_T^{\,2}\right)}\right]-\frac{(x_0-\bar{x})^2}{\sum(x_i-\bar{x})^2}\right\}^{-1}. \qquad (3.4)$$

When $\theta < \theta_0 \neq 0$ and $n_T = n_R = n$, an approximate formula for sample size

determination per drug is given as

$$n = \left\{ \left[\left(\frac{\theta_0 - \theta}{t(\frac{\beta}{2}, 2n-2) + t(\alpha, 2n-2)} \right)^2 \frac{1}{\left(S_R^2 + S_T^2 \right)} \right] - \frac{(x_0 - \bar{x})^2}{\sum (x_i - \bar{x})^2} \right\}^{-1} \quad \text{(3.5)}$$

Then the sample size per drug can be allocated to each dose. For example, with each

allocation of k doses, the required sample size per dose is n/k.
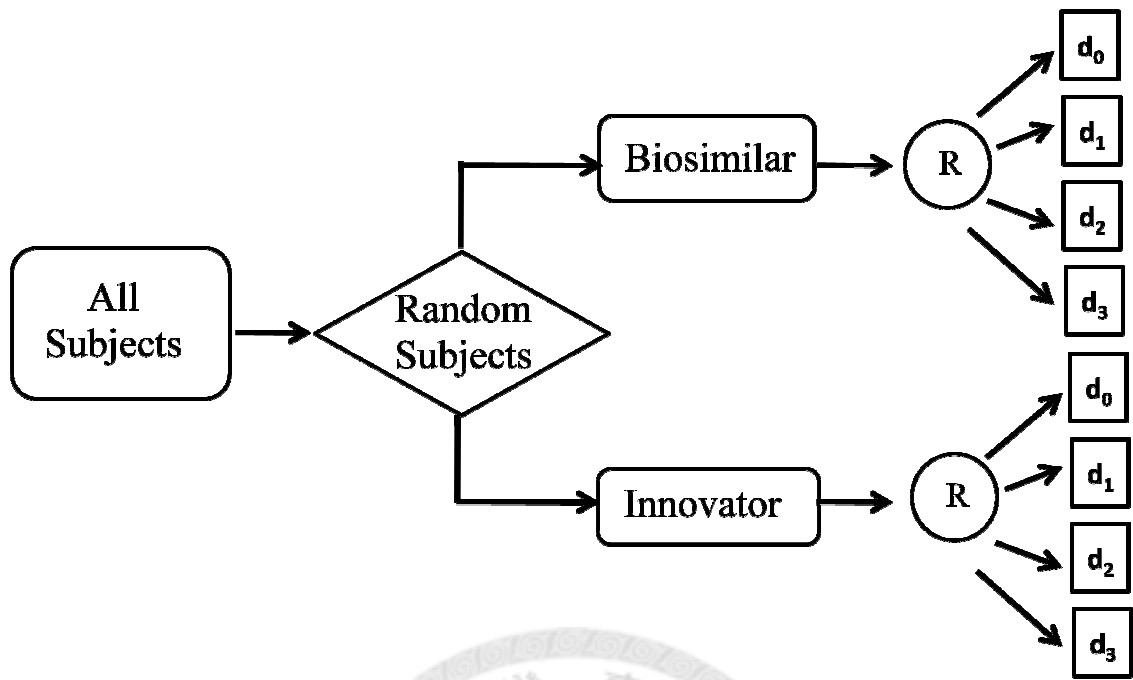
Figure 3.1: Design (a) for evaluation of extrapolation ability
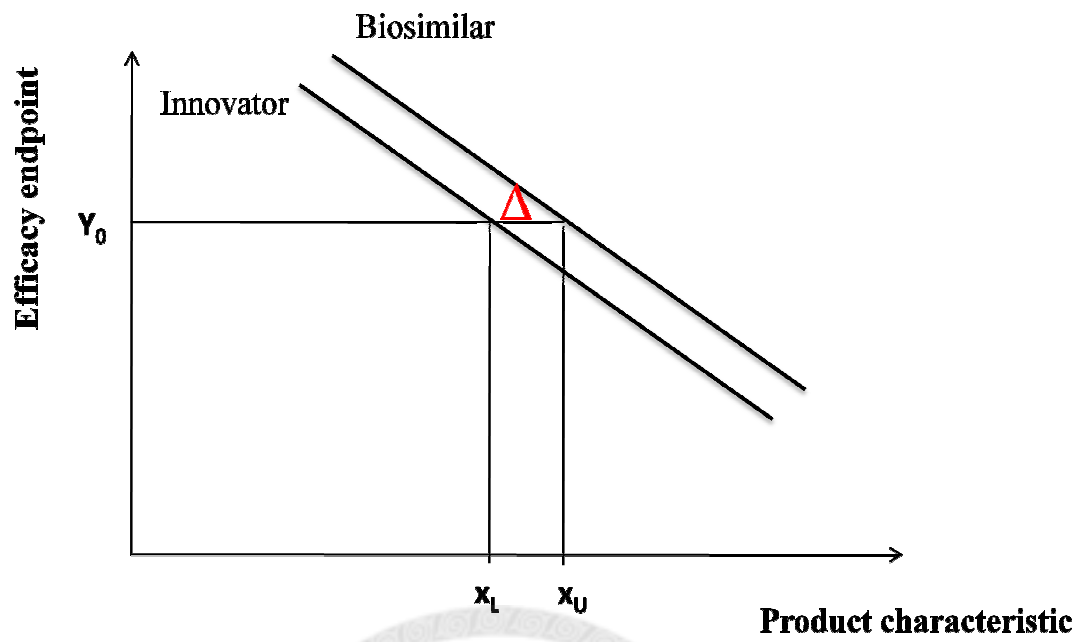
Figure 3.2: Application of the parallel line assay to evaluation of the extrapolability of based on product characteristics
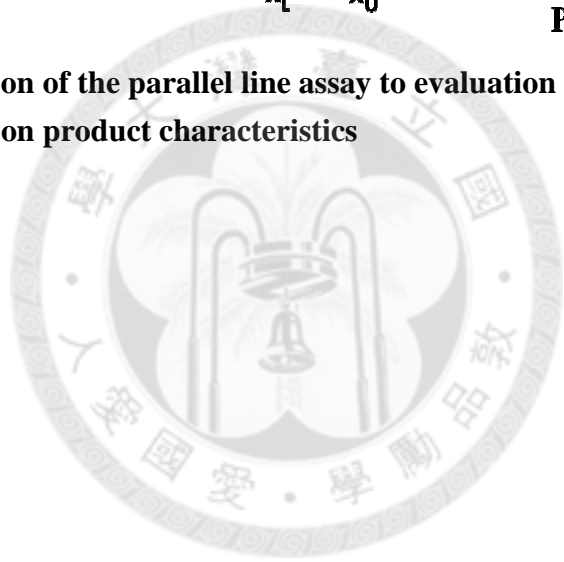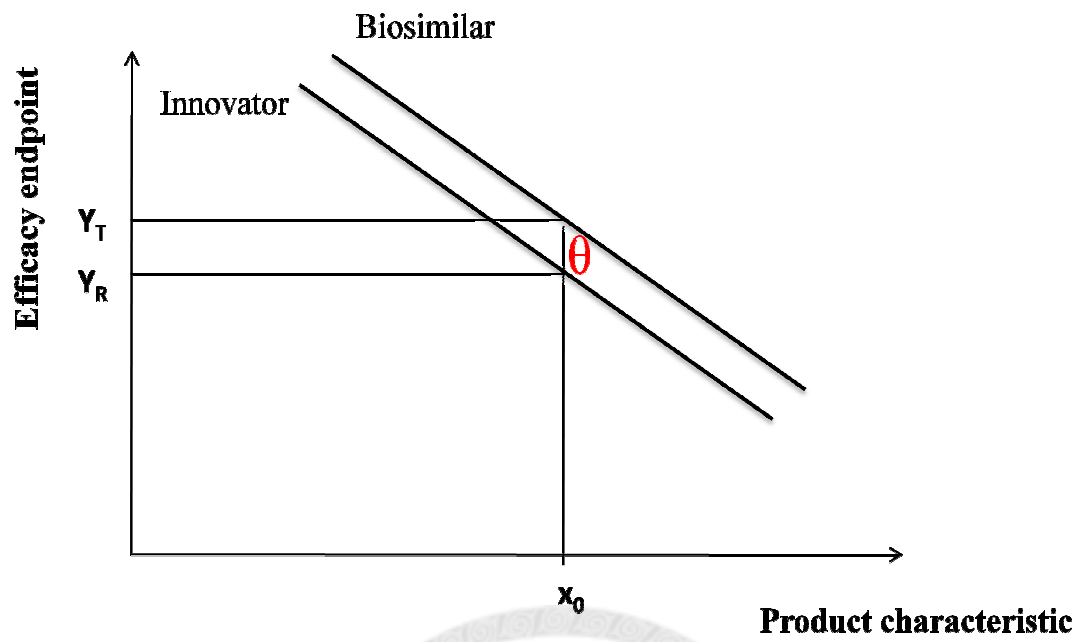
Figure 3.3: Application of the parallel line assay to evaluation of the extrapolability of based on clinical responses.
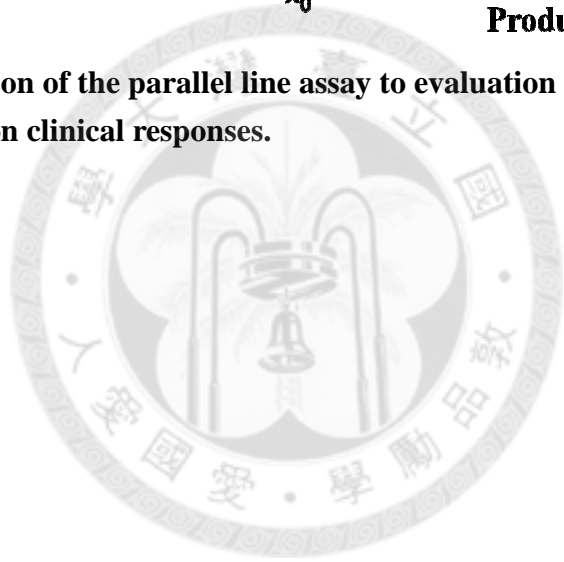
Chapter 4 Numerical Example

In this chapter we apply the proposed statistical methods derived in chapter 3 to a numerical data set. We want to test if the AUC in the innovative drug product and biosimilar are close, then clinical response values are also close.

Because we can not get real data, we use the Fortran to generate the data to evaluate the biosimilar drug. The purpose of this section is to illustrate the proposed methods with a numerical data set. The steps for generation of data are fronded below :

Step 1: Determine the relationship between dose and AUC. For each dose generate 20 AUC under normal assumption with mean=10, 20, 30, 40, 50 and variance =1.

Step 2: To calculate the mean of AUC (x_R or x_T) and use the mean to generate the data set of clinical responses(y) based on the following linear regression models.

$$Y_R = 100 - x_R + \varepsilon_R$$
$$Y_T = 100 - x_T + \varepsilon_T$$

ε_R and ε_T are generated from normal assumption with mean=0 and variance =144

Both proposed methods are illustrated with the above data set which is given in Appendix A.

The mean characteristics and mean clinical response with the correspond standard deviation for biosimilar and innovator's products are provided in Table 4.1. Table 4.2

gives the ANOVA tables for biosimilar and innovator's products.

The linear regression line for the biosimilar and innovator's products are given respectively as $\hat{y}_T = 100 - x_T$, $\hat{y}_R = 100 - x_R$. First, we should test the linear regression model. The F-values for the lack of fit test for both products, because $F_{R,LOF} = 0.174$ and $F_{T,LOF} = 0.876$ which are less than $F(0.05, 3, 95) = 2.7$, therefore we fail to reject the null hypothesis that the two regression models are linear at the 5% significance level.

Next we verify whether the two linear regression lines are parallel, the corresponding sums of square and cross products are provided in Table 4.3.

The residual mean squares for the innovator's and biosimilar products are given as

$$S_R^2 = \frac{{}_R Syy - b_R \cdot {}_R Sxy}{n_r - 2} = 118.334 \quad and$$

$$S_T^2 = \frac{{}_T Syy - b_T \cdot {}_T Sxy}{n_t - 2} = 140.067$$

$$F = 1.184 < F(0.1, 98, 98) = 2.14$$

Therefore the null hypothesis of homogenous variances is not rejected at the 5% significance level.

The pooled residual mean squares is then estimates as

$$S_p^2 = \frac{1}{2}(118.334 + 140.067) = 129.2$$

It follow the estimated variance of b_R-b_T is

$$V(b_R - b_T) = 129.2 \left[\frac{1}{19612.75} + \frac{1}{19612.75} \right] = 0.0132$$

Therefore the test statistic for hypothesis of equal slope is

$$t = \frac{\left|-0.8437 - (-1.002)\right|}{\sqrt{0.0132}} = 1.379 < t_{0.05,196} = 1.65$$

As results, null hypothesis of equal slopes is not rejected at 5% significance level

The estimated common slope is then given as

$$b_c = \frac{-16548.043 + (-19652.91)}{19612.75 + 19612.75} = -0.9229$$

It follow that the two estimated linear regression lines are given as:

$$y_R = 97.2205 - 0.9929x$$
$$y_T = 96.935 - 0.9929x$$

An estimated relative potency is given as

$$\hat{\Delta} = \frac{a_R - a_T}{b_c} = \frac{97.2205 - 96.935}{-0.9229} = -0.2875$$

Since

$$g = \frac{t^2_{\frac{0.1}{2},196} S^2}{b_c^2 \left[\sum (x_r - \overline{x}_r)^2 + \sum (x_t - \overline{x}_t)^2 \right]} = \frac{1.65^2 \cdot 129.2}{(-0.9229)^2 (19612.75 + 19612.75)} = 0.011 < 0.1$$

We can ignore g-value is computation of the 90% confidence interval for Δ. The lower

and upper limits of the 90% confidence interval for Δ.

$$(\hat{\Delta}_U, \hat{\Delta}_L) = \hat{\Delta} \pm \frac{t_{\alpha/2} S}{|b_c|} \left[\left(\frac{1}{n_R} + \frac{1}{n_T} \right) + \frac{\hat{\Delta}^2}{\sum (x_R - \overline{x}_R)^2 + \sum (x_T - \overline{x}_T)^2} \right]^{\frac{1}{2}}$$

$$= -0.2875 \pm \frac{1.65 \cdot \sqrt{129.2}}{|-0.9229|} \left[\left(\frac{1}{100} + \frac{1}{100} \right) + \frac{(-0.2875)^2}{19612.75 + 19612.75} \right]^{\frac{1}{2}}$$

$$= -0.2875 \pm 2.874$$

$$= (-3.162, \ 2.5865)$$

The 90% confidence interval for $\hat{\Delta}$ is (-3.162, 2.5865). Let $-\delta_L = \delta_U = 10$. Since the

90% confidence interval (-3.162, 2.5865) is contained within (-10, 10), then we can

conclude that the biosimilar product is similar to the innovator product at 5% significance level.

On the other hand, the estimated vertical distance of the two regression lines is

$$\hat{\theta} = a_R - a_T = 0.2855$$

Its corresponding estimated variance is given as

$$\hat{Var}(\hat{\theta}) = \left(\frac{1}{n} + \frac{(x_0 - \overline{x})^2}{\sum(x_i - \overline{x})^2}\right)\left(S_R^2 + S_T^2\right) = \left(\frac{1}{100} + \frac{(30 - 30.0615)^2}{19612.75^2}\right)(2 \times 129.2) = 2.584$$

It follows that the 90% confidence interval for θ is (-2.3712, 2.9422)

If the equivalence limits θ_0 is set at 10, since the 90% confidence interval (-2.3712, 2.9422) is completely within (-10,10) , then the null hypothesis is rejected and biosimilar product is similar to the innovator product at 5% significance level.

Suppose that $\theta = 0$, the required sample size for hypothesis (3.2) for 80% power at the 5% significance level is given as

$$n = \left\{\left[\left(\frac{\Delta}{t(\frac{\beta}{2}, 2n-2) + t(\alpha, 2n-2)}\right)^2 \frac{1}{\left(S_R^2 + S_T^2\right)}\right] - \frac{(x_0 - \overline{x})^2}{\sum(x_i - \overline{x})^2}\right\}^{-1} = 28.25$$

The sample size is 29 per group. However, there are 5 doses, therefore we round 29 up to 30 so that with an equal allocation, the sample size for each dose is 6 per products.

Table 4.1: The Descriptive Statistics of innovator's product and biosimilar product

Mean characteristic of Innovator's products	Mean	10.06	20.44	30.18	39.71	49.93
	Standard deviation	1.5	0.98	0.87	1.12	1.54
Mean characteristic of Biosimilar products	Mean	9.65	19.93	28.9	40.15	50.18
	Standard deviation	0.59	1	0.84	0.93	0.86
Innovator's products	Mean	85.69	78.74	68.61	62.11	52.23
	Standard deviation	9.2	12.03	10.99	11.55	11.12
Biosimilar products	Mean	87.46	80.42	68.9	62.23	46.95
	Standard deviation	11.79	11.06	11.46	10.85	13.87

Table 4.2: The ANOVA table of lack of fit test

Source of Variation	df	Innovator's product			Biosimialr product		
		Sum of Squares	Mean Squares	F-statistics	Sum of Squares	Mean Squares	F-statistics
Regression(R)	1	13962.225	13962.225		19693.152	19693.152	
Residual error(E)	98	11596.045	118.327		13725.538	140.057	
Lack of fit (L)	3	63.448	21.149	0.174	369.398	123.133	0.876
Pure error(P)	95	11532.600	121.396		13356.140	140.591	
Total(T)	99	25558.270			33418.690		

Table 4.3: The Descriptive Statistics of innovator's product and biosimilar product

group	n	Sxx	Sxy	Syy	xbar	ybar	a	b
Innovator's	100	19612.75	-16548.04	25558.27	30.0615	69.4769	94.841	-0.8437
Biosimilar	100	19612.75	-19652.91	33418.69	30.0615	69.1914	99.315	-1.002

Chapter 5 Simulation Study

In this chapter, Compaq Visual Fortran 6.5 and IMSL's STAT/LIBRARY FORTRAN subroutines RNNOR were used in simulation study to generate the data to do simulation to compare the coverage probability and power with different sample size, variance and dose. A flow chart of the simulation study is given in Figure 5.1.

5.1 Simulation Process

First we generate the data of AUC, for each dose generates 20 AUC under normal assumption with specified means and variances. The number of doses is 3, 5, 7. In addition, the RNNOR subroutine of Compaq Visual Fortran 6.5/IMSL was employed to generate normal deviates with variance 1 and dose 10 to k10 by 10, where k is the number of the dose. The responses were generated using following regression equation, and where x_R (x_T) are the mean AUC at each dose.

$$Y_R = 100 - x_R + \varepsilon_R$$
$$Y_T = 100 - x_T + \varepsilon_T \ .$$

ε are i.i.d. with mean 0 and standard deviation from 12 to 24 by 6.

Then we use the statistical method mentioned in Chapter 3 to determine the 90% confidence interval. Then we repeated for 10000 times to determine the empirical power, empirical size and coverage probability. The equivalence limits are set to be ±10 for

both methods based on product characteristic and clinical responses. The empirical size

is estimated the proportion of 10000 random samples for which the confidence intervals

are totally contained within (-10, 10) when $\mu_T - \mu_R = \pm 10$.

The empirical power is calculated as the proportion of 10000 random samples in

which the confidence interval falls entirely within (-10, 10), where

$\mu_T - \mu_R = -5, 0$ and 5.

The coverage probability is computed as the proportion of 10000 confidence intervals

containing the specified $\mu_T - \mu_R = -10$ to 10 by 5

At the 5% nominal significance level, a simulation study with 10000 samples

that 95% of the empirical sizes will be between 0.0457 and 0.0542 if the proceed

procedure can adequately control the type one error rate at its nominal level. Similarly,

90% coverage probability will be between 0.8941 and 0.9059 if the proceed procedure

can provide adequate coverage probability at the 90% level.

5.2 Simulation Result

The simulation results based on relative potency are given from Table 5.1 to Table

5.3. In general, the empirical size decreases as the variability increases or the number of

the doses increase. 36.67% (33/90) of empirical sizes are within (0.0457, 0.0542). This

indicates that the method based on relative potency can not adequately control the size.

When σ = 12 and the number of the doses is fewer than 5, the method based on relative

potency yields the empirical sizes greater than 0.0542. Therefore, the method based on

relative potency is liberal when the variability is small and the number of doses is fewer.

However when σ = 24 and the number of doses is 5 or above, the empirical sizes are

below 0.0457. This reveals that the method based on relative potency becomes much

more conservative when the variability is higher and the number of doses is larger.

The simulation results based on clinical responses are provided from Table 5.4 to

Table 5.6. When σ = 12, all 30 empirical sizes are within (0.0457, 0.0542). However,

when σ > 12, all empirical size are below 0.0457. The empirical sizes based on clinical

responses are smaller than 0.01 when σ = 24. Therefore, the method based on clinical

responses can adequately control the size when variability is small. It becomes

extremely conservative when variability is large.

The sample size for the simulation study is to provide a 80% power when

$\mu_{\mathrm{T}} - \mu_{\mathrm{R}} = 0$ and σ = 12 using the formula based on the clinical responses. This is

confirmed by the fact that empirical power is at least 80% when $\mu_{\mathrm{T}} - \mu_{\mathrm{R}} = 0$ and σ = 12.

On the other hand, when $\mu_{\mathrm{T}} - \mu_{\mathrm{R}} = 0$ and σ=12, the empirical power based on the

relative potency are around 80% though they are smaller than those based on clinical

response.

In general, the empirical power for both methods decreases dramatically as the

variability increases for the sample size investigated in the simulation study. On the other hand, the empirical power increases slightly as the number of doses increases or the sample size increases. None of the coverage probabilities given from Table 5.1 to Table 5.6 are smaller than 0.8941. This suggested that the confidence intervals constructed by both methods can provide sufficient probability.

Table 5.1: Simulation result of power and coverage probability based on relative potency under σ=12

dose	n		-10	-5	0	5	10
3	26	power	0.0648	0.4130	0.7303	0.4198	0.0703
		coverage probability	0.9231	0.9200	0.9220	0.9214	0.9182
3	27	power	0.0706	0.4250	0.7394	0.4223	0.0706
		coverage probability	0.9186	0.9174	0.9157	0.9170	0.9161
3	28	power	0.0686	0.4373	0.7684	0.4387	0.0726
		coverage probability	0.9188	0.9174	0.9183	0.9162	0.9151
3	29	power	0.0672	0.4475	0.7831	0.4451	0.0672
		coverage probability	0.9186	0.9177	0.9198	0.9173	0.9182
3	30	power	0.0665	0.4574	0.7961	0.4622	0.0675
		coverage probability	0.9204	0.9206	0.9185	0.9203	0.9193
5	26	power	0.0585	0.4196	0.7776	0.4330	0.0560
		coverage probability	0.9078	0.9103	0.9081	0.9070	0.9100
5	27	power	0.0588	0.4346	0.7915	0.4396	0.0590
		coverage probability	0.9034	0.8991	0.9021	0.9014	0.9017
5	28	power	0.0593	0.4457	0.8153	0.4475	0.0601
		coverage probability	0.9038	0.9029	0.9051	0.9041	0.9032
5	29	power	0.0546	0.4615	0.8359	0.4606	0.0570
		coverage probability	0.9071	0.4615	0.8359	0.9608	0.9043
5	30	power	0.0545	0.4692	0.8441	0.4794	0.0560
		coverage probability	0.9090	0.9069	0.9061	0.9057	0.9075
7	26	power	0.0535	0.4195	0.7882	0.4347	0.0537
		coverage probability	0.9059	0.9070	0.9048	0.9042	0.9058
7	27	power	0.0567	0.4380	0.8038	0.4410	0.0558
		coverage probability	0.8977	0.8957	0.8986	0.8986	0.8977
7	28	power	0.0550	0.4472	0.8315	0.4499	0.0553
		coverage probability	0.9015	0.8999	0.9005	0.9018	0.9005
7	29	power	0.0513	0.4626	0.8510	0.4613	0.0532
		coverage probability	0.9054	0.9068	0.9032	0.9025	0.9021
7	30	power	0.0504	0.4744	0.8624	0.4816	0.0536
		coverage probability	0.9042	0.9028	0.9032	0.9022	0.9028

Table 5.2: Simulation result of power and coverage probability based on relative potency under σ=18

dose	n		-10	-5	0	5	10
3	26	power	0.0617	0.1979	0.2967	0.2041	0.0676
		coverage probability	0.9282	0.9392	0.9482	0.9406	0.9235
3	27	power	0.0669	0.2065	0.3099	0.2142	0.0677
		coverage probability	0.9241	0.9376	0.9437	0.9358	0.9217
3	28	power	0.0672	0.2229	0.3381	0.2257	0.0685
		coverage probability	0.9231	0.9345	0.9420	0.9349	0.9214
3	29	power	0.0667	0.2325	0.3475	0.2266	0.0657
		coverage probability	0.9251	0.9366	0.9448	0.9355	0.9260
3	30	power	0.0662	0.2374	0.3700	0.2427	0.0684
		coverage probability	0.9268	0.9393	0.9424	0.9367	0.9242
5	26	power	0.0500	0.1745	0.2725	0.1802	0.0502
		coverage probability	0.9166	0.9202	0.9188	0.9178	0.9181
5	27	power	0.0505	0.1844	0.2874	0.1858	0.0507
		coverage probability	0.9095	0.9079	0.9089	0.9103	0.9102
5	28	power	0.0526	0.2033	0.3204	0.2026	0.0524
		coverage probability	0.9132	0.9125	0.9146	0.9125	0.9141
5	29	power	0.0495	0.2195	0.3364	0.208	0.0525
		coverage probability	0.9164	0.9150	0.9151	0.9140	0.9125
5	30	power	0.0513	0.2211	0.3645	0.2189	0.0526
		coverage probability	0.9156	0.9156	0.9143	0.9124	0.9138
7	26	power	0.0439	0.1662	0.2664	0.1685	0.0454
		coverage probability	0.9115	0.9111	0.9094	0.9095	0.9100
7	27	power	0.0477	0.1783	0.2829	0.1791	0.0473
		coverage probability	0.9013	0.9000	0.9029	0.9026	0.9032
7	28	power	0.0476	0.1978	0.3125	0.1922	0.0461
		coverage probability	0.9040	0.9037	0.9049	0.9054	0.9054
7	29	power	0.0457	0.2110	0.3333	0.2075	0.0486
		coverage probability	0.9097	0.9107	0.9072	0.9071	0.9060
7	30	power	0.0465	0.2151	0.3656	0.2214	0.0498
		coverage probability	0.9075	0.9068	0.9064	0.9056	0.9062

Table 5.3: Simulation result of power and coverage probability based on relative potency under σ=24

dose	n		-10	-5	0	5	10
3	26	power	0.0420	0.0907	0.1207	0.0967	0.0489
		coverage probability	0.9303	0.9516	0.9667	0.9520	0.9253
3	27	power	0.0479	0.0992	0.1245	0.1054	0.0505
		coverage probability	0.9245	0.9519	0.9662	0.9550	0.9230
3	28	power	0.0506	0.1060	0.1428	0.1148	0.0526
		coverage probability	0.9235	0.9485	0.9663	0.9494	0.9218
3	29	power	0.0521	0.1208	0.1471	0.1087	0.0481
		coverage probability	0.9266	0.9515	0.9625	0.9477	0.9273
3	30	power	0.0521	0.1184	0.1576	0.1194	0.0541
		coverage probability	0.9267	0.9516	0.9633	0.9513	0.9250
5	26	power	0.0237	0.0504	0.0688	0.0507	0.0239
		coverage probability	0.9271	0.9311	0.9327	0.9307	0.9292
5	27	power	0.0236	0.0510	0.0675	0.0545	0.0248
		coverage probability	0.9223	0.9212	0.9229	0.9221	0.9208
5	28	power	0.0256	0.0613	0.0842	0.0650	0.0275
		coverage probability	0.9235	0.9253	0.9268	0.9262	0.9236
5	29	power	0.0274	0.0701	0.0915	0.0638	0.0275
		coverage probability	0.9256	0.9259	0.9254	0.9254	0.9219
5	30	power	0.0292	0.0700	0.1026	0.0734	0.0295
		coverage probability	0.9252	0.9259	0.9266	0.9238	0.9245
7	26	power	0.0148	0.0349	0.0441	0.0336	0.0152
		coverage probability	0.9185	0.9188	0.9173	0.9159	0.9166
7	27	power	0.0162	0.0330	0.0452	0.0374	0.0157
		coverage probability	0.9088	0.9065	0.9073	0.9090	0.9088
7	28	power	0.0173	0.0413	0.0541	0.0451	0.0169
		coverage probability	0.9101	0.9091	0.9112	0.9108	0.9108
7	29	power	0.0187	0.0543	0.0671	0.0458	0.0207
		coverage probability	0.9150	0.9160	0.9146	0.9129	0.9125
7	30	power	0.0184	0.0510	0.0743	0.0536	0.0207
		coverage probability	0.9134	0.9124	0.9111	0.9101	0.9105

Table 5.4: Simulation result of power and coverage probability based on clinical response under σ=12

dose	n		-10	-5	0	5	10
3	26	power	0.0478	0.4161	0.8013	0.4279	0.0455
		coverage probability	0.9067	0.9067	0.9067	0.9067	0.9067
3	27	power	0.0498	0.4355	0.8130	0.4337	0.0502
		coverage probability	0.9000	0.9000	0.9000	0.9000	0.9000
3	28	power	0.0488	0.4472	0.8396	0.4435	0.0495
		coverage probability	0.9017	0.9017	0.9017	0.9017	0.9017
3	29	power	0.0478	0.4636	0.8592	0.4536	0.0485
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
3	30	power	0.0474	0.4707	0.8703	0.4771	0.0489
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
5	26	power	0.0475	0.4149	0.8020	0.4279	0.0459
		coverage probability	0.9066	0.9066	0.9066	0.9066	0.9066
5	27	power	0.0496	0.4363	0.8125	0.4376	0.0502
		coverage probability	0.9002	0.9002	0.9002	0.9002	0.9002
5	28	power	0.0489	0.4464	0.8405	0.4437	0.0496
		coverage probability	0.9015	0.9015	0.9015	0.9015	0.9015
5	29	power	0.0479	0.4639	0.8602	0.4541	0.0484
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
5	30	power	0.0474	0.4706	0.8707	0.4767	0.0488
		coverage probability	0.9038	0.9038	0.9038	0.9038	0.9038
7	26	power	0.0480	0.4153	0.8022	0.4291	0.0461
		coverage probability	0.9059	0.9059	0.9059	0.9059	0.9059
7	27	power	0.0498	0.4361	0.8128	0.4370	0.0502
		coverage probability	0.9000	0.9000	0.9000	0.9000	0.9000
7	28	power	0.0489	0.4473	0.8399	0.4445	0.0494
		coverage probability	0.9017	0.9017	0.9017	0.9017	0.9017
7	29	power	0.0475	0.4634	0.8595	0.4546	0.0486
		coverage probability	0.9039	0.9039	0.9039	0.9039	0.9039
7	30	power	0.0477	0.4714	0.8701	0.4768	0.0483
		coverage probability	0.9040	0.9040	0.9040	0.9040	0.9040

Table 5.5: Simulation result of power and coverage probability based on clinical response under σ=18

dose	n		-10	-5	0	5	10
3	26	power	0.0369	0.1464	0.2477	0.1527	0.0354
		coverage probability	0.9067	0.9067	0.9067	0.9067	0.9067
3	27	power	0.0395	0.1646	0.2732	0.1667	0.0402
		coverage probability	0.9000	0.9000	0.9000	0.9000	0.9000
3	28	power	0.0407	0.1839	0.3006	0.1752	0.0396
		coverage probability	0.9017	0.9017	0.9017	0.9017	0.9017
3	29	power	0.0414	0.1972	0.3236	0.1696	0.0427
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
3	30	power	0.0423	0.2046	0.3595	0.2140	0.0449
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
5	26	power	0.0367	0.1471	0.2484	0.1532	0.0358
		coverage probability	0.9066	0.9066	0.9066	0.9066	0.9066
5	27	power	0.0395	0.1650	0.2738	0.1661	0.0402
		coverage probability	0.9002	0.9002	0.9002	0.9002	0.9002
5	28	power	0.0408	0.1853	0.3010	0.1761	0.0396
		coverage probability	0.9015	0.9015	0.9015	0.9015	0.9015
5	29	power	0.0415	0.1977	0.3244	0.1961	0.0426
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
5	30	power	0.0423	0.2045	0.3604	0.2144	0.0447
		coverage probability	0.9038	0.9038	0.9038	0.9038	0.9038
7	26	power	0.0371	0.1474	0.2477	0.1539	0.0362
		coverage probability	0.9059	0.9059	0.9059	0.9059	0.9059
7	27	power	0.0398	0.1643	0.2735	0.1661	0.4010
		coverage probability	0.9000	0.9000	0.9000	0.9000	0.9000
7	28	power	0.0408	0.1860	0.3014	0.1747	0.0392
		coverage probability	0.9017	0.9017	0.9017	0.9017	0.9017
7	29	power	0.0409	0.1980	0.3243	0.1962	0.0429
		coverage probability	0.9039	0.9039	0.9039	0.9039	0.9039
7	30	power	0.0427	0.2051	0.3603	0.2144	0.0442
		coverage probability	0.9040	0.9040	0.9040	0.9040	0.9040

Table 5.6: Simulation result of power and coverage probability based on clinical response under σ=24

dose	n		-10	-5	0	5	10
3	26	power	0.0026	0.0087	0.0087	0.0065	0.0031
		coverage probability	0.9067	0.9067	0.9067	0.9067	0.9067
3	27	power	0.0045	0.0077	0.0117	0.0087	0.0049
		coverage probability	0.9000	0.9000	0.9000	0.9000	0.9000
3	28	power	0.0042	0.0108	0.0173	0.0126	0.0040
		coverage probability	0.9017	0.9017	0.9017	0.9017	0.9017
3	29	power	0.0052	0.0146	0.0220	0.0136	0.0059
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
3	30	power	0.0071	0.0215	0.0287	0.0190	0.0087
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
5	26	power	0.0026	0.0086	0.0085	0.0063	0.0033
		coverage probability	0.9066	0.9066	0.9066	0.9066	0.9066
5	27	power	0.0043	0.0078	0.0123	0.0091	0.0046
		coverage probability	0.9002	0.9002	0.9002	0.9002	0.9002
5	28	power	0.0046	0.0111	0.0170	0.0126	0.0037
		coverage probability	0.9015	0.9015	0.9015	0.9015	0.9015
5	29	power	0.0052	0.0149	0.0219	0.0137	0.0059
		coverage probability	0.9037	0.9037	0.9037	0.9037	0.9037
5	30	power	0.0072	0.0210	0.0283	0.0190	0.0088
		coverage probability	0.9038	0.9038	0.9038	0.9038	0.9038
7	26	power	0.0027	0.0085	0.0088	0.0058	0.0037
		coverage probability	0.9059	0.9059	0.9059	0.9059	0.9059
7	27	power	0.0044	0.0079	0.0124	0.0092	0.0046
		coverage probability	0.9000	0.9000	0.9000	0.9000	0.9000
7	28	power	0.0043	0.0114	0.0175	0.0122	0.0039
		coverage probability	0.9017	0.9017	0.9017	0.9017	0.9017
7	29	power	0.0052	0.0148	0.0220	0.0138	0.0061
		coverage probability	0.9039	0.9039	0.9039	0.9039	0.9039
7	30	power	0.0072	0.0213	0.0289	0.0190	0.0084
		coverage probability	0.9040	0.9040	0.9040	0.9040	0.9040

The flow chart contains the following boxes and text:

Generate the AUC under Normal assumption with mean and variance

Three branches below:

Dose=3
Mean =10, 20. 30
Variance=1

Dose=5
Mean =10, 20. 30, 40, 50
Variance=1

Dose=7
Mean =10, 20. 30, 40, 50, 60, 70
Variance=1

Calculate the mean of AUC and use the mean to generate the clinical response

$$y_r = 1\,0\,0 - x_r + \varepsilon_r$$
$$y_t = 1\,0\,0 - x_t + \varepsilon_t$$

Obtain the confidence interval of the methods based on relative potency

Obtain the confidence interval of the methods based on clinical response

Coverage probability

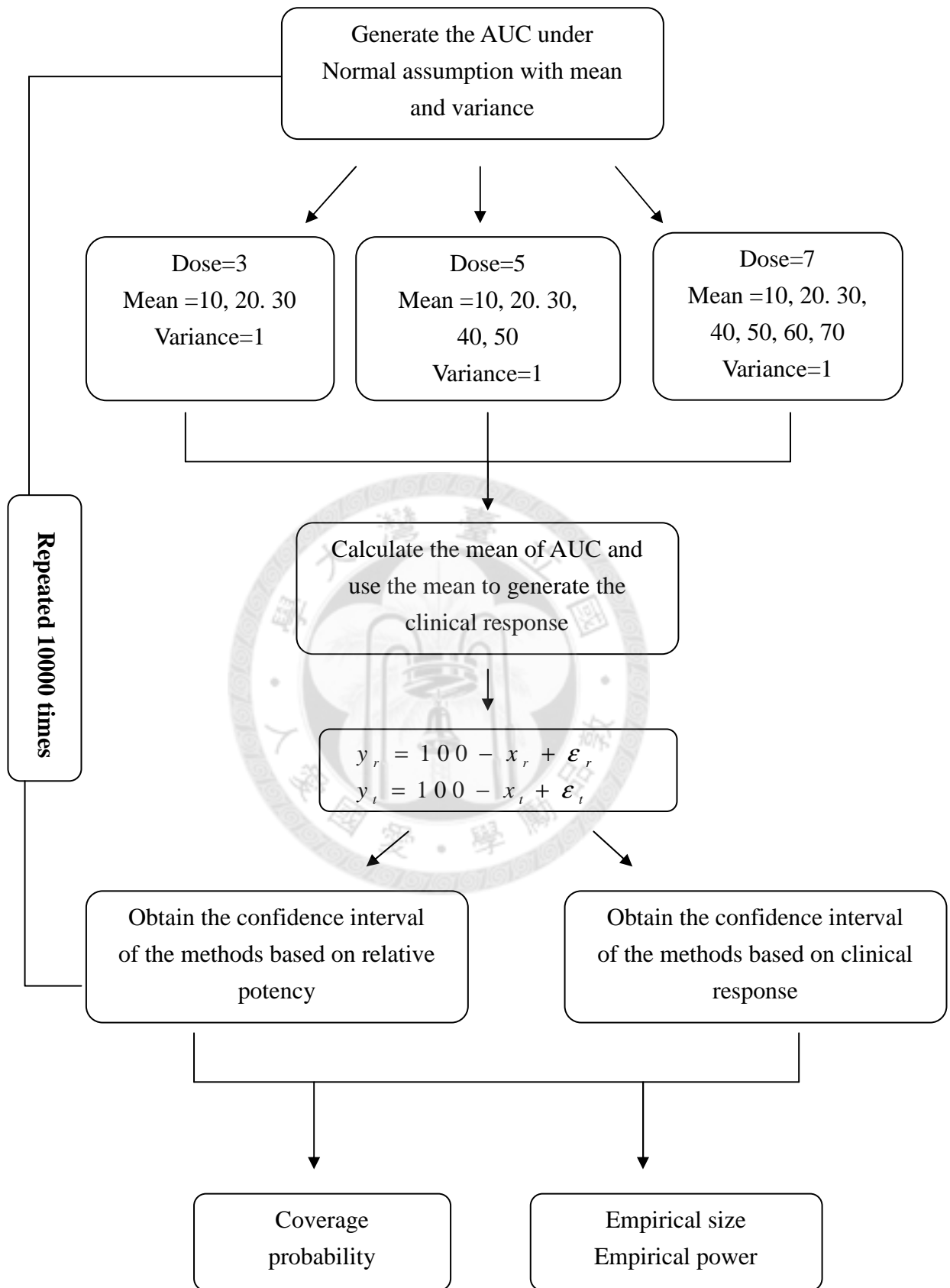**Empirical size Empirical power**

Repeated 10000 times

Figure 5.1: Flow chart of simulation process

Chapter 6 Discussion and Conclusion

Following the results of simulation, when the variance increases, the error of size decrease, and the empirical power also decreases. And the number of doses also affects the empirical power. Under the same sample size, when the number of doses increases, the empirical power also increases. In summary, when sample size is large enough with an reasonable variance, we can choose the dose appropriately.

From the

$$\left[(\bar{x}_R - \bar{x}_T) + \frac{1}{1-g} \left\{ (\Delta - \bar{x}_R + \bar{x}_T) \pm \frac{t_{\alpha/2} S}{|b_c|} \sqrt{(1-g)(\frac{1}{n_R} + \frac{1}{n_T}) + \frac{(\Delta - \bar{x}_R + \bar{x}_T)^2}{\sum (x_R - \bar{x}_R)^2 + \sum (x_T - \bar{x}_T)^2}} \right\} \right],$$

we will therefore wish to choose the design so that

  i.   t-value is small,

  ii.   the common slope b_c is large,

  iii. $(\frac{1}{n_R} + \frac{1}{n_T})$ is small,

  iv. $(\Delta - \bar{x}_R + \bar{x}_T)$ is small, and

  v. $\sum S_{xx}$ is large.

Then the relative potency will be small (Finney, 1979).

When the variance increases, the empirical power decreases. So we could increase the sample size to increase empirical power. But when we increase the sample size the cost of the experiment also increases.

In the Chapter 2, we want to know whether two linear regression lines are parallel. In other words we need to check whether the slopes are equal between the two regression lines. So we let the null hypothesis be $\beta_R = \beta_T$ and the alternative hypothesis be $\beta_R \neq \beta_T$. When the Eq. (2.7) less than $\mathrm{t}\left(\alpha/2,\ \mathrm{n}_R + \mathrm{n}_T - 4\right)$, we fail to reject the null hypothesis. But it dose not imply that the slopes are equal between the two regression lines. Therefore, we can apply the concept of equivalence to formulate the null and alternative hypothesis. As a result, when we reject the null hypothesis, it ensures that the slopes are equivalent between the two regression lines.

In the Chapter 3 we proposed the sample size based on clinical endpoints. However, the sample size based on relative potency is not provided. This is the future goal of our research. In the Chapter 5 we generate the data of AUC under normal assumption with mean and standard deviation 1 when we increase the standard deviation to 4 and 8. The additional simulation results on the coverage probability, empirical power and empirical size are quite consistent.

In this thesis we assume the clinical response is a continuous variable. But many clinical responses are categorical data or censored data. Application of the parallel line assay to assess biosimilar products may be adapted using the logistic regression for binary data and the proportional hazard model for censored data. However further research is required for these applications.

References

沈明來(2000), 生物檢定統計法 第一版, 九州圖書文物有限公司

Chow, S. C. and Liu, J. P. (2008). Design and Analysis of Bioavailability and Bioequivalence Studies. 3rd ed. Marcel Dekker, Inc. New York.

Chow, S. C. and Liu, J. P. (2010). Statistical Assessment of Biosimilar Products. Journal of Biopharmaceutical Statistics 20:10-30

Finney, D.J. (1979). Statistical Method in Biological Assay. 3rd Ed. Oxford University Press, New York, New York.

Generic Pharmaceutical Association and Company Reports(2005)

ICH (2005). Q5E Guideline on Comparability of Biotechnological/Biological Products Subject to Changes in Their Manufacturing Process. Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, the US Food and Drug Administration, Rockville, Maryland, USA

Keith O. Webber (2007). Biosimilars: Are we there yet? Presented at Biosimilars 2007, George Washington University, Washington DC, USA.

Schellekens, H. (2004). How similar do 'biosimilar' need to be? Nat Biotechnol, 22, 1357-1359.

http://www.sandoz.com/site/en/product_range/more_about_biosimilars/index.shtml

http://en.wikipedia.org/wiki/Biosimilar (2010.04)

Appendix A

The data of numerical example

Table A.1: The data of AUC

Innovator's products					Biosimilar products				
9.86	9.40	9.17	11.38	10.66	9.96	8.76	8.49	9.50	9.62
12.68	10.41	11.76	10.03	10.97	9.83	9.98	8.91	8.22	9.50
9.28	10.35	9.84	8.76	10.16	9.52	9.93	10.70	10.32	8.41
8.01	10.26	11.33	7.81	9.01	10.18	9.53	10.82	10.04	10.70
21.34	20.53	20.22	19.42	21.55	19.36	21.95	19.81	19.55	19.72
20.57	20.92	20.38	19.95	20.46	19.88	19.22	19.93	18.81	20.13
21.09	19.97	21.07	19.93	20.54	20.16	19.53	19.00	18.74	20.57
21.28	16.97	21.24	21.00	20.30	19.23	20.03	22.00	21.97	19.02
30.30	28.68	31.17	29.57	30.88	30.75	29.28	28.67	28.81	29.00
31.09	30.33	29.83	30.47	29.93	30.33	29.43	31.16	29.10	29.80
31.75	29.77	29.18	29.84	28.20	29.28	30.66	28.77	30.75	30.23
30.38	29.59	30.88	29.96	31.80	31.77	29.32	30.76	30.74	29.36
39.14	40.50	39.70	38.79	39.98	38.00	39.91	40.09	39.87	38.73
40.65	39.63	39.69	39.29	41.00	40.89	39.08	40.13	42.43	40.73
39.39	42.07	37.82	39.21	38.24	40.73	39.34	40.86	39.35	41.12
38.33	39.76	39.37	40.16	41.40	40.28	40.28	40.03	40.78	40.42
50.45	50.59	48.94	52.19	50.73	50.79	51.51	51.61	48.87	50.73
50.06	48.31	49.34	51.14	49.50	50.07	50.19	50.47	49.26	50.37
49.37	49.34	50.03	47.96	49.58	49.62	49.59	48.59	50.56	48.95
52.67	49.02	48.31	51.21	49.86	50.27	51.05	49.72	49.49	51.87

Table A.2: The data of clinical response

Innovator's products					Biosimilar products				
89.47	75.04	71.79	83.91	85.42	111.51	93.15	71.99	70.92	98.87
87.86	89.66	76.82	68.58	83.89	88.17	95.34	88.47	97.56	81.15
84.22	89.12	98.40	93.83	70.86	100.52	89.91	87.78	63.14	80.89
92.15	84.30	99.79	90.38	98.38	90.40	70.60	84.44	95.64	88.69
71.87	102.98	77.28	74.22	76.19	101.77	98.56	80.20	78.89	59.28
78.13	70.26	78.74	65.34	81.08	82.88	80.33	86.99	88.19	78.83
81.50	73.90	67.57	64.45	86.42	67.17	89.04	72.38	69.72	79.79
70.35	79.94	103.59	103.25	67.78	83.11	75.70	60.67	91.30	83.61
78.88	61.20	53.86	55.52	57.78	54.74	82.10	68.82	76.17	63.48
73.74	62.94	83.74	59.06	67.46	84.83	67.55	78.84	61.37	68.29
61.19	77.69	55.05	78.82	72.62	50.09	57.12	79.71	64.08	78.60
91.10	61.63	78.96	78.75	62.19	61.25	93.50	54.45	69.57	63.50
36.26	59.18	61.37	58.71	45.04	67.11	59.46	59.67	38.09	61.58
70.99	49.27	61.82	89.42	69.01	67.19	45.64	67.16	55.15	51.97
69.08	52.42	70.58	52.44	73.72	70.10	67.08	55.87	55.44	75.09
63.68	63.60	60.71	69.65	65.31	60.81	64.71	67.42	89.67	65.37
59.52	68.23	69.36	36.53	58.88	65.73	50.96	34.17	24.38	39.24
50.90	52.33	55.74	41.17	54.52	66.40	69.73	58.73	41.97	44.85
45.55	45.20	33.16	56.82	37.53	40.27	55.25	43.57	31.50	41.50
53.32	62.62	46.71	43.96	72.51	33.48	25.41	55.35	50.24	66.22

Appendix B

Fortran code for simulation program

```
!sstats.lib smaths.lib sstatd.lib smathd.lib
PROGRAM MAIN
!use IMSL
INTEGER ISEED,ISEED1, NOUT, NR,i,j,k,q
REAL XM,XSTD,t(5),x_bar,yt_bar,yr_bar,Syyr,Syyt,
Sxx,Sxy,nm,m,g,tsc,mu,ml,DFD,DFN,F,FIN,P,DF,TV,TIN
REAL R1(6),R2(30),R3(30),yr(30),yt(30),x(30)
EXTERNAL RNNOA, RNSET, UMACH, SSCAL, SADD, FIN , TIN
CALL UMACH (2, NOUT)
OPEN(8, FILE='M confidence interval.txt')
print*,"q="
read(*,*) q
!====產生 100 筆亂數====================
ISEED = 25557
!ISEED1 = 234
CALL RNSET (ISEED)
!CALL RNSET (ISEED1)
do k=1,10000
!=====================t,F value=========
NR2=30

P=0.9
DFN=NR2-2
DFD=NR2-2
F=FIN(P,DFN,DFD)
!print*, F

P=0.95
DF=2*NR2-4
TV=TIN(P,DF)
!print*, TV
```

```fortran
!========= ===================
NR1 = 6
XSTD = 1
do I=1,5 !產生 mean=10,20,30,40,50
XM = I*10
!CALL RNSET (ISEED)
CALL RNNOA (NR1, R1)
CALL SSCAL (NR1,XSTD,R1,1)
CALL SADD (NR1,XM,R1,1)
!WRITE (NOUT,99999) R1
99999 FORMAT ('normal random deviates: ', 20F8.4)
t(I)=SUM(R1)/NR1
end do
!===============將 5 個 array 轉成 100array=============
NR2 = 30
average_t=sum(t)/5
!WRITE(8,*)'Mean of X is', X
do i = 1, 5
do j = 1, NR2
if (((i-1)*6 .lt. j) .and. (j .le. i*6) ) then
x(j)=t(i)
end if
end do
end do
!print*,'x array',x
x_bar=sum(x)/NR2
!print*,'x_bar',x_bar
!==========error term ~ N(0,12)===== =========
CALL UMACH (2, NOUT)
NR2 = 30
XSTD = 36
XM=0
!CALL RNSET (ISEED)
CALL RNNOA (NR2, R2)
CALL SSCAL (NR2,XSTD,R2,1)
CALL SADD (NR2,XM,R2,1)
!WRITE (NOUT,9999) R2
9999 FORMAT ('error normal random : ', 100F8.4)
```

```
CALL UMACH (2, NOUT)
NR2 = 30
XSTD = 36
XM=0
!CALL RNSET (ISEED1)
CALL RNNOA (NR2, R3)
CALL SSCAL (NR2,XSTD,R3,1)
CALL SADD (NR2,XM,R3,1)
!WRITE (NOUT,999) R3

999 FORMAT ('error normal random : ', 100F8.4)
!==================x_bar, y_bar=====================

yr=100+q-x+R2
yt=100-x+R3
!PRINT*,' yr is', yr
!PRINT*,' yt is', yt
!print*,'x_bar',x_bar
yr_bar=sum(yr)/NR2
!PRINT*,' yr_bar is',yr_bar
yt_bar=sum(yt)/NR2
!PRINT*,' yt_bar is', yt_bar
!===============sum of square=================

Sxx=sum((x-x_bar)**2)
!PRINT*,' Sxx is', Sxx
Syyr=sum((yr-yr_bar)**2)
!PRINT*,' Syyr is', Syyr
Syyt=sum((yt-yt_bar)**2)
!PRINT*,' Syyt is', Syyt
Sxyr=sum(((x-x_bar)*(yr-yr_bar)))
!PRINT*,' Sxyr is', Sxyr
Sxyt=sum(((x-x_bar)*(yt-yt_bar)))
!PRINT*,' Sxyt is', Sxyt
!================= ==========
br=Sxyr/Sxx
!print*,'br=',br
ar=yr_bar-br*x_bar
```

```
!print*,'ar=',ar

bt=Sxyt/Sxx
!print*,'bt=',bt
at=yt_bar-bt*x_bar
!print*,'at=',at
!=========== ============
Sr_square=(syyr-(br*sxyr))/(NR2-2)
!print*,'Sr^2=',Sr_square
St_square=(syyt-(bt*sxyt))/(NR2-2)
!print*,'St^2=',St_square
F1=Sr_square/St_square
!print*,'F=',F

if(F1 .lt. F)then
Sp_square=(Sr_square+St_square)/2
!print*,'Sp^2',Sp_square
end if
V=Sp_square*(2/Sxx)
t1=(abs(br-bt))/(sqrt(v))

if(t1 < TV) then
bc=(Sxyr+Sxyt)/(Sxx+Sxx)
!print*,'bc=',bc
end if
ar_new=yr_bar-bc*x_bar
at_new=yt_bar-bc*x_bar
!print*,'ar_new=',ar_new
!print*,'at_new=',at_new
M=(ar_new-at_new)
!print*,'M=',M
sigma_t=sum((((yt-at_new-bc*x)**2)/(NR2-2))
!print*,sigma_t
sigma_r=sum((((yr-ar_new-bc*x)**2)/(NR2-2))
var=((1.0/NR2)+(((30-x_bar)**2)/Sxx))*(sigma_t+sigma_r)
var_root=sqrt(var)
!print*,var_root
```
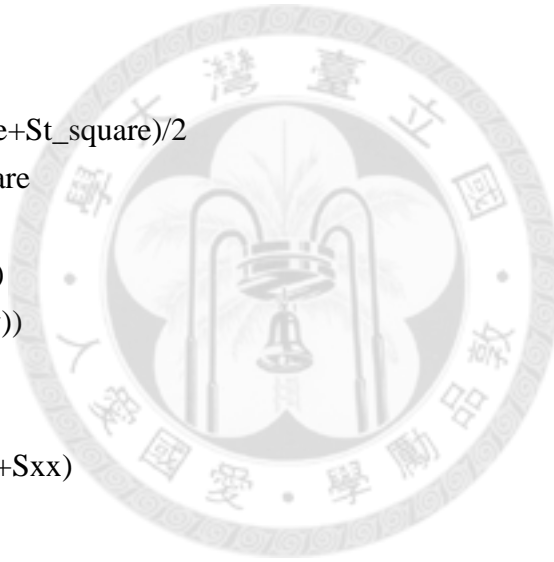
```fortran
mu=M+(tv*sqrt(var))
ml=M-(tv*sqrt(var))

!print*,k,M,ml,mu
!===================count power=========
cc1=0
if((ml > -10) .AND. (mu < 10)) cc1 = 1
oc1=oc1+cc1
!===== ======count coverage probability================
cc2=0
if((ml <   q) .AND. (mu >   q)) cc2 = 1
oc2=oc2+cc2
!=======sample size==================
sample=9.1827*(1/(Sp_square*2))
size=(sample-(((30-x_bar)**2)/Sxx))
sample_size=1/size
!print*,sample_size
!write(8,*) k,ml,mu
!,sample_size
!write(8,*) oc1,oc2
end do
write(*,*)"power=", oc1
write(*,*)"converage probability=", oc2
END program
```
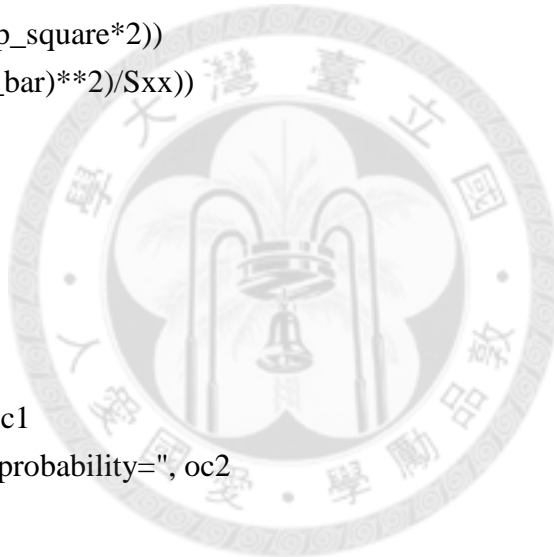
```fortran
!sstats.lib smaths.lib sstatd.lib smathd.lib
PROGRAM MAIN
!use IMSL
INTEGER ISEED,ISEED1, NOUT, NR,i,j,k,muk,power,oc1,oc2,q
REAL XM,XSTD,t(5),x_bar,yt_bar,yr_bar,Syyr,Syyt, nm,
Sxx,Sxy,m,g,tsc,mu,ml,t1,t11,t111,t1111,t11111,DFD,DFN,F,FIN,P,DF,TV,TIN
REAL R1(7),R11(5),R111(3),R1111(5),R11111(7),R2(27),R3(27),yr(27),yt(27),x(27)
EXTERNAL RNNOA, RNSET, UMACH, SSCAL, SADD, FIN , TIN
CALL UMACH (2, NOUT)
```

```
OPEN(8, FILE='M confidence interval.txt')

print*,"q="
read(*,*) q
!====產生 100 筆亂數===================
ISEED = 12345
!ISEED1 = 234
CALL RNSET (ISEED)
!CALL RNSET (ISEED1)
do k=1,10000
!=========================================
NR2=27
P=0.9
DFN=NR2-2
DFD=NR2-2
F=FIN(P,DFN,DFD)
!print*, F

P=0.95
DF=2*NR2-4
TV=TIN(P,DF)
!print*, TV
!===============mean =10   ======================
NR1 = 7
XSTD = 1
XM = 10
!CALL RNSET (ISEED)
CALL RNNOA (NR1, R1)
CALL SSCAL (NR1,XSTD,R1,1)
CALL SADD (NR1,XM,R1,1)
!WRITE (NOUT,99999) R1
99999 FORMAT ('normal random deviates: ', 20F8.4)
t1=SUM(R1)/NR1
!write(*,*)"-------------t1"
!write(*,*) t1
!=================mean =20================
NR11 = 5
XSTD = 1
```

```
XM = 20
!CALL RNSET (ISEED)
CALL RNNOA (NR11, R11)
CALL SSCAL (NR11,XSTD,R11,1)
CALL SADD (NR11,XM,R11,1)
!WRITE (NOUT,99998) R11
99998 FORMAT ('normal random deviates: ', 20F8.4)
t11=SUM(R11)/NR11
!write(*,*)"-------------t11"
!write(*,*) t11
!=================mean =30=================
NR111 = 3
XSTD = 1
XM = 30
!CALL RNSET (ISEED)
CALL RNNOA (NR111, R111)
CALL SSCAL (NR111,XSTD,R111,1)
CALL SADD (NR111,XM,R111,1)
!WRITE (NOUT,99997) R111
99997 FORMAT ('normal random deviates: ', 20F8.4)
t111=SUM(R111)/NR111
!write(*,*)"-------------t111"
!write(*,*) t111
!==================mean =40==================
NR1111= 5
XSTD = 1
XM = 40
!CALL RNSET (ISEED)
CALL RNNOA (NR1111, R1111)
CALL SSCAL (NR1111,XSTD,R1111,1)
CALL SADD (NR1111,XM,R1111,1)
!WRITE (NOUT,99996) R1111
99996 FORMAT ('normal random deviates: ', 20F8.4)
t1111=SUM(R1111)/NR1111
!write(*,*)"-------------t1111"
!write(*,*) t1111
!==================mean =50==================
NR11111 = 7
```

```
XSTD = 1
XM = 50
!CALL RNSET (ISEED)
CALL RNNOA (NR11111, R11111)
CALL SSCAL (NR11111,XSTD,R11111,1)
CALL SADD (NR11111,XM,R11111,1)
!WRITE (NOUT,99995) R11111
99995 FORMAT ('normal random deviates: ', 20F8.4)
t11111=SUM(R11111)/NR11111
!write(*,*)"------------t11111"

!write(*,*) t11111


!===== ========combine array=========================
x=(/R1,R11,R111,R1111,R11111/)
x_bar=(t1+t11+t111+t1111+t11111)/5
!write(*,*)"------------x"

!write(*,*) x

!write(*,*) x_bar

!===========error term ~ N(0,12)============ =========
CALL UMACH (2, NOUT)
NR2 = 27
XSTD = 12
XM=0
!CALL RNSET (ISEED)
CALL RNNOA (NR2, R2)
CALL SSCAL (NR2,XSTD,R2,1)
CALL SADD (NR2,XM,R2,1)
!WRITE (NOUT,9999) R2
9999 FORMAT ('error normal random : ', 100F8.4)
CALL UMACH (2, NOUT)
NR2 = 27
XSTD = 12
XM=0
!CALL RNSET (ISEED1)
```

```fortran
CALL RNNOA (NR2, R3)
CALL SSCAL (NR2,XSTD,R3,1)
CALL SADD (NR2,XM,R3,1)
!WRITE (NOUT,999) R3


999 FORMAT ('error normal random : ', 100F8.4)
!=================x_bar, y_bar===================
yr=100+q-x+R2
yt=100-x+R3
!PRINT*,' yr is', yr
!PRINT*,' yt is', yt
!print*,'x_bar',x_bar
yr_bar=sum(yr)/NR2
!PRINT*,' yr_bar is',yr_bar
yt_bar=sum(yt)/NR2
!PRINT*,' yt_bar is', yt_bar
!===============sum of square=========================

Sxx=sum((x-x_bar)**2)
!PRINT*,' Sxx is', Sxx
Syyr=sum((yr-yr_bar)**2)
!PRINT*,' Syyr is', Syyr
Syyt=sum((yt-yt_bar)**2)
!PRINT*,' Syyt is', Syyt
Sxyr=sum(((x-x_bar)*(yr-yr_bar)))
!PRINT*,' Sxyr is', Sxyr
Sxyt=sum(((x-x_bar)*(yt-yt_bar)))
!PRINT*,' Sxyt is', Sxyt
!==================== ===============
br=Sxyr/Sxx
!print*,'br=',br
ar=yr_bar-br*x_bar
!print*,'ar=',ar

bt=Sxyt/Sxx
!print*,'bt=',bt
at=yt_bar-bt*x_bar
!print*,'at=',at
```
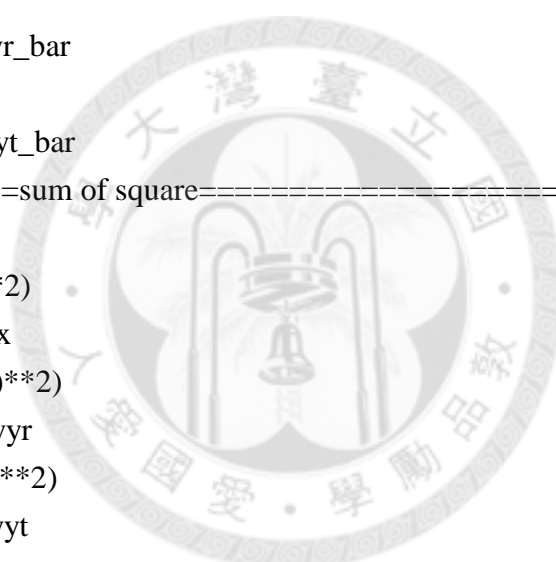
```fortran
!==================== ====================
Sr_square=(syyr-(br*sxyr))/(NR2-2)
!print*,'Sr^2=',Sr_square
St_square=(syyt-(bt*sxyt))/(NR2-2)
!print*,'St^2=',St_square
F1=Sr_square/St_square
!print*,'F=',F

if(F1 .lt. F)then
Sp_square=(Sr_square+St_square)/2
!print*,'Sp^2',Sp_square
end if
V=Sp_square*(2/Sxx)
t1=(abs(br-bt))/(sqrt(v))

if(t1 < TV) then
bc=(Sxyr+Sxyt)/(Sxx+Sxx)
!print*,'bc=',bc
end if
ar_new=yr_bar-bc*x_bar
at_new=yt_bar-bc*x_bar
!print*,'ar_new=',ar_new
!print*,'at_new=',at_new
M=(at_new-ar_new)/bc
!print*,'M=',M
g=((TV**2)*Sp_square)/((bc**2)*(Sxx+Sxx))
!print*,'g=',g
tsc=(TV*sqrt(Sp_square))/abs(bc)
nm=(2.0/NR2)+((M**2)/(Sxx+Sxx))
!print*,nm
mu=M+(tsc*sqrt(nm))
ml=M-(tsc*sqrt(nm))

!print*,k,M,ml,mu
!==================count power==================
cc1=0
if((ml > -10) .AND. (mu < 10)) cc1 = 1
oc1=oc1+cc1
```
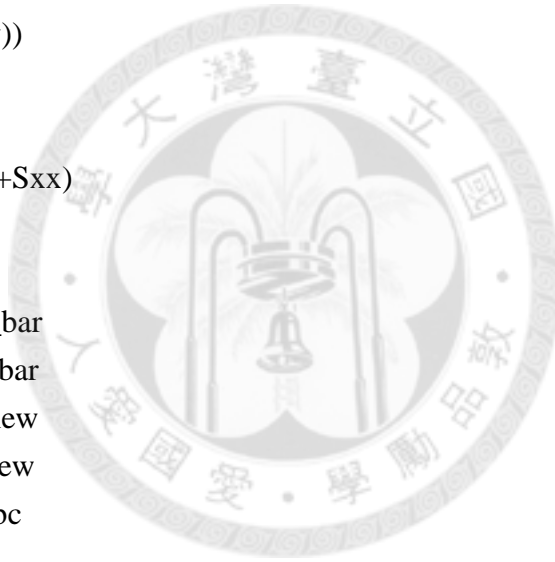
```
!==== =======count converage probability================
cc2=0
if((ml < q) .AND. (mu > q)) cc2 = 1
oc2=oc2+cc2


!========sample size============
sample=9.1827*(1/(Sp_square*2))
size=(sample-(((30-x_bar)**2)/Sxx))
sample_size=1/size


!print*,sample_size



write(8,*) k,ml,mu
!,sample_size
!write(8,*) oc1,oc2
end do
write(*,*)"power=", oc1
write(*,*)"coverage probability=", oc2
END program
```