STATISTICAL PROPERTIES OF RATIO MEASURES BASED ON THE PRE- AND POST-DATA

TAKAHARU YAMABE

MARCH 2012

STATISTICAL PROPERTIES OF RATIO MEASURES BASED ON THE PRE- AND POST-DATA

A dissertation submitted to
THE GRADUATE SCHOOL OF ENGINEERING SCIENCE
OSAKA UNIVERSITY
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN ENGINEERING

BY
TAKAHARU YAMABE

MARCH 2012

謝辞

本論文の作成におきましては，多くの方にご指導・ご支援を賜りました．ここに深くお礼を申し上げます．

指導教員の白旗慎吾先生には，本論文を通して，多大なご教示を頂きました．本稿を丁寧にご査読いただき，多くの貴重なご指摘を頂戴いたしました．ダブリンで開催された国際学会（ISI, 2011）にご一緒させていただいた際には，奥様と共に気さくな雰囲気で声をかけていただくことで，リラックスした雰囲気で発表に望むことができました．心よりお礼を申し上げるとともに，今後もますますのご高配のほど，よろしくお願い申し上げます．大阪大学教授（大学院基礎工学研究科統計数理講座）の狩野 裕先生には，本論文の初稿を査読いただき，公聴会では貴重なご指摘を賜りました．大阪大学教授（大学院基礎工学研究科数理計量ファイナンス講座）の内田雅之先生には，本論文の初稿を査読いただき，公聴会では貴重なご指摘を賜りました．上記の先生方にはお礼申し上げます．

NPO 法人医学統計研究会（Biostatistical Research Association: BRA）の理事長である後藤昌司先生には，本研究の主題をご提示いただくとともに，本論文およびそれに関わる公表論文作成，学会発表などの全ての過程でご指導を賜りました．大阪大学教授時代に教授室ではじめてお会いした際，「学問とは哲学がなければ駄目で，化学は学問とは認めない」という言葉は強烈に今でも脳裏に焼きついています．「化学…」という箇所は今でも納得がいきませんが，「哲学が大事であること」，「遊学一如の世界も大事であること」，「たくさんの仲間を大事にすること（金持ちというより人持ちになること）」など，学問に加えて人生の教訓となるたくさんの教えを学ぶことができ，現在の業務・生活の糧となっているのは，後藤先生の厳しくも愛情あふれるご指導のおかげです．本当にありがとうございました．

本論文の作成にあたり，BRA の皆様には多大なご支援と心温まるご配慮を頂きました．長崎大学の柴田義貞先生には，BRA のシンポジウム・フォーラム等でお話しする機会を通して，大変に啓発されました．研究会での厳しい指導，鋭いご指摘に加えて暖かいご助言を賜る機会に恵まれたため，私の研究で重要なヒントを得ることができ，そして柴田先生の研究（特に東日本大震災後の原発に関する研究）に対する姿勢には大変感銘を受けました．大分大学教授の越智義道先生には，大分で筆者の研究発表後にご意見をいただきました．鹿児島高等専門学校教授の藤崎恒晏先生には，私の父が同じ高等専門学校で教鞭をとっていることもあり，BRA の会合でお会いするたびに気さくに話しかけて頂きました．大阪大学准教授の坂本亘先生には本論文の初稿を確認

いただきました．ダブリンで開催された国際学会（ISI，2011）にご一緒させていただいた際には，研究会等で見せる鋭い顔に加えて，白旗先生と共に気さくに声をかけていただき安心して発表が出来ました．心よりお礼申し上げます．大阪大学准教授の濱﨑俊光先生には博士後期課程に進学する際に色々なご助言を頂きました．弘前大学准教授の杉本知之先生には BRA の諸会合でお会いすると温かく励ましていただきました．山梨大学准教授の下川敏雄先生には山梨大学での研究会などでお世話になりました．先生の成果物の多さと行動力のには大変刺激を受けました．兵庫医科大学講師の大門貴志先生には統計数理研究所でベイズの講義を拝聴する機会に恵まれました．先生のわかりやすい講義により未知の世界であったベイズに興味がわきました．上記の先生方に重ねて御礼申し上げます．

魚井技術士事務所の魚井 徹博士には，BRA の会合でお会いした際に，励ましのお言葉を頂戴いたしました．臨床情報研究センター [財団法人先端医療振興財団] の松原義弘博士には BRA 会合でお会いした際にお声をかけていただきました．株式会社フィールドワークスの木田義之さんには，BRA 会合でお会いした際に幾度も激励のお言葉を声をかけていただきました．株式会社ソリューションズラボの志賀 功さんには大分に訪れた際，東京や大阪のシンポジウムでお声をかけていただきました．株式会社富士通大分ソフトウェアラボラトリの衛藤俊寿博士には，大分統計談話会の折に遊学をご一緒させていただきました．第一三共株式会社の佐藤俊之博士には，BRA の会合でお会いするたびに貴重なご意見を頂きました．小野薬品工業の冨金原悟博士には，筆者の発表について実用的な観点からご指摘を頂きました．また，学会終了後，課題検討会等のお酒の席でご馳走になりました．ファイザー株式会社の栗林和彦博士には，筆者が大学院に入学し，学位を取得しようとしていた時に色々とご助言頂きました．株式会社クリニカル スタディ サポートの磯村達也さんには，BRA のシンポジウム等でご一緒させて頂き，製薬会社とは別の視点から医療の現場でおきている事柄を教えて頂きました．ファイザー株式会社の河合統介博士には，BRA の先輩として，会社の上司として，共同研究者として数多くのご助言と鋭いご指摘を賜りました．また酒席でも周りに配慮し，皆を盛り上げ，何時も楽しい時間を過ごさせて頂きました，大変感謝しています．あすか製薬の藤澤正樹博士には，定例シンポジウム等，BRA の行事にご一緒させて頂き，研究に対する姿勢等，人生の先輩として貴重なご意見を頂きました．エーザイ株式会社の高瀬貴夫さんには，BRA 定例研究会，大分統計談話会等で貴重なご意見を頂きました．また，大分空港の海甲でご馳走になった関サバ，関アジの味は忘れられません．協和発酵キリンの古川泰信さんには，ご自身の研究である生物学的同等性に関する学会発表資料を頂き，勤行の一助とさせて頂きました．アステラス製薬株式会社の伊藤雅憲博士には，筆者の研究に対して参考となる論文を頂きました．また，BRA シンポジウム，遊学の会合を企画する能力と気力には大変刺激を受け，楽しい時を過ごさせて頂いています．ノバルティスファーマの池田公俊さんには，ご自身の勤行，特にアメリカ滞在における貴重な体験談を教えて頂きました．株式会社ベルシステム 24 の金水龍さんには，BRA の諸会合でご一緒させて頂きました．アスビオファーマ株式会社の永久保太士博士には，てんかんの文献データを教えて頂きました．興和株式会社の丸尾和司博士には，共同研究者として色々相談に乗って頂きました．特に数理的な観点から深く，そして正確

な御助言には大変助けられましたし，筆者が研究でくじけそうになった時には御自身の経験から温かい励ましの言葉を頂きました．亀戸での研究相談，Skype による議論は大変思い出に残っております．これからも宜しくお願いします．株式会社 GREE の元垣内広毅さんには，お会いするたびに色々な仕事上の経験をお伺いして大変に刺激を受けました．大塚製薬工場の大江基貴さんには，ご自身の研究である ROC 曲線の推測について興味深い知見と深い考察を拝聴して大変勉強になりました．大日本住友製薬の中村将俊さんには大阪でお会いした時に色々とお世話になりました．ファイザー株式会社の五十川直樹さんには，ご自身の研究の 1 つである差の分布について，詳しくご教授頂き，私の研究を前に進める大きなヒントを頂きました．加えて，同じ会社ということもあり遊学におけるユーモアな立ち振る舞いを拝見させていただきました．大阪大学大学院博士後期課程の山口裕介さんには，大学関係の資料の取り寄せにご協力いただき，お世話になりました．また，五十川直樹さんと共にダブリンで遊学をご一緒させていただきました．大阪大学大学院博士前期課程の吉川隆範さん，大山秀輔さん，横山隼人さんには帰阪した際に色々とお世話頂きました．日本臓器製薬の尾崎寿昭さんと株式会社ベルシステム 24 の池田敏広さんには BRA のフォーラム等，楽しい酒食を共にさせて頂きました．トーアエイヨー株式会社の川端ゆみこさんには，計算機統計学会でご発表された資料を頂き，私の研究のヒントを頂きました．株式会社新日本科学の古賀正さんには，BRA のシンポジウムで BRA のフォーラム，シンポジウム等でご一緒させて貴重なご自身の経験談を紹介していただきました．特に，「FDA と相談した体験談」をお伺いした時には大変啓発されました．後藤昌司先生の奥様の後藤孚さんには，お会いするたびに声をかけて頂き，千里中央の事務所を訪問した際には時々，手作りの美味しい昼食をごちそうになりました．BRA 書記の亀山日出子さんには千里中央の事務所を訪問した際に美味しいコーヒーを入れて頂き，また筆者の英語のレビューをしていただきました．上記の方々に御礼申し上げます．ファイザー株式会社臨床統計部第 2 統計グループリーダーの丸山奈美博士には業務と研究にご配慮いただき，温かい励ましを頂きました．勤行，学問に真摯に取り組む姿勢にはいつも啓発されました．同グループの吉山保さんには，ガバペンチンのデータについて示唆に富んだご意見を頂戴いたしました．大倉征幸博士には R のシミュレーションプログラムについて，多大なご助言を頂きました．中水流嘉臣さんには，社会人の同期として勤行・遊学共にお付き合いいただきました．豊泉滋之さんには，勤行，特にモデルの検討，についてご自身で検討されている内容を教えて頂き，大変参考になりました．皆様に御礼申し上げます．

最後に絶えず，筆者の身を案じ，励ましてくれた妻と両親，妹弟に感謝いたします．

Abstract

In a clinical trial, we sometimes evaluate the treatment effect based on the ratio measures which requires pre- and post-data of treatment intervention. As a measure of ratio, percent change from baseline (PC) which is defined as $PC = (X_2 - X_1)/X_1$ is often used in a trial. And, symmetrized percent change (SPC) which is defined as $SPC = (X_2 - X_1)/(X_1 + X_2)$ is sometimes also used in trials(Berry,1989). Though the statistical properties of PC were investigated on condition that pre- and post-data are assumed as bivariate normal distribution in past research, PC is said to have some difficulties to apply the statistical analysis based on the parametric methods (Asakura *et al.*, 2011; Senn & Julious, 2009). On the other hand, SPC is said to have good performance based on a limited simulation, but is said to have difficulties in interpretation (Berry, 1989; Berry & Ayers, 2006).

As I mentioned in the above paragraph, PC and SPC as the ratio measures are investigated in some aspects. However, past findings are based on limited research such as the investigations of PC assumed as the bivariate normal distribution in pre- and post-data. In a clinical trial, data follows not only normal distribution but also positive skew distribution such as log-normal distribution or more positive skew distribution than log-normal(Maruo *et al.*, 2008). Therefore, we need to investigate the statistical properties of two ratio measures, PC and SPC, in various distributions of pre- and post-data. In this paper, we declare the probability distribution function (pdf) of two ratio measures, percent change (PC) and symmetrized percent change (SPC), and evaluate the relationship between the skewness of two ratio measures and the distribution of pre- and post-data. Next, we evaluate the performance of two ratio measures to detect the treatment difference within pre- and post-data or between two groups based on the simulation and propose how to apply two measures in various situations. In addition, we declare the relationship between ratio measure (SPC) and coefficient of variation (CV).

Acknowledgments

The author would like to express his deep and sincere gratitude to many people who gave much suggestions, helps and encouragements throughout the preparation of his dissertation.

The author appreciates Professor Shingo Shirahata of Osaka University for providing helpful and useful comments. Professor Yutaka Kano of Osaka University provided the helpful comments. Professor Masayuki Uchida of Osaka University also provided the helpful comments. The author is deeply grateful to them.

The author especially appreciates Dr. Masashi Goto of Biostatistical Research Association (BRA), NPO, who led the author to the theme of their research, provided important and useful suggestions and encouraged me to initiate this research.

Professor Yoshisada Shibata of Nagasaki University for valuable comments and advice about the interpretation of his research. Associate professor Wataru Sakamoto of Osaka University reviewed this thesis and helped his research. The author would like to thank them.

The author wishes to thank all of members of BRA. In BRA meetings, he received helpful advices and comments, especially he would like to thank Dr. Kazuhiko Kuribayashi, Dr. Toshimitsu Hamasaki, Dr. Norisuke Kawai, Mr. Tadashi Koga, Dr. Satoru Fukinbara, Dr. Masaki Fujisawa, Dr. Tomoyuki Sugimoto, Dr. Toshio Shimokawa, Mr. Takao Takase, Mr. Yasunobu Furukawa, Dr. Masanori Ito, Mr. Kimitoshi Ikeda, Dr. Takashi Nagakubo, Dr. Kazushi Maruo, Mr. Hiroki Motogaito, Mr. Motoki Ohe, Mr. Masatoshi Nakamura, Mr. Naoki Isogawa and Mr. Yusuke Yamaguchi.

The author would like to thank all the members of clinical statistics dept. at Pfizer Japan for their continuing kindness and substantial support.

Finally, the author is grateful to his parents, brothers and wife for their support.

Notations

notation	definition/example	explanation
general		
$\mathrm{E}[\cdot]$	$\mathrm{E}[X]$	expectation
$\mathrm{Var}[\cdot]$	$\mathrm{Var}[X]$	variance
$\xi_{\cdot}$	$\xi_{0.5}$	percentile
distribution		
X_i		random variable on pre- and post-data
x_i		observed value on pre- and post-data
λ_i		shape parameter (transformation parameter)
μ_i		location parameter
σ_i		scale parameter
BN	$\mathrm{BN}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$	bivariate normal distribution
BLN	$\mathrm{BLN}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$	bivariate log normal distribution
BPN	$\mathrm{BPN}(\lambda_1, \lambda_2, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$	bivariate power-normal distribution
$f_{\cdot}(\cdot)$	$f_{\mathrm{BPN}}(x_1, x_2)$	probability density function (pdf)
$F_{\cdot}(\cdot)$	$F_{\mathrm{BPN}}(x_1, x_2)$	cumulative distribution function (cdf)
$\phi(\cdot)$	$\phi(x)$	pdf of standardized normal distribution
$\Phi(\cdot)$	$\Phi(x)$	cdf of standardized normal distribution

Contents

Section 1

Introduction

1.1 Background

The designs with pre- and post-data fall under the broad category of paired data analysis. Paired data arise when the same experimental unit, such as a person or laboratory animal, is measured on some variable in two different timings or at the same time under different testing conditions. A type of the design with pre- and post-data is when subjects receive a treatment intervention prior to the measurement of the post-data, after collecting the pre-data. And the question of interest is either that there are differences among groups or changes in an individual over time. For example, an object of a clinical trial is to compare the treatment groups with intervention, and it is said that pre-defined measures for evaluating the intervention is very important (Tsubaki, 1999).

In the case of a treatment evaluation of disease, and especially in the evaluation of the efficiency of a particular drug, we sometimes use an index based on the change seen in pre- and post-treatment data of the drug, specifically used in a certain disease area. Generally, the index (measure) is considered as a categorical scale, ordinal scale or interval scale. In particular, the measure of the continuous data is based on a difference or a ratio of pre-(X_1) and post(X_2) data.

The appropriate measure is selected according to the balance of both the clinical and statistical points of view. The clinical point comes from ease of the interpretation and the statistical point comes from the ease of data analyses based on the normal distribution. In a particular experiment, a choice of difference or ratio as the primary measure of treatment effect may not be obvious. Statistically, the principal reasons to adjust for baseline, usually presented in relation to analysis of covariance, are to remove concomitant variation in the response and improve the precision of treatment comparisons (Steel & Torrie, 1980). Furthermore, summary statistics of an adjusted response should be independent response (Kaiser, 1989). Another relevant question is about the kind of effect anticipated. For example, is it additive, multiplicative or neither?

Compared with difference, ratio measures are not always investigated from the statistical point of view. Thus, this paper focuses on the statistical property of ratio measures. In later paragraphs of this section, firstly, we show the statistical properties of general ratio measures. Secondly, we review the past findings of two ratio measures. Lastly, we show some applicable examples of these two ratio measures in clinical trials and some examples of the difference often seen in clinical trials.

Statistical properties of general ratio measures. Relative change scores as ratio measure require the pre- and post-data to be continuous random variables. Thus ensuring the change score (difference) to be a continuous random variable, relative change scores also require the pre- and post-data to be the same type of measurement made using the same device and have equal units of measurement. Although pre- and post-data have the same units, relative change scores are often unitless or expressed as percentages.

Relative change scores convert the pre- and post-data into a proportional change score, C, expressed as either raw change (difference) or absolute change. The formula to convert pre- and post-data can be written as

$$C = \frac{X_2 - X_1}{X_1}$$

in the case of raw change, or

$$C = \frac{|X_2 - X_1|}{X_1}$$

in the case of absolute change, where C is the change score, X_2 is the post-data and X_1 is the pre-data(Bonate, 2000). Note that the numerator is a difference whereas the denominator scale is the pre-data. A variant of these equations is to multiply the proportional change scores by 100 thereby converting them to percent change scores. If $C = 0$, no change has occured. A positive relative change score indicates that the post-data was greater than the pre-data, whereas a negative relative change indicates that the post-data was less than the pre-data. One criticism of relative change score is in the choice of the scaling term or denominator. Consider an individual whose initial score is 3 and whose final score is 7. Using Eq. 1.1, this represents a 133 % increase from baseline. However, if a patient scores a 7 initially and deteriorates to a 3, a -57 % decrease has occured. Hence, different denominator terms result in different transformations and estimates of change.

Proportional and percent change score fall under a family of transformations known as change functions. Törnqvist, *et al.,*(1985) formally defined a change function as a real-value function $C(X_1, X_2)$ of positive arguments, $C : \mathbf{R}_2^+ \to \mathbf{R}$ with the following properties:

1. $C(X_1, X_2) = 0, \quad$ if $X_1 = X_2$
2. $C(X_1, X_2) > 0, \quad$ if $X_1 > X_2$
3. $C(X_1, X_2) < 0, \quad$ if $X_1 < X_2$
4. C is a continuous increasing function of X_2 when X_1 is fixed.
5. $\forall a : a > 0 \;\rightarrow\; C(aX_1, aX_2) = C(X_1, X_2)$

The last property merely states that the function is independent of units of measurement. The property $C : \mathbf{R}_2^+ \rightarrow \mathbf{R}$ states that a two-demensional vector $(\mathbf{R}_2^+)$ is mapped into a one-dimensional vector $(\mathbf{R})$ by the function C. It can be shown that both proportional percent change functions meet these requirements. It can also be shown that difference scores represent another valid type of change function.

By setting $a = 1/X_1$ in property 5, Törnqvist, *et al.*,(1985) have shown that almost every indicator of relative change can be expressed as a function of X_2/X_1 alone. Hence, the change function can be expressed as an alternate function dependent solely on X_2/X_1. Formally, there exists a function H, such that

$$C(X_1, X_2) = H\left(\frac{X_2}{X_1}\right) = C\left(1, \frac{X_2}{X_1}\right)$$

with properties:

1. $H\left(\frac{X_2}{X_1}\right) = 0, \quad$ if $\frac{X_2}{X_1} = 1$
2. $H\left(\frac{X_2}{X_1}\right) > 0, \quad$ if $\frac{X_2}{X_1} > 1$
3. $H\left(\frac{X_2}{X_1}\right) < 0, \quad$ if $\frac{X_2}{X_1} < 1$
4. H is a continuous increasing function of its argument $\frac{X_2}{X_1}$.
5. $H\left(\frac{aX_2}{aX_1}\right) = H\left(\frac{X_2}{X_1}\right)$ trivially

Table 1.1 shows a variety of other relative change functions proposed by Törnqvistnqvist, *et al.*(1985) and their simplification into functions of $Y = (X_2/X_1)$. Here, $\mathrm{K}(X_1, X_2)$ is any mean of X_1 and X_2.

Table 1.1: Relative change functions and their simplification into functions of $Y(= X_2/X_1)$ as presented by Törnqvist *et al.*(1985)

Mapping	Function	Mapping	Function
$\dfrac{X_2 - X_1}{X_1}$	$Y - 1$	$\dfrac{X_2 - X_1}{2(X_1^{-1} + X_2^{-1})^{-1}}$	$\dfrac{1}{2}(Y-1)(1+1/Y)$
$\dfrac{X_2 - X_1}{X_2}$	$1 - \dfrac{1}{Y}$	$\dfrac{X_2 - X_1}{\min(X_1, X_2)}$	$\dfrac{Y-1}{\min(1, Y)}$
$\dfrac{X_2 - X_1}{(X_1 + X_2)/2}$	$\dfrac{Y-1}{(1+Y)/2}$	$\dfrac{X_2 - X_1}{\max(X_1, X_2)}$	$\dfrac{Y-1}{\max(1, Y)}$
$\dfrac{X_2 - X_1}{\sqrt{X_1 X_2}}$	$\dfrac{Y-1}{\sqrt{Y}}$	$\dfrac{X_2 - X_1}{\mathrm{K}(X_1, X_2)}$	$\dfrac{Y-1}{\mathrm{K}(1, Y)}$

Statistical properties of two ratio measures. In this paragraph, two ratio measures which has been applied in the clinical trial are shown. As a measure of ratio, percent change from baseline (PC), $PC = (X_2 - X_1)/X_1$, is often used in a trial. In addition, symmetrized percent change (SPC), $SPC = (X_2 - X_1)/(X_1 + X_2)$, are sometimes also used in trials(Berry,1989). Bonate(2000) and Törnqvistnqvist, *et al.* (1985) shows the modified SPC which is defined as the mean of two values for a numerator which is $(X_2 - X_1)/\left\{\frac{1}{2} \times (X_1 + X_2)\right\}$.

The PC means "the proportion of increase (or decrease) for pre-value", and is acceptable from the clinical point of view because of the easy interpretation. On the other hand, some statistical difficulties are pointed out to PC. Senn & Julious(2009) said that the statistical analysis based on the parametric are not recommended for PC, because PC(or ratio of two values) is not normal even if pre- and post-data are normal. Asakura *et al.*,(2011) investigated the statistical properties of ratio on condition that two values are normal, summarized the statistical issues of ratio and gave a warning for using the ratio to the estimation of effect. Pharm-Gia *et al.*(2006) gave the exact closed form expression of the density of X_2/X_1 , where X_1 and X_2 are normal random variables, in terms of Hermite and confluent hypergeometric functions, and show the skewness distribution in some situation. On the other hand, Berry (1989) introduced the SPC as the modified percent change with good statistical properties in the medical field. Brouwers & Mohr(1989) argued that the advantage of using SPC over the PC is that the transformed variable dose not depend on the denominator used in the transformation and the resultant distribution is symmetrical about its mean. Berry & Ayers(2006) showed the simulation results under independent, additive and multiplicative correlation structures of pre- and post-data for parametric and nonparametric analyses. And Berry & Ayers(2006) concluded that simple ANOVA on SPC had power equal or greater than alternative analysis methods except for independence structure. However, the interpretation of SPC may not be intuitive for those accustomed to thinking in terms of PC. For example, if SPC is -0.1 or -0.2, then the post-data shows to reduce form pre-data, but it is difficult to interpret the value of -0.1 or -0.2. Concerning this point, Koti(2001) suggested that SPC is obscurant in nature. However, the same can be said for many statistical methods that are valuable in making inferences, such as taking the logarithm and most nonparametric tests(Berry & Ayers, 2006). For interpretability of analysis results, Berry (1989) suggested transforming SPC to the PC scale using the inverse transformation: robust percent change $RPC = 2 \times SPC/(1 - SPC)$. For example, if SPC is equal to -0.25 for a particular treatment arm, then $RPC = -0.4$.

Application example of difference or ratio measures in clinical trial. As some example of measures, the difference which is defined as $D = X_2 - X_1$ is used for the treatment evaluation for patients with high-blood pressure based on the diastolic blood pressure or systolic blood pressure (Adachi *et al.*, 2009), for patients with pain, such as neuropathic pain or pain of osteoarthritis of the knee, based on the 11-point rating scale or 100mm visual analog scale (Satoh et al., 2010: Lane *et al.*, 2010) and for patients with glaucoma based on the ocular pressure (Kitazawa *et al.*, 2009)

As the ratio measures, the percent change (PC) which is defined as $PC = (X_2 - X_1)/X_1 = (X_2/X_1) + 1$ are often used for treatment evaluation. On the other hand, symmetrized percent change which is defined as $SPC = (X_2 - X_1)/(X_1 + X_2) = \{(X_2/X_1) - 1\} / \{1 + (X_2/X_1)\}$ are sometimes used. As examples of clinical evaluation, PC are applied to the treatment evaluation of patients with high-density lipoprotein cholesterol (Adachi *et al.*, 2009), of patients with urge to urinate or urge incontinence based on the number of acraturesis (Homma *et al.*, 2003), of patients with climacteric disorder based on the number of hot flush (Endrikat *et al.*, 2007). SPC is applied to the treatment evaluation of patients with partial epilepsy based on the seizure frequency(Yamauchi *et al.*, 2006) and evaluation of male patients with osteoporotic fracture based on the physical activity (anney et al., 2010).

1.2 Motivation

As I mentioned in the previous section, PC and SPC as the ratio measures are investigated in some aspects. However, past findings are based on limited researches such as the investigations of PC assumed as the bivariate normal distribution in pre- and post-data. In a clinical trial, data follows not only normal distribution but also positive skew distribution such as log-normal distribution or more positive skew distribution than log-normal(Maruo *et al.*, 2008). Therefore, we need to investigate the statistical properties of two ratio measures, PC and SPC, in various distributions of pre- and post-data. In this paper, we show more deeply investigation of two ratio measures as follows,

- We derive the probability distribution function (pdf) of two ratio measures, percent change (PC) and symmetrized percent change (SPC)
- We evaluate the relationship between the skewness of two ratio measures and the distribution of pre- and post-data
- We evaluate the performance of two ratio measures to detect the treatment difference within pre- and post-data or between two groups based on the simulation
- We propose how to apply the two measures in various situations
- We show the relationship between ratio measure (SPC) and coefficient of variation (CV)

1.3 Components of this paper

In section 2, we define the three kinds of distributions of the pre- and post-data, which are bivariate normal, bivariate log-normal and bivariate power normal distribution, and review some statistical properties of the distributions. In section 3, we derive the probability density function (pdf) of ratio measures and evaluate the skewness of distribution in each condition. In addition, we declare the relationship between ratio measures and coefficient of variation between pre- and post data with correlation. In section 4, we conduct simulations to evaluate the performance to detect the treatment difference within pre- and post-data or between two groups based on the simulations. In addition, we show a case example to apply SPC. In section 5, we describe the summary results, findings of this research and future investigation plan.

Section 2

Definition of the distributions for pre- and post-data

In this chapter, firstly, we introduce bivariate normal distribution and bivariate log-normal distribution assumed as pre- and post-data distribution generally used. However, distributions of pre- and post-data in real situations such as in clinical trials are sometimes not based on these two distributions. Therefore, we also introduce the bivariate power normal distribution and will evaluate properties of ratio measures comprehensively in a later chapter.

2.1 Commonly used distributions for pre- and post-data

Bivariate normal distribution (BN). Let the random variables $X_i(i = 1, 2)$ denote the response of pre- and post-data following bivariate normal distribution, and the variables satisfy $(X_1, X_2) \sim \mathrm{BN}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$, where μ_i is the location parameters, σ_i is the scale parameters and ρ is the correlation parameter between two random variables of pre- and post-data. Then, the probability density function (pdf) of random variable $X_i(i = 1, 2)$ which follows a bivariate normal distribution is,

$$\begin{aligned} f_{BN}(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \\ &\quad \times \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}\right]. \end{aligned}$$

Bivariate log-normal distribution (BLN). Let the positive random variables $X_i(i = 1, 2)$ denote the response of pre- and post-data following bivariate log-normal distribution, and the variables satisfy $(X_1, X_2) \sim \mathrm{BLN}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. Then, the pdf of random variable $X_i (i = 1, 2)$ which follows a bivariate log-normal distribution is,

$$f_{BLN}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}x_1x_2}$$
$$\times \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{\log x_1 - \mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{\log x_1 - \mu_1}{\sigma_1}\right)\left(\frac{\log x_2 - \mu_2}{\sigma_2}\right) + \left(\frac{\log x_2 - \mu_2}{\sigma_2}\right)^2\right\}\right]$$

2.2 Comprehensive distribution for pre- and post-data

Bivariate power normal distribution(BPN). Bivariate power normal distribution is a parametric class of probability distributions which includes the bivariate truncated normal and the bivariate log-normal as a special case. The bivariate power normal distribution is on the basis of the Box and Cox power-transformation which is defined by positive random variables $X_i (i = 1, 2)$

$$X_j^{(\lambda_j)} = \begin{cases} \dfrac{X_j^{\lambda_j} - 1}{\lambda_j} & \lambda_j \neq 0 \\ \log X_j & \lambda_j = 0 \end{cases} \tag{2.1}$$

where the range of $X_j^{(\lambda_j)}$ is $-1/\lambda_j < X_j^{(\lambda_j)} < +\infty$ when $\lambda_j > 0$ and is $-\infty < X_j^{(\lambda_j)} < -1/\lambda_j$ when $\lambda_j < 0$.

Let a power transformed variables $X_i^{(\lambda_i)}$ of X_i denote the truncated bivariate normal distribution with mean vector $\mu = (\mu_1, \mu_2)^{\mathrm{T}}$ and variance covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}. \tag{2.2}$$

Then, (X_1, X_2) is to have the bivariate power-normal distribution if the marginal pdf is

$$f_{BPN}(x_1, x_2) = \frac{x_1^{\lambda_1 - 1} x_2^{\lambda_2 - 1}}{A(\boldsymbol{K})} g_{BPN}\left(x_1^{\lambda_1 - 1}, x_2^{\lambda_2 - 1}\right), \quad x_1, x_2 > 0 \tag{2.3}$$

where

$$g_{BPN}\left(x_1^{\lambda_1 - 1}, x_2^{\lambda_2 - 1}\right) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{Q(x_1^{(\lambda_1)}, x_2^{(\lambda_2)})}{2}\right\} \tag{2.4}$$

and

$$Q(x_1^{(\lambda_1)}, x_2^{(\lambda_2)}) = \frac{1}{1-\rho^2} \times$$
$$\left\{\left(\frac{x_1^{(\lambda_1)} - \mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1^{(\lambda_1)} - \mu_1}{\sigma_1}\right)\left(\frac{x_2^{(\lambda_2)} - \mu_2}{\sigma_2}\right) + \left(\frac{x_2^{(\lambda_2)} - \mu_2}{\sigma_2}\right)^2\right\}$$

where λ_j, μ_j and σ_j are shape, location and scale parameters and ρ is a correlation parameter between $X_1^{(\lambda_1)}$ and $X_2^{(\lambda_2)}$ (Goto & Hamasaki, 2002 : Hamasaki & Goto, 2002). $A(\boldsymbol{K})$ is the probability proportional constant term and is given by,

$$A(\boldsymbol{K}) = \int_{a_2}^{b_2} \int_{a_1}^{b_1} \phi\,(x_1, x_2 : \rho)\, dx_1 dx_2, \tag{2.5}$$

in terms of the joint pdf of the bivariate standard normal distribution

$$\phi\,(x_1, x_2 : \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{ -\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho)^2} \right\}, \tag{2.6}$$

with the values of a_j and b_j given by in the following,

- $a_j = -k_j,\ b_j = +\infty$ when $\lambda_j > 0$
- $a_j = -\infty,\ b_j = +\infty$ when $\lambda_j = 0$
- $a_j = -\infty,\ b_j = -k_j$ when $\lambda_j < 0$

and the standardized truncation point k_j is given by

$$k_j = \frac{\lambda_j \mu_j + 1}{\lambda_j \sigma_j}, \quad j = 1, 2. \tag{2.7}$$

The power normal distribution fits a large variety of distributions, because it has the shape parameter. Goto *et al.* (1983) mentioned four considering points about the inclusive model.

1. The consistency of logic about statistical analyses process.
2. The flexibility of the model.
3. The ease of the model fitting evaluation.
4. The ease of computation.

Parameter Setting of BPN. In the previous paragraph, a bivariate power normal distribution, $BPN\ (\lambda_1, \lambda_2, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$ is introduced as a distribution of pre- and post-data. In this paragraph, we consider the reduction of location parameter from pre- to post-data ($\mu_1 > \mu_2$) with same shape and scale parameters between pre- and post-data ($\lambda_1 = \lambda_2 = \lambda$, $\sigma_1 = \sigma_2 = \sigma$). This means $(X_1, X_2) \sim BPN(\lambda, \lambda, \mu_1, \mu_2, \sigma, \sigma, \rho)$.
However, it is difficult to set these location (μ_i) and scale (σ_i) parameters for simulations, since

it is difficult to interpret the values of these parameters which vary greatly depending on the value of λ_i. Therefore, we use the median of original scale ($\xi_{0.5}$) as location-related parameter and the $\tau = (\xi_{0.75} - \xi_{0.25})/\xi_{0.5}$ as scale-related parameter for easiness of parameter setting and interpretation (Maruo & Goto, 2012; Maruo, *et al.,* 2011). ξ_p is the $100p$ percentile of power normal distribution and is given by

$$\xi_p = \begin{cases} \{\lambda(\mu + \sigma z_{p^*}) + 1\}^{\frac{1}{\lambda}}, & \lambda \neq 0, \\ \exp(\mu + \sigma z_p), & \lambda = 0, \end{cases}$$

where z_p and z_p^* are the p and p^* percentile of standard normal distribution, and p^* is given by

$$p^* = \begin{cases} 1 - A(K)(1-p), & \lambda > 0, \\ A(K)p, & \lambda < 0. \end{cases}$$

Moreover, the change of location parameter between pre- and post-data defines from percent change from pre-data of original scale (R), and the relationship is defined as {the median of post-data original scale}$= \xi_{0.5} \times (100 - R)/100$ ($0 < R < 100$). In summary, the distributions are identified based on reparametrization method $\{\lambda, \xi_{0.5}, \tau, R, \rho\}$ instead of $\{\lambda, \mu_1, \mu_2, \sigma, \rho\}$. The detail of reparametrization method is shown in appendix (Maruo, *et al.*, 2011; Maruo & Goto, 2012).

Figure 2.1 shows the pdf of BPN with the parameters of reparametrization method. In the figure, shape parameters are from -1 to +1 by 1 ($\lambda = -1, 0, +1$), scale-like parameters are from 0.2 to 0.8 by 0.2 ($\tau = 0.2, 0.4, 0.6, 0.8$), median of pre-data is 100 ($\xi_{0.5} = 100$), percent change from pre-data is 0 ($R = 0$) and correlation parameter is $\rho = 0.8$. The distribution is positive skewed when λ is less than 1, and the scale becomes large when the value of τ increases.

Applications of BPN to clinical data. It is expected that BPN is applicable to distributions of various clinical data, because BPN includes the shape parameters (λ) and can set the various distributions including more skewed distributions. For example, Goto & Uesaka (1980) presented the $\hat{\lambda}$ of blood serum component of laboratory test. Maruo *et al.*(2008) applied the univariate power normal distribution to various laboratory test data and estimated the shape parameter $\hat{\lambda}$ with the range between -1 and 0.25 as shown in figure 2.1 and evaluated the loss of information when we assume the normal or log-normal distribution to laboratory data. Hamasaki & Goto (2002) applied the BPN to the clinical data in both diastolic blood pressure (DBP) and systolic blood pressure (SBP) of the clinical trial to evaluate the treatment effect of calcium blocker, and said that SBP would be more positive skewed distribution than log-normal distribution because of $\hat{\lambda} < 0$ and DBP would be normal distribution because of $\hat{\lambda} \approx 1$ in the data. Goto *et al.*(2007) applied the power normal distribution to partial epilepsy data and estimated $\hat{\lambda} \approx 0$,

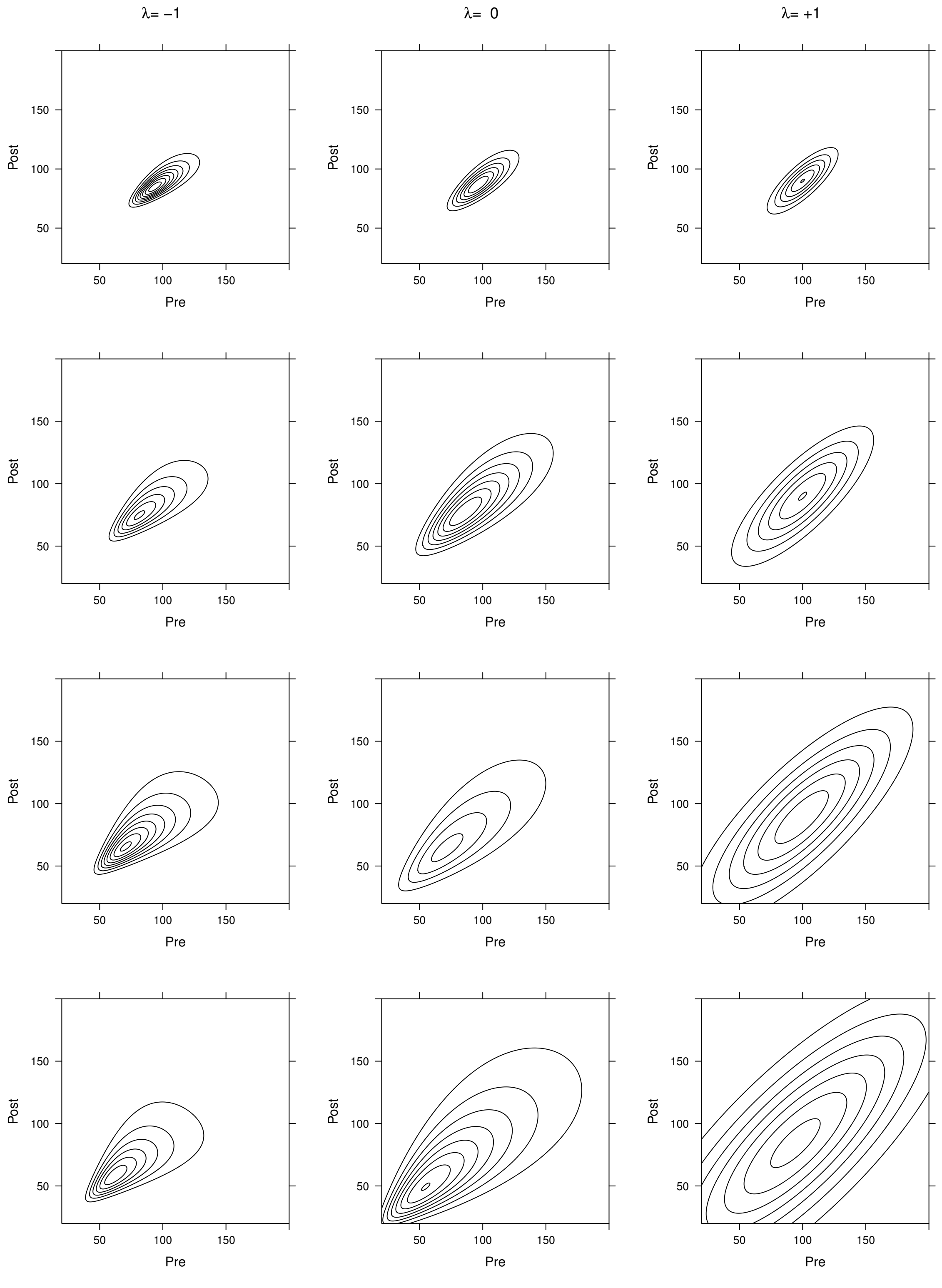


Figure 2.1: PDF of BPN with $\lambda = -1, 0, +1$ and $\tau = 0.2$(upper), $\tau = 0.4$, $\tau = 0.6$, $\tau = 0.8$(bottom))

which means that it is appropriate to analyze the data based on the log-normal distribution.

Table 2.1: $\hat{\lambda}$ of laboratory test: the modification of table 2 in Maruo *et al.*(2008)

Laboratory test	$\hat{\lambda}$	Laboratory test	$\hat{\lambda}$
ALP	0.25	TC	0.25
GOT	-1	TG	-0.25
GPT	-0.5	HDL-C	0
γ-GTP	-0.5		

Section 3

Statistical properties of ratio measures

3.1 Evaluation based on the bivariate normal distribution

Probability density function of PC and SPC. Pham-Gia *et al.* (2006) gave the exact closed form expression of the density of X_1/X_2 , where X_1 and X_2 are normal random variables, in terms of Hermite and confluent hypergeometric functions. In this section, we give the probability density function of PC and SPC based on Pham-Gia *et al.* (2006) .

Let X_1 and X_2 be the two random variables of bivariate normal distribution with parameters $\mathrm{BN}(\mu_1, \mu_2, \sigma, \sigma, \rho)$. Strictly speaking, suppose that truncated bivariate normal distribution, $\mathrm{TBN}(\mu_1, \mu_2, \sigma, \sigma, \rho)$, for X_1 and X_2, because we consider the data which is $X_1 \geq 0, X_2 \geq 0$. Then the distribution of PC is

$$h_{\mathrm{BN(PC)}}(v) = \frac{K_1}{2(1-\rho)(1+v)+v^2} H_{-2}(\xi_1(v)), \tag{3.1}$$

where $H_{-2}(\cdot)$ is the Hermite function,

$$H_{-2}(z) = \int_0^\infty t e^{-t^2-2tz} dt$$

and

$$\xi_1(v) = -\frac{(1-\rho)(\mu_1+\mu_2)+(\mu_2-\rho\mu_1)v}{\sigma\sqrt{2(1-\rho^2)\{2(1-\rho)(1+v)+v^2\}}},$$

$$K_1 = \frac{\sqrt{1-\rho^2}}{\pi\Phi_2(0,0;-\mu_1,-\mu_2,\sigma,\sigma,\rho)} \times \exp\left\{-\frac{\mu_1^2-2\rho\mu_1\mu_2+\mu_2^2}{2(1-\rho^2)\sigma^2}\right\}.$$

And the pdf of SPC is

$$h_{\mathrm{BN(SPC)}}(w) = \frac{K_2}{1-\rho+(1+\rho)w^2}H_{-2}(\xi_2(w)) \tag{3.2}$$

where $H_{-2}(\cdot)$ is also the Hermite function as well as the case of $h_{\mathrm{BN(PC)}}(v)$ and

$$\xi_2(w) = -\frac{1}{2\sigma\sqrt{(1-\rho^2)\{1-\rho+(1+\rho)w^2\}}} \times \{(1-\rho)(\mu_1+\mu_2)+(1+\rho)(\mu_2-\mu_1)w\},$$

$$K_2 = \frac{\sqrt{1-\rho^2}}{\pi\Phi_2(0,0;-\mu_1,-\mu_2,\sigma,\sigma,\rho)} \times \exp\left\{-\frac{\mu_1^2-2\rho\mu_1\mu_2+\mu_2^2}{2(1-\rho^2)\sigma^2}\right\},$$

where $\Phi_2(\cdot)$ is the cumulative distribution function of standard normal distribution.

Consider the situation with small coefficients of variation which are (σ/μ_1 or σ/μ_2) to ignore the affect of truncation. Then, $\Phi_2(0,0;-\mu_1,-\mu_2,\sigma,\sigma,\rho)$ is approximated by 1, and we can assume the situation that $X_1, X_2 \geq 0$. Figure 3.1 shows the pdf of PC and SPC with the parameters that $\mu_2-\mu_1 = 0, -0.3, -1$,$\sigma = 5, 10$ and $\rho = 0, 0.4, 0.8$ to figure out the shapes of the pdf. The pdf of SPC is symmetrical comared to the pdf of PC based on this figure.

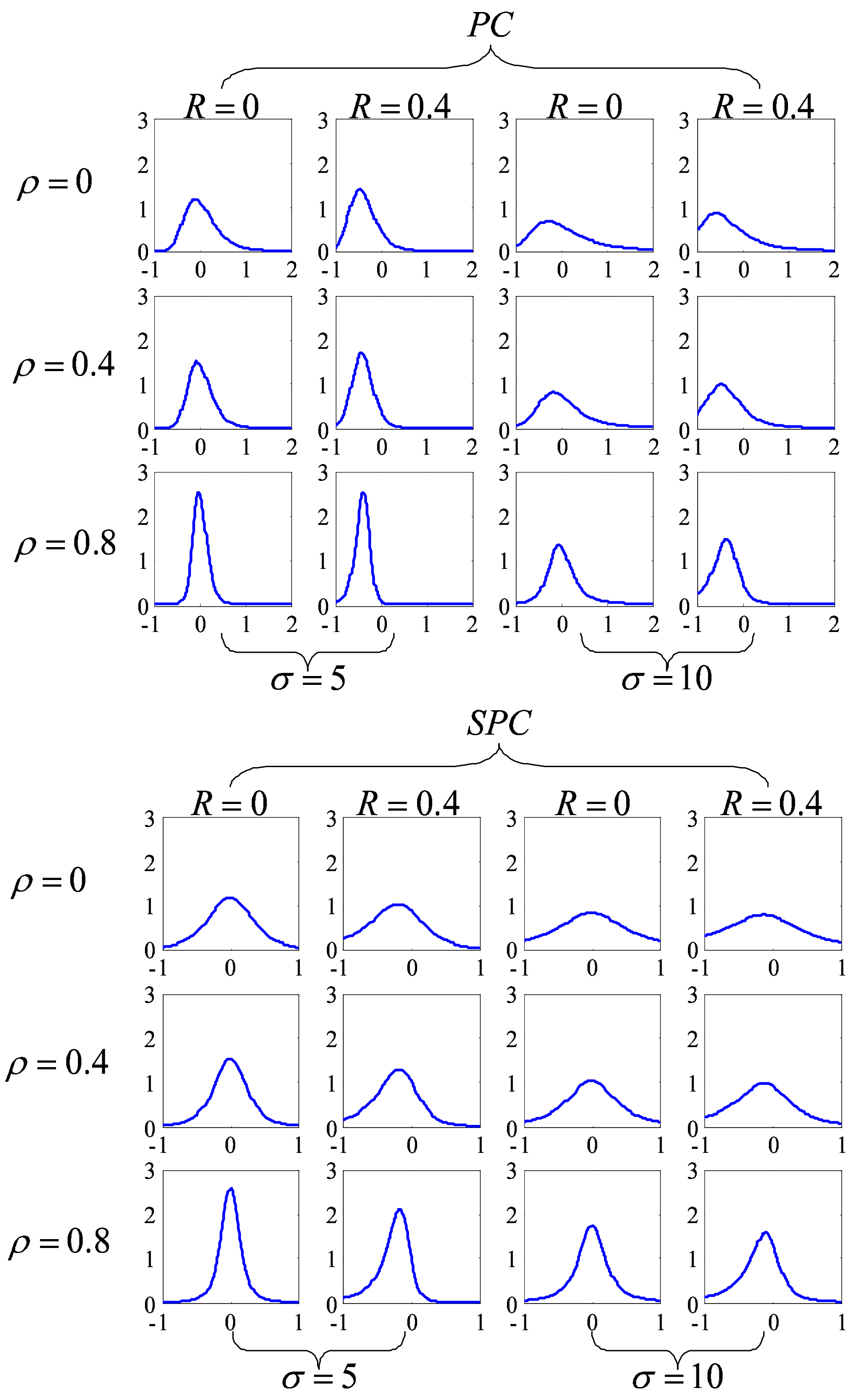


Figure 3.1: The pdf of PC and SPC based on bivariate normal distribution

Skewness of PC and SPC. Figure 3.2 shows the skewness of PC or SPC calculated based on each pdf. We assume that two random variables, X_1, X_2, are based on the bivariate normal distribution with the parameters $\mathrm{BN}(\mu_1, \mu_2, \sigma^2, \sigma^2, \rho)$ which are set within the range of $\mu_1 = 10$, $\mu_2 = 9$, $\sigma = 1$ and $\rho = 0 \sim 0.9$. The PC do not skew so much and so the difference of skewness between PC and SPC became small.

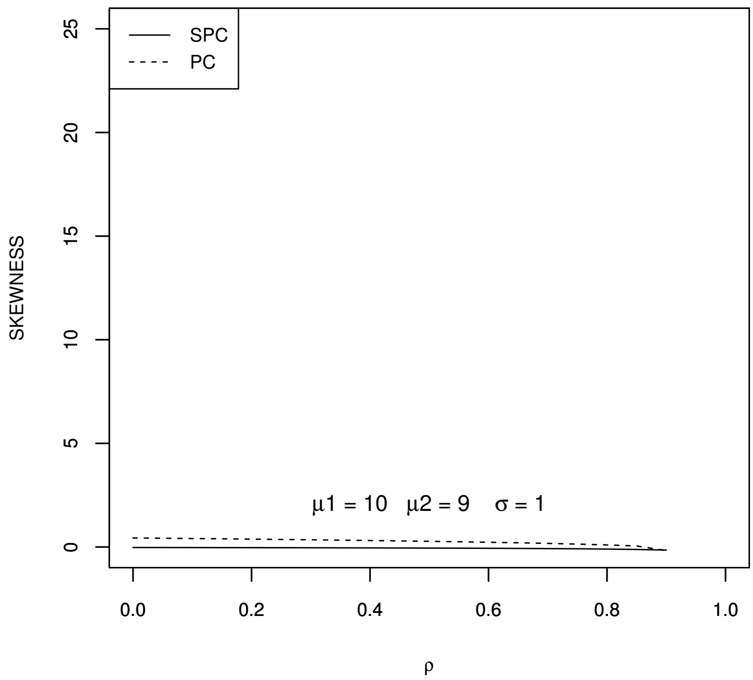


Figure 3.2: The relationship between ρ and skewness of PC and SPC based on bivariate normal distribution

3.2 Evaluation based on the bivariate log-normal distribution

Probability density function of PC and SPC. Let X_1 and X_2 be two random variables of bivariate log-normal distribution with parameters $\mathrm{BLN}(\mu_1, \mu_2, \sigma, \sigma, \rho)$. Then, we define the pdfs of PC is $h_{\mathrm{BLN(PC)}}(v)$ and the pdfs of SPC is $h_{\mathrm{BLN(SPC)}}(w)$, and these pdfs are

$$
\begin{aligned}
h_{\mathrm{BLN(PC)}}(v) &= \frac{1}{2\sigma\,(1+v)\,\sqrt{\pi\,(1-\rho)}} \\
&\quad \times \exp\left[-\frac{1}{4\sigma^2\,(1-\rho)}\left\{\log\,(1+v)-(\mu_2-\mu_1)\right\}^2\right], \qquad (3.3)
\end{aligned}
$$

$$
\begin{aligned}
h_{\mathrm{BLN(SPC)}}(w) &= \frac{1}{\sigma(1-w^2)\sqrt{\pi(1-\rho)}} \\
&\quad \times \exp\left[-\frac{1}{4\sigma^2(1-\rho)}\left\{\log\left(\frac{1-w}{1+w}\right)+\mu_2-\mu_1\right\}^2\right], \qquad (3.4)
\end{aligned}
$$

where μ_1 and μ_2 are the mean of log-transformed two variables (X_1 and X_2), σ^2 is the variance and ρ is the correlation.

Figure 3.3 shows the pdf of SPC and PC with the parameters that $R = 0, 0.4$, $\sigma = 0.5, 1$, $\rho = 0, 0.4, 0.8$ to figure out the shapes of the pdf. The R is the median of percent change about post-data which is calculated by the $\exp(\mu_2) = (1 - R)\exp(\mu_1)$. $R = 0$ means that the median of pre-data is same as the median of post-data, and $R = 0.4$ means the median of post-data had the 40 % reduction from pre-data. In these figures, all PC shows the positive skew distribution. On the other hand, SPC shows the symmetrized distributions.

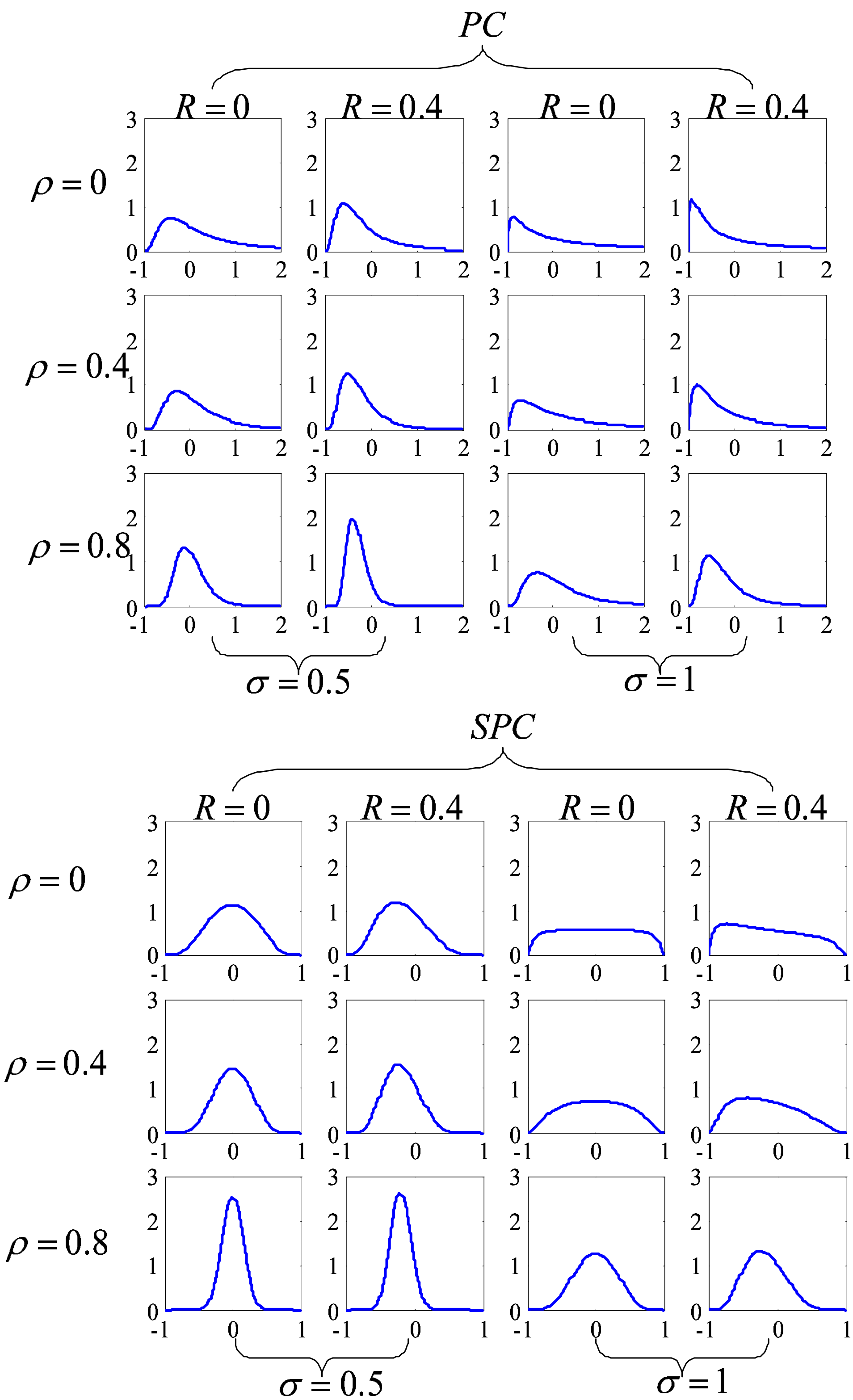


Figure 3.3: The pdf of PC and SPC based on bivariate log-normal distribution

Skewness of PC and SPC. In this paragraph, we evaluated the skewness of PC and SPC calculated by numerical integration method based on the pdfs, quantitatively. Figure 3.4 shows the relationship between correlation and skewness of PC and SPC. The parameter combinations used for the skewness calculation in this figure are that $\mu_1 = 1$, $\mu_2 = 0.9$, $\sigma = 1$ and $\rho = 0 \sim 0.9$. The skewness of SPC is smaller than PC without regard to correlation.

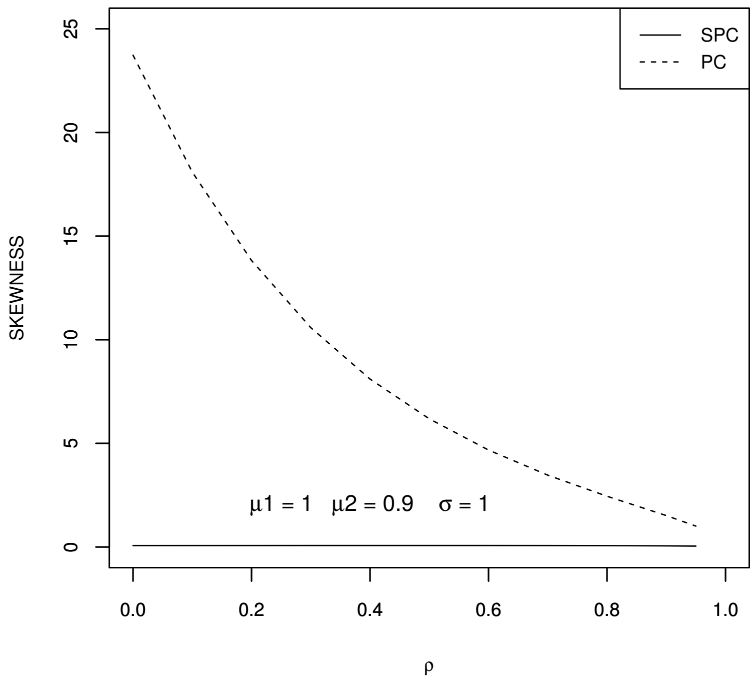


Figure 3.4: The relationship between ρ and skewness of PC and SPC based on bivariate log-normal distribution

3.3 Evaluation based on the bivariate power-normal distribution

Probability density function of PC and SPC. Let X_1 and X_2 be the two positive random variables of bivariate power normal distribution with parameters $\mathrm{BPN}(\lambda, \lambda, \mu_1, \mu_2, \sigma, \sigma, \rho)$. The pdfs of PC or SPC are calculated by using the variable transformation method. The two random variables, X_1 and X_2, of BPND in section 2 are transformed to $X_1 = U$ and $X_2 = U(1 + V)$ for PC, which is equal to $U = X_1$ and $V = (X_2 - X_1)/X_1$. On the other hand, the two variables are also transformed to $X_1 = U$ and $X_2 = U \times (1 + W)/(1 - W)$ for SPC, which is equal to

$U = X_1$ and $W = (X_2 - X_1)/(X_1 + X_2)$. Then, the pdfs of PC as $h_{\mathrm{BPN(PC)}}(v)$ is given by,

$$\begin{aligned} h_{\mathrm{BPN(PC)}}(v) &= \int u \times \frac{u^{2\lambda-2}(1+v)^{\lambda-1}}{2\pi\sigma^2\sqrt{1-\rho^2}A(\boldsymbol{K})} \\ &\quad \times \exp\left[-\frac{1}{2}\left(\mathbf{M}^{(\lambda)} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{M}^{(\lambda)} - \boldsymbol{\mu}\right)\right] du \end{aligned} \tag{3.5}$$

where $A(\boldsymbol{K})$ is the probability proportional constant term shown in section 2 and

$$\begin{aligned} \mathbf{M}^{(\lambda)} &= \left(u^{(\lambda)},\ \ \{u\,(1+v)\}^{(\lambda)}\right), \\ \boldsymbol{\mu} &= (\mu_1, \mu_2)\,, \\ \boldsymbol{\Sigma} &= \begin{pmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{pmatrix}. \end{aligned}$$

And the components of $\mathbf{M}^{(\lambda)}$ are

$$u^{(\lambda)} = \begin{cases} \dfrac{u^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log u & \lambda = 0 \end{cases}$$

and

$$\{u\,(1+v)\}^{(\lambda)} = \begin{cases} \dfrac{u^\lambda\,(1+v)^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log\{u(1+v)\} & \lambda = 0. \end{cases}$$

Next, the pdf of SPC as $h_{\mathrm{BPN(SPC)}}(w)$ is given by

$$\begin{aligned} h_{\mathrm{BPN(SPC)}}(w) &= \int \frac{u}{(1-w)^2} \times \frac{u^{\lambda-1}\left\{\dfrac{1+w}{1-w} \times u\right\}^{\lambda-1}}{\pi\sigma^2\sqrt{1-\rho^2}A(\boldsymbol{K})} \\ &\quad \times \exp\left[-\frac{1}{2}\left(\mathbf{N}^{(\lambda)} - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{N}^{(\lambda)} - \boldsymbol{\mu}\right)\right] du \end{aligned} \tag{3.6}$$

where

$$\mathbf{N}^{(\lambda)} = \left(u^{(\lambda)},\ \ \left(\frac{1+v}{1-v} \times u\right)^{(\lambda)}\right)$$

and $A(\boldsymbol{K})$ is also the probability proportional constant term shown in section 2, $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is the same as ones of PC. The components of $\mathbf{N}^{(\lambda)}$ are

$$u^{(\lambda)} = \begin{cases} \dfrac{u^{\lambda}-1}{\lambda} & \lambda \neq 0 \\ \log u & \lambda = 0 \end{cases}$$

and

$$\left(\frac{1+v}{1-v}\times u\right)^{(\lambda)} = \begin{cases} \dfrac{\left(\dfrac{1+v}{1-v}\times u\right)^{\lambda}-1}{\lambda} & \lambda \neq 0 \\ \log\left(\dfrac{1+v}{1-v}\times u\right) & \lambda = 0. \end{cases}$$

3.3.1 Definition of the distribution of skewness

When the shape parameter of power normal distribution is equal to or larger than 0 generally, it is possible to calculate any moment and skewness. However, skewness cannot be calculated in $-3 \leq \lambda < 0$, because three order moment does not exist(Goto *et al.*, 1983). In this section, we define the alternative criterion about skewness, which is

$$\eta = \frac{\xi_{0.975} - \xi_{0.5}}{\xi_{0.5} - \xi_{0.025}}.$$

When the distributions become more symmetrical, η will become nearer one. And, when η is larger than one, the distributions become more positively skewed. On the other hand, the distributions become negative skew, when η is less than 0.

3.3.2 Detection of factors to affect the skewness of distributions

Factors to affect the distribution of skewness of PC or SPC were investigated based on ANOVA. In this analysis, η of PC or SPC is set as response value, shape(λ), scale-like(τ), percent change from pre-data (R) and correlation (ρ) were included as factor. In addition, the interactions between two factors of four parameters were also included in the model. In the clinical data, such as epilepsy (Goto *et al.*, 2007) and laboratory data (Uesaka & Goto:1980, Maruo *et al.*: 2007) are based on the positive skew distribution in many cases, therefore shape parameters are set from -1 to 1 by 0.5 ($\lambda = -1, -0.5, 0, 0.5, +1$). Percent change from pre-data are set from 10 to 40 by 10($R = 10, 20, 30, 40$). Scale-like parameter are set from 0.2 to 1.0 by 0.2 ($\tau = 0.2, 0.4, 0.6, 0.8, 1.0$). Correlation parameter are set from 0.2 to 0.8 by 0.2 ($\rho = 0.2, 0.4, 0.6, 0.8$). The η of PC or SPC were calculated based on the numerical integral for all combinations of

Table 3.1: Sum of square, F value and contribution rate of PC or SPC based on the ANOVA

	The η of PC is response.			The η of SPC is response.		
Factor	Sum of square	F value	Cont. Rate (%)	Sum of square	F value	Cont. Rate (%)
λ	52.573	273.359	9.94	0.316	258.289	38.97
τ	315.690	1641.466	59.70	0.091	74.131	11.18
ρ	106.838	555.514	20.20	0.238	193.977	29.27
R	9.605	49.941	1.82	0.013	10.693	1.61
$\lambda \times \tau$	17.868	92.904	3.38	0.021	16.852	2.54
$\lambda \times \rho$	3.266	16.979	0.62	0.009	7.729	1.17
$\lambda \times R$	1.993	10.365	0.38	0.059	48.415	7.30
$\tau \times \rho$	19.814	103.026	3.75	0.007	5.660	0.85
$\tau \times R$	1.126	5.854	0.21	0.015	12.637	1.91
$\rho \times R$	0.006	0.031	near 0	0.042	34.439	5.20

these four parameter, which were 400 combination cases (= 5 levels of λ $\times$ 4 levels of R $\times$ 5 levels of τ $\times$ 4 levels of ρ) as total.

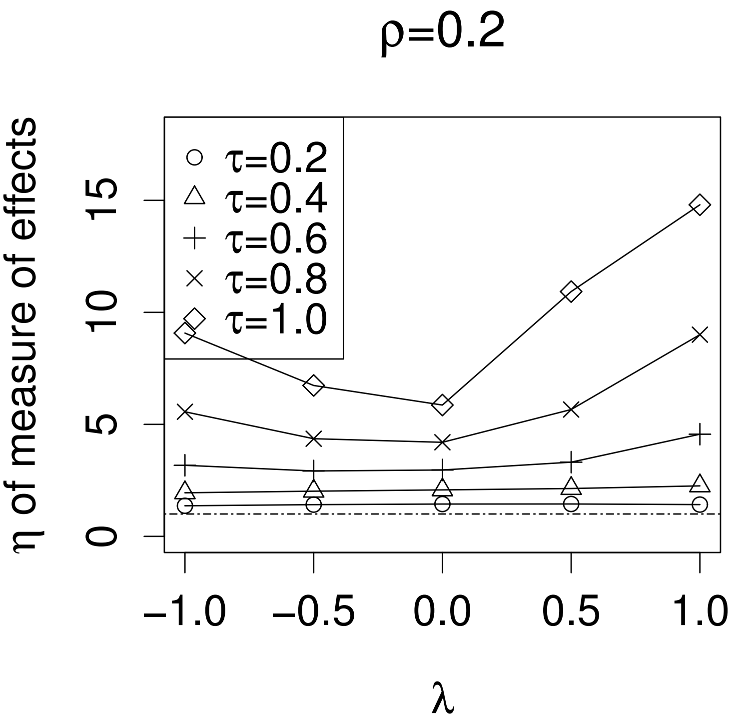
Table 3.1 shows the sum of squares, F value and contribution rate which is defined as the sum of squares in each factor is divided by the sum of squares in total factors $\times$ 100. For PC, the τ was the largest contribution to distribution of skewness (59.70 %). The second largest contribution was ρ (20.20 %), and the third was λ (9.94 %). However, the R did not have a high contribution to the distribution of skewness (1.82 %). For interactions between two factors of PC, $\lambda \times \tau$ (3.38 %) and $\tau \times \rho$ (3.75 %) had more contribution than others. On the other hand, the contribution rate of τ, ρ and λ for SPC which had high contribution for PC were 11.18 %, 29.27 % and 38.97 % respectively. These three factors of SPC were also contributed highly as well as PC. The contribution of R for SPC was also low (1.61 %). For interactions, SPC had a different trend to PC, and the contribution of $\lambda \times R$ (7.30 %) and $\rho \times R$ (5.20 %) was high. However, the sum of squares of SPC in each factor was much smaller than PC, and it was shown that each factor of SPC was contribution less to skewness of distributions.

3.3.3 Graphical evaluation of skewness of distribution

In this section, we evaluated the effect of three parameters which had a high contribution to skewness of PC, graphically. The three parameters were λ, τ and ρ, and the R which had less contribution to the skewness of PC were fixed as 10 %. Figure 3.5 to 3.8 shows the relationship between the skewness of the distribution (η) of PC or SPC and the three parameters of BPN (λ, τ and ρ). The λ set 5 levels which are -1, -0.5, 0, +0.5 and +1. The τ set 4 levels which were 0.2, 0.4, 0.6 and 0.8. The ρ set 4 levels which were 0.2, 0.4, 0.6 and 0.8.

For PC, the η increased with the absolute value of λ ($\lambda = 1$ or -1) and this trend became remarkable especially when τ was equal to or more than 0.6. And the η increased with τ increasing or decreases with ρ increasing. When $\lambda < 0$, the distribution of pre- and post-data became more positively skewed than log-normal distribution, and might have the case that post-value (X_1) was much larger than post-value (X_2). Then, the η of PC was larger than one and the η increased with increasing τ. When $\lambda > 0$, the distribution of pre- and post-data became more negative skew and occured the value near 0. Especially, when τ was large, the truncation in the left side occured ($A(\mathbf{K}) < 1$) and a lot of values near $X_1 = 0$ generated. Then, the η of PC increased. For ρ, the η of PC increased with decreasing ρ, because the difference between X_1 and X_2 became large.

For SPC, the skewness of the distribution (η) was almost one in all conditions, and this means that all distributions of SPC show almost all symmetry in all combinations of the BPN parameters.



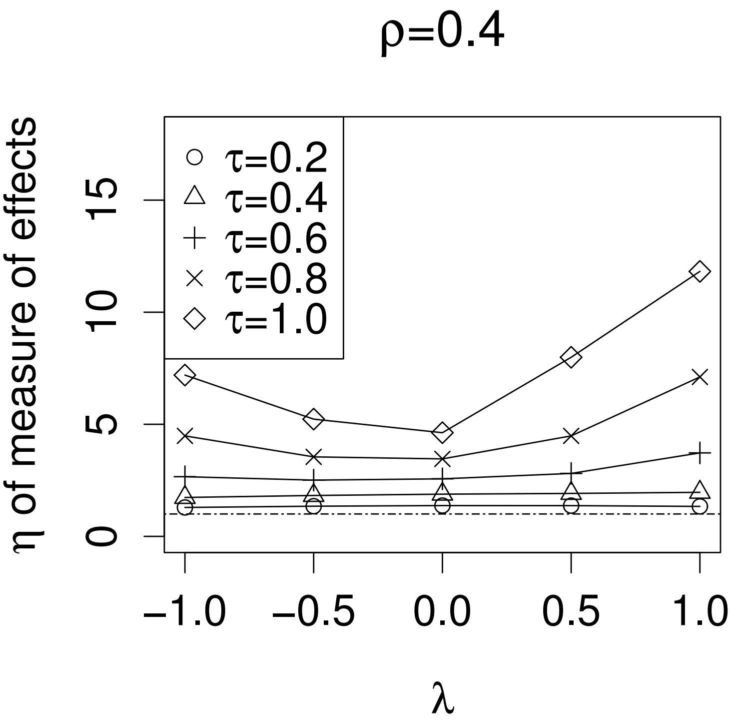


Figure 3.5: Relationship between η of PC and λ of BPN ($\rho = 0.2, 0.4$ and $R = 10$ %)

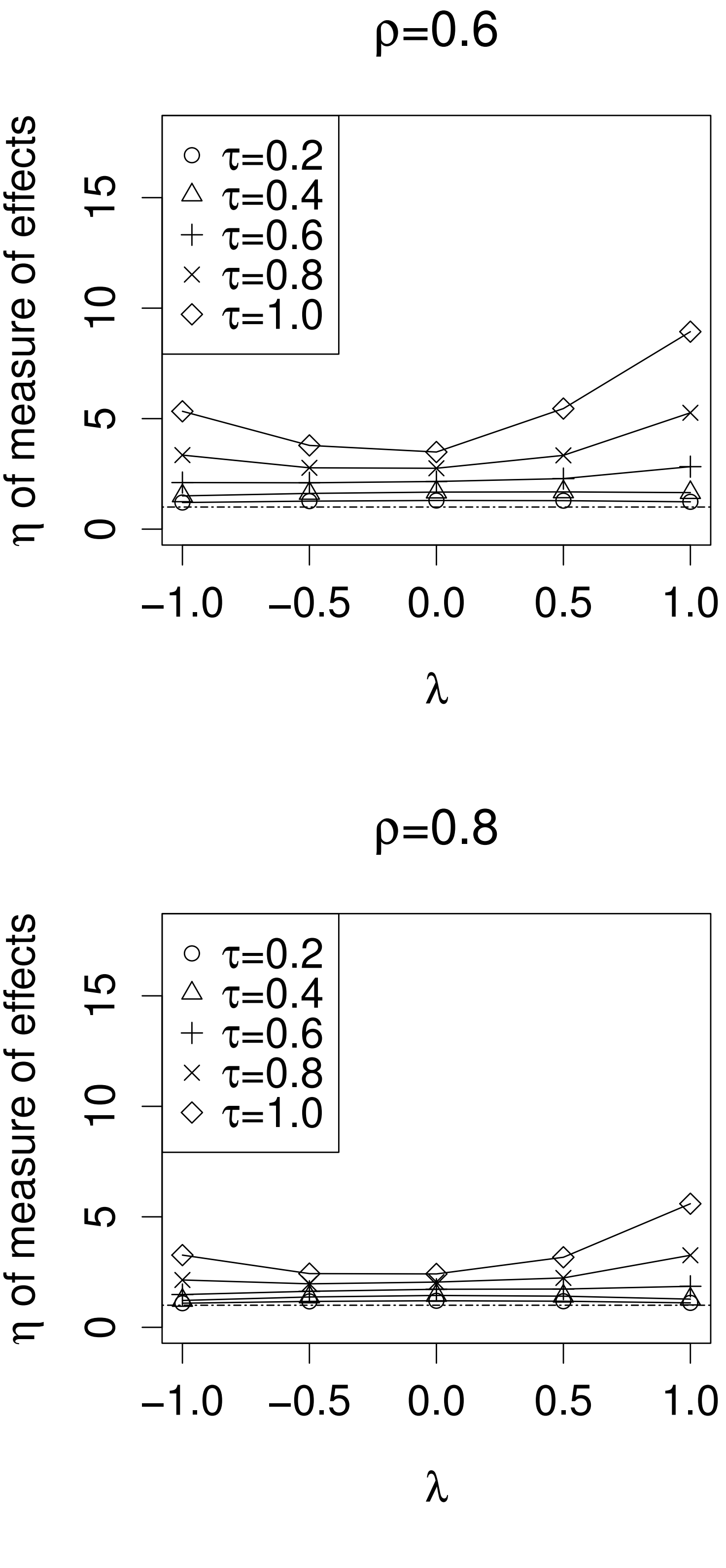


Figure 3.6: Relationship between η of PC and λ of BPN ($\rho = 0.6, 0.8$ and $R = 10$ %)

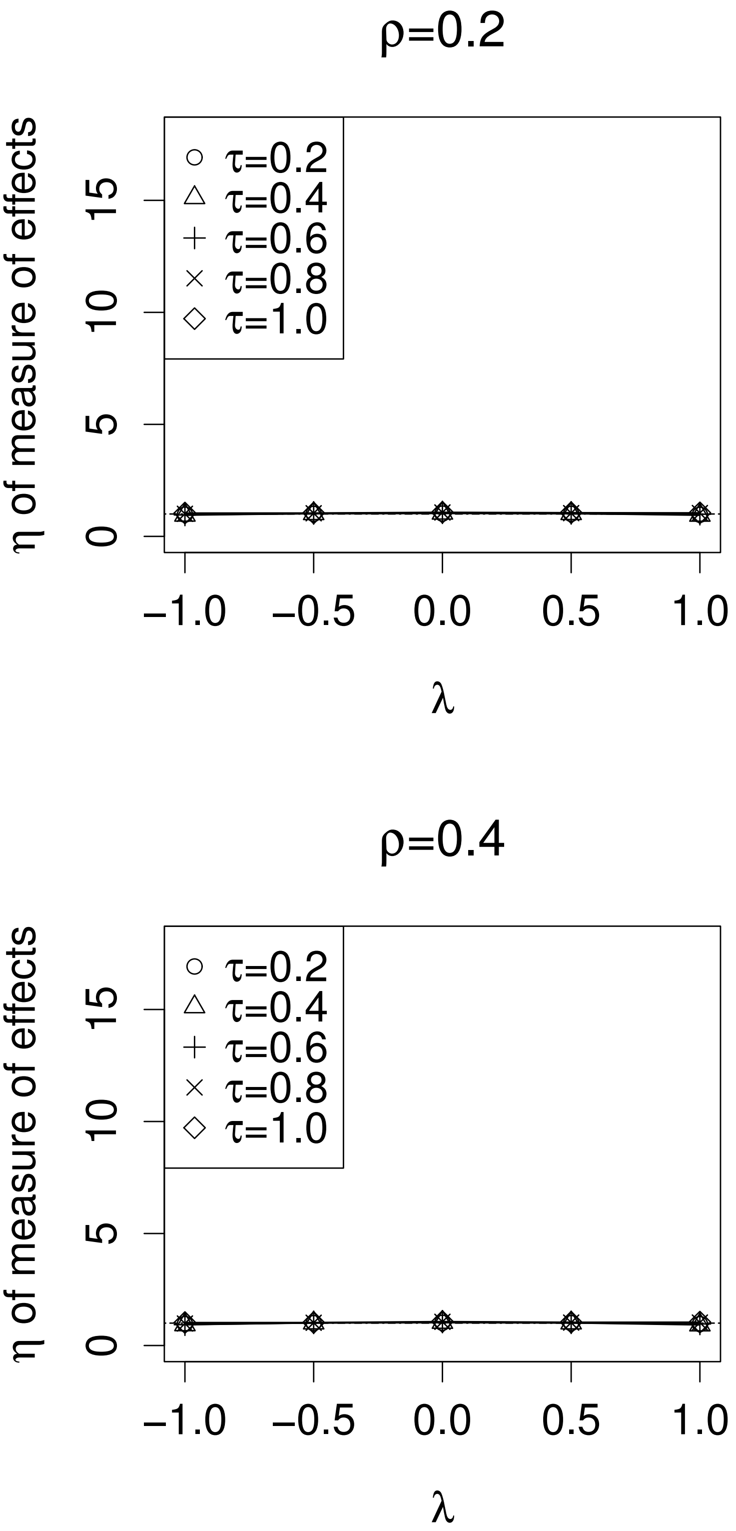
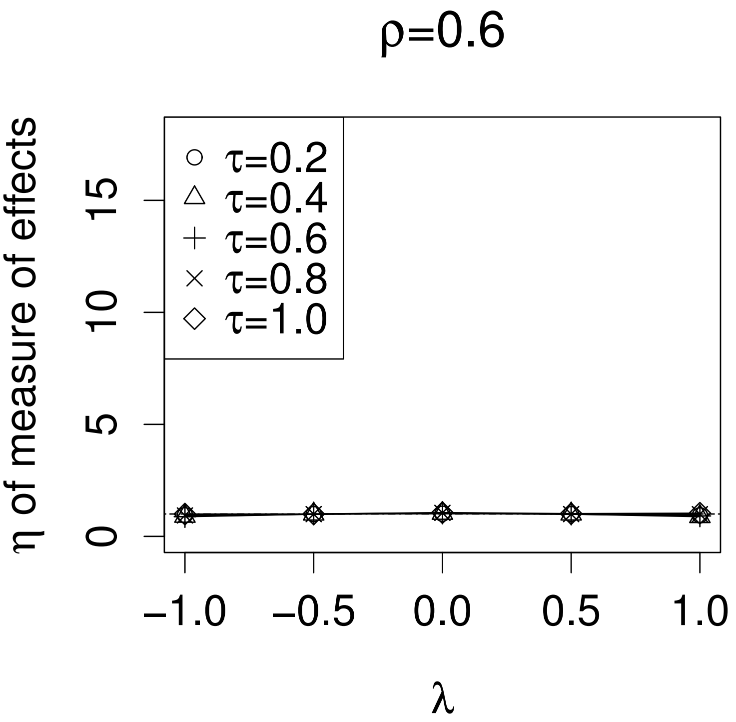


Figure 3.7: Relationship between η of SPC and λ of BPN ($\rho = 0.2, 0.4$ and $R = 10$ %)



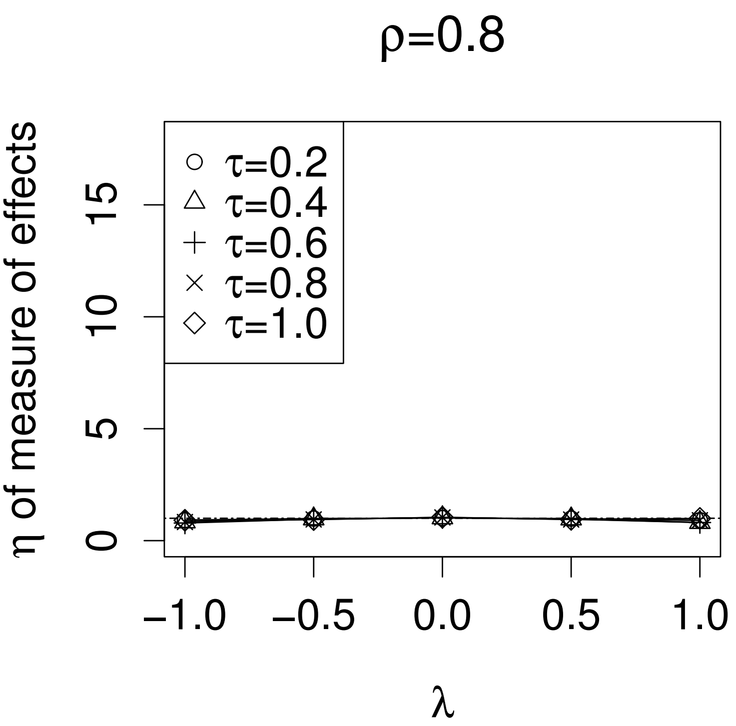


Figure 3.8: Relationship between η of SPC and λ of BPN ($\rho = 0.6, 0.8$ and $R = 10$ %)

3.4 Relationship between symmetrized percent change and coefficient of variation

On the other hand, we considered an another measure which is defined as $(X_2 - X_1)/\frac{1}{2}(X_2 + X_1)$. The numerator of the measure is a difference of two data and the denominator is a mean. When we regard the difference of denominator as an index of variation, the measure may be considered as a variation srandarized by mean, such as a coefficient of variation. In fact, this measure is called Variability (%) in the bioanalytical field and is used for evaluating the level of reproducibility of assay results using incurred samples (Mario *et al.*, 2007 and Douglas *et al.*, 2009). The variability is used in a fixed error limit method and a model similar to the familiar 4-6-X QC criteria can be applied. For small molecules (non-ligand binding) two thirds of the repeat samples (X_2) should agree within 20 % and for ligand-binding assay, two thirds of the repeat samples should agree within 30 %. The variability (% difference) should be calculated using the mean of the original and repeat results as described by the following formula:

$$\text{Variability}(\%) = \frac{\text{Repeat}(X_2) - \text{Original}(X_1)}{\frac{1}{2}(X_1 + X_2)} \times 100$$

Graphical comparison of pdf. In this paragraph, we investigate the relationship between SPC (Variability) and CV of two samples. Figure 3.9 shows the relationship between SPC and CV. The CV of two samples is given by $|X_2 - X_1|/(X_1 + X_2)$ and the numerator of this formula is replaced by the difference of two samples with absolute value.

To figure out the distribution of SPC and CV graphically, we show the histogram of SPC and CV in figure 3.10 on condition that two samples follows the bivariate normal distribution (BN) and bivariate log-normal distribution (BLN) . The upper graph is based on the BN with parameters of $\text{BN}(10, 9, 1, 1, \rho)$ and the bottom is on the BLN with parameters of $\text{BLN}(10, 9, 1, 1, \rho)$. The ρ is from 0.2 to 0.8 by 0.2 in all graphs. We generate the 10,000 random samples with each parameters and create histograms. From these figures, the distribution of CV is the distributions folded back negative value of SPC to positive, because of the formula of absolute value of numerator.

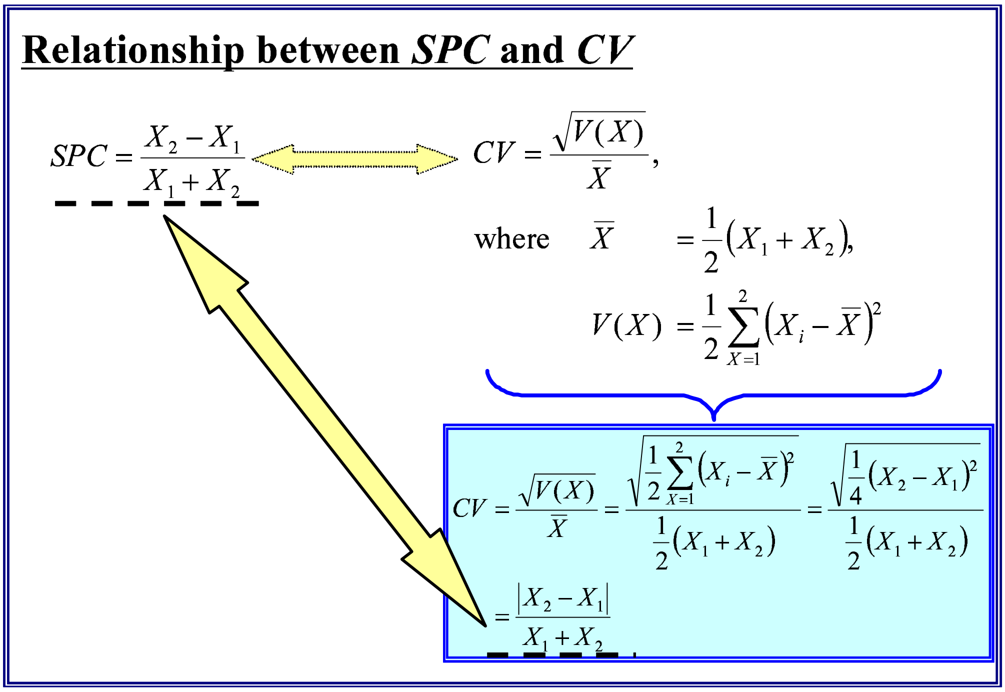


Figure 3.9: The relationship between SPC and CV of two samples

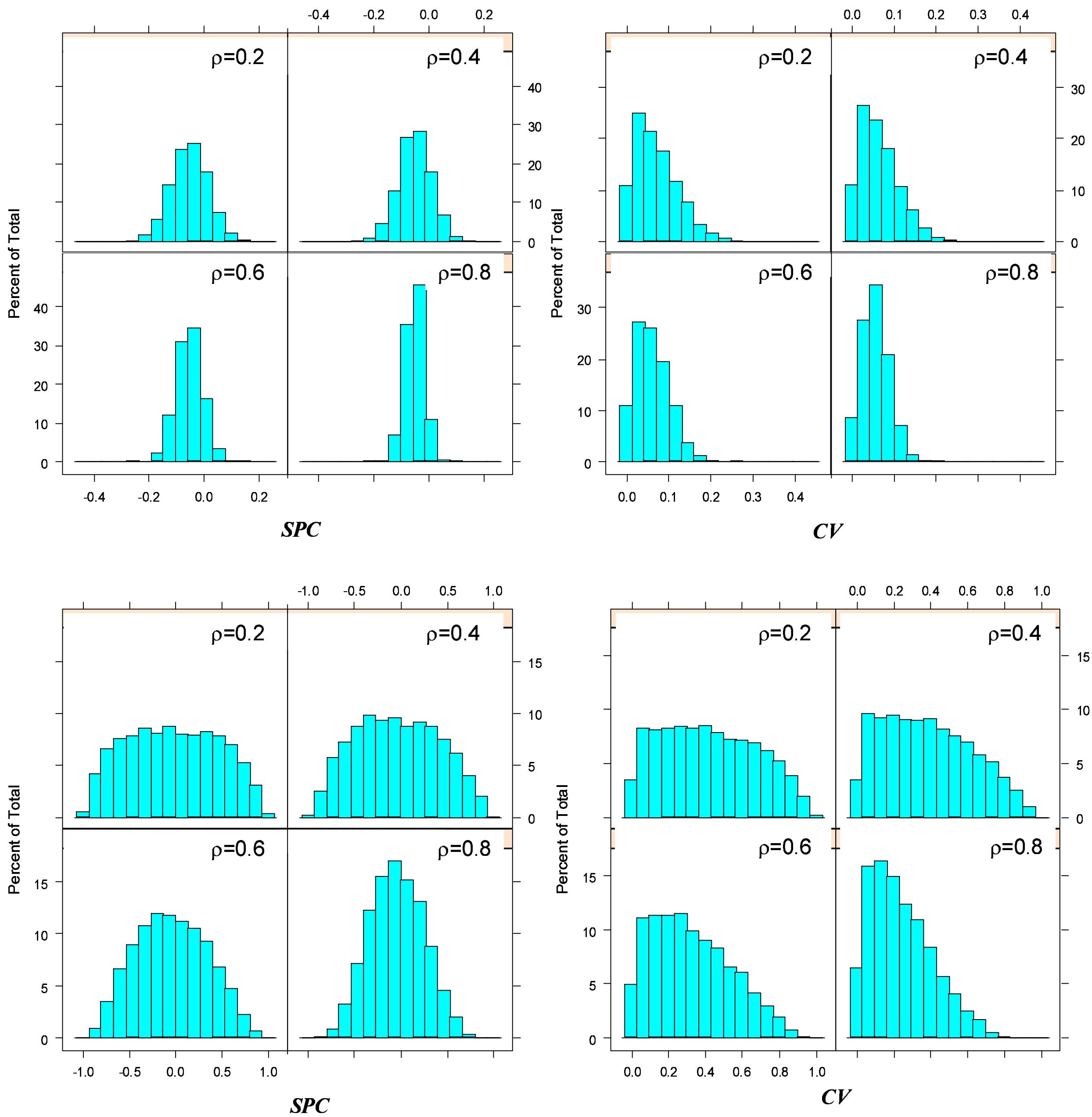


Figure 3.10: The distribution of SPC and CV (upper is BN and bottom is BLN)

Section 4

Simulations and case studies

In this section, we evaluate the effect of the statistical test results based on the simulation in case the distributions of PC or SPC do not follow the assumed distribution in each test, such as normal. In addition, we also show the case example to apply the power normal distribution to SPC.

4.1 Simulation 1: One sample comparison

4.1.1 Design of simulation 1

We consider the situation where treatment effect is to reduce the post-data from pre-data, which is $R > 0$, and then we investigate the power of one-sample test for PC or SPC. The objective of this simulation is to evaluate the relationship between the distribution of pre- and post data based on the BPN and the power of the one-sample test about PC or SPC.

Hypothesis of the statistical test. In this simulation, the following hypotheses with 0.05 of significance level are set for three measures PC, SPC, DTS. The DTS is called "Difference on Transformed Scale" and is defined as $X_2^{(\lambda)} - X_1^{(\lambda)}$. One-sample t-test is used for PC, SPC and DTS and Wilcoxon Signed Rank Test (WSRT) are also used for PC and SPC. The hypotheses of interest are based on one-side and are as follows,

$$\begin{aligned} &\mathrm{H}_0 : \theta = 0, \\ \text{vs} \quad &\mathrm{H}_1 : \theta < 0, \end{aligned}$$

where θ is the expected value or median of PC, SPC and DTS(expected value only) in one sample test. In addition, we evaluate the hypotheses based on the two sides,

$$\begin{aligned} &\mathrm{H}_0 : \theta = 0, \\ \text{vs} \quad &\mathrm{H}_1 : \theta \neq 0. \end{aligned}$$

Parameter setting. Table 4.1, 4.2 and 4.3 shows the parameters combination of BPN assumed as pre- and post-data distribution. We set the 5 levels λ $(= -1, -0.5, 0, 0.5, 1)$, 4 levels τ $(= 0.2, 0.4, 0.6, 0.8)$, 4 levels ρ $(= 0.2, 0.4, 0.6, 0.8)$ and R(=0, 10 %), and then calculate the μ_1, μ_2, σ_1 and σ_2 based on the reparametrization method (Maruo, *et al.*, 2011; Maruo & Goto, 2012). Sample size is calculated based on the DTS(t-test), because DTS is normal in many cases and has highest power. Minimum sample size to exceed the power of 0.8 for DTS(t-test) sets in both hypotheses.

We calculate the proportion of significance per total numbers of simulations about DTS, PC and SPC, when one-sample t-test or Wilcoxon Signed Rank Test (PC and SPC only) are applied. Total numbers of simulation is 100,000 times. The (tentative) type I error rate is defined as the proportion of significance when $R = 0$, and the (tentative) power is defined as the proportion of significance when $R = 10$. We can evaluate the loss of information based on the difference from the power of PC or SPC to 0.8, because sample size sets near the value of 0.8 for power of DTS.

Table 4.1: The combination of parameters for simulation and sample size ($\lambda = -1$ and -0.5)

λ	med1	R	τ	ρ	μ_1	μ_2	σ_1	σ_2	n(one side)	n(two sides)
-1	100	10	0.2	0.2	0.990	0.989	0.00147	0.00147	18	44
-1	100	10	0.2	0.4	0.990	0.989	0.00147	0.00147	13	34
-1	100	10	0.2	0.6	0.990	0.989	0.00147	0.00147	9	24
-1	100	10	0.2	0.8	0.990	0.989	0.00147	0.00147	5	12
-1	100	10	0.4	0.2	0.990	0.989	0.00286	0.00286	66	166
-1	100	10	0.4	0.4	0.990	0.989	0.00286	0.00286	49	126
-1	100	10	0.4	0.6	0.990	0.989	0.00286	0.00286	33	84
-1	100	10	0.4	0.8	0.990	0.989	0.00286	0.00286	17	42
-1	100	10	0.6	0.2	0.990	0.989	0.00415	0.00415	143	368
-1	100	10	0.6	0.4	0.990	0.989	0.00415	0.00415	106	270
-1	100	10	0.6	0.6	0.990	0.989	0.00415	0.00415	71	180
-1	100	10	0.6	0.8	0.990	0.989	0.00415	0.00415	36	90
-1	100	10	0.8	0.2	0.990	0.989	0.00549	0.00549	288	726
-1	100	10	0.8	0.4	0.990	0.989	0.00549	0.00549	208	534
-1	100	10	0.8	0.6	0.990	0.989	0.00549	0.00549	134	340
-1	100	10	0.8	0.8	0.990	0.989	0.00549	0.00549	65	166
-0.5	100	10	0.2	0.2	1.80	1.79	0.0148	0.0148	19	48
-0.5	100	10	0.2	0.4	1.80	1.79	0.0148	0.0148	14	36
-0.5	100	10	0.2	0.6	1.80	1.79	0.0148	0.0148	10	24
-0.5	100	10	0.2	0.8	1.80	1.79	0.0148	0.0148	5	12
-0.5	100	10	0.4	0.2	1.80	1.79	0.0291	0.0291	72	182
-0.5	100	10	0.4	0.4	1.80	1.79	0.0291	0.0291	53	136
-0.5	100	10	0.4	0.6	1.80	1.79	0.0291	0.0291	36	92
-0.5	100	10	0.4	0.8	1.80	1.79	0.0291	0.0291	18	46
-0.5	100	10	0.6	0.2	1.80	1.79	0.0427	0.0427	157	396
-0.5	100	10	0.6	0.4	1.80	1.79	0.0427	0.0427	115	290
-0.5	100	10	0.6	0.6	1.80	1.79	0.0427	0.0427	77	194
-0.5	100	10	0.6	0.8	1.80	1.79	0.0427	0.0427	39	100
-0.5	100	10	0.8	0.2	1.80	1.79	0.0553	0.0553	260	658
-0.5	100	10	0.8	0.4	1.80	1.79	0.0553	0.0553	192	492
-0.5	100	10	0.8	0.6	1.80	1.79	0.0553	0.0553	130	328
-0.5	100	10	0.8	0.8	1.80	1.79	0.0553	0.0553	66	166

Table 4.2: The combination of parameters for simulation and sample size ($\lambda = 0$ and $+0.5$)

λ	med1	R	τ	ρ	μ_1	μ_2	σ_1	σ_2	n(one side)	n(two sides)
0	100	10	0.2	0.2	4.61	4.50	0.148	0.148	20	50
0	100	10	0.2	0.4	4.61	4.50	0.148	0.148	15	38
0	100	10	0.2	0.6	4.61	4.50	0.148	0.148	10	26
0	100	10	0.2	0.8	4.61	4.50	0.148	0.148	5	14
0	100	10	0.4	0.2	4.61	4.50	0.295	0.295	78	198
0	100	10	0.4	0.4	4.61	4.50	0.295	0.295	59	148
0	100	10	0.4	0.6	4.61	4.50	0.295	0.295	39	100
0	100	10	0.4	0.8	4.61	4.50	0.295	0.295	20	50
0	100	10	0.6	0.2	4.61	4.50	0.438	0.438	172	436
0	100	10	0.6	0.4	4.61	4.50	0.438	0.438	129	328
0	100	10	0.6	0.6	4.61	4.50	0.438	0.438	86	218
0	100	10	0.6	0.8	4.61	4.50	0.438	0.438	43	110
0	100	10	0.8	0.2	4.61	4.50	0.578	0.578	299	758
0	100	10	0.8	0.4	4.61	4.50	0.578	0.578	224	570
0	100	10	0.8	0.6	4.61	4.50	0.578	0.578	150	380
0	100	10	0.8	0.8	4.61	4.50	0.578	0.578	75	190
0.5	100	10	0.2	0.2	18.0	17.0	1.48	1.48	21	54
0.5	100	10	0.2	0.4	18.0	17.0	1.48	1.48	16	40
0.5	100	10	0.2	0.6	18.0	17.0	1.48	1.48	11	28
0.5	100	10	0.2	0.8	18.0	17.0	1.48	1.48	6	14
0.5	100	10	0.4	0.2	18.0	17.0	2.97	2.97	84	208
0.5	100	10	0.4	0.4	18.0	17.0	2.97	2.97	63	160
0.5	100	10	0.4	0.6	18.0	17.0	2.97	2.97	42	106
0.5	100	10	0.4	0.8	18.0	17.0	2.97	2.97	21	54
0.5	100	10	0.6	0.2	18.0	17.0	4.45	4.45	185	478
0.5	100	10	0.6	0.4	18.0	17.0	4.45	4.45	141	352
0.5	100	10	0.6	0.6	18.0	17.0	4.45	4.45	94	240
0.5	100	10	0.6	0.8	18.0	17.0	4.45	4.45	47	118
0.5	100	10	0.8	0.2	18.0	17.0	5.93	5.93	337	860
0.5	100	10	0.8	0.4	18.0	17.0	5.93	5.93	253	628
0.5	100	10	0.8	0.6	18.0	17.0	5.93	5.93	167	424
0.5	100	10	0.8	0.8	18.0	17.0	5.93	5.93	83	210

Table 4.3: The combination of parameters for simulation and sample size ($\lambda = +1$)

λ	med1	R	τ	ρ	μ_1	μ_2	σ_1	σ_2	n(one side)	n(two sides)
1	100	10	0.2	0.2	99.0	89.0	14.8	14.8	22	56
1	100	10	0.2	0.4	99.0	89.0	14.8	14.8	17	42
1	100	10	0.2	0.6	99.0	89.0	14.8	14.8	11	28
1	100	10	0.2	0.8	99.0	89.0	14.8	14.8	6	14
1	100	10	0.4	0.2	99.0	89.0	29.7	29.7	89	226
1	100	10	0.4	0.4	99.0	89.0	29.7	29.7	66	166
1	100	10	0.4	0.6	99.0	89.0	29.7	29.7	45	112
1	100	10	0.4	0.8	99.0	89.0	29.7	29.7	22	56
1	100	10	0.6	0.2	98.2	88.2	45.2	45.2	232	590
1	100	10	0.6	0.4	98.2	88.2	45.2	45.2	169	436
1	100	10	0.6	0.6	98.2	88.2	45.2	45.2	112	280
1	100	10	0.6	0.8	98.2	88.2	45.2	45.2	54	138
1	100	10	0.8	0.2	93.3	83.3	64.2	64.2	628	1560
1	100	10	0.8	0.4	93.3	83.3	64.2	64.2	440	1100
1	100	10	0.8	0.6	93.3	83.3	64.2	64.2	271	696
1	100	10	0.8	0.8	93.3	83.3	64.2	64.2	128	322

4.1.2 Results of simulation 1

Figure 4.1 shows the (tentative) type I error rate with one side in each parameter combination. Tentative type I error rate of three measures were nearly equal to or less than significance level (0.05) in all parameter combinations. Especially, the (tentative) type I error rate of PC(t-test) and PC(WSRT) were much less than 0.05. This was because the absolute value order of positive value was larger than negative value order because of positive skew distribution and expectation of PC was more than 0, even if $\mu_1 > \mu_2$ ($R > 0$).

On the other hand, figure 4.3 shows the (tentative) type I error rate with both sides hypothesis ($\mathrm{H}_1 : \theta \neq 0$) in each parameter combination. From these figures, the (tentative) type I error rate of PC was much larger than 0.05 when ρ was small and τ was large, because PC had significance in the positive expectation value ($\theta > 0$).

Next, we show the results of (tentative) power with one side hypothesis in figure 4.2. The (tentative) power of DTS(t-test) was nearly equal to 0.8. The (tentative) powers of PC which was PC(WSRT) and PC(t-test) were less than 0.7 in all parameters combination. This trend became larger when τ was large or ρ was small. Especially, the (tentative) power of PC(t-test) was nearly equal 0 when τ was equal to or larger than 0.6 and ρ was equal to or less than 0.4. The condition to decrease the (tentative) power of PC depended on the skewness of distribution. The lager the distribution of skewness was (The larger the η is), the less the (tentative) power became. The (tentative) power of SPC which are SPC(WSRT) and SPC(t-test) was nearly equal to 0.8 when $\tau = 0.4$, and was a little less than 0.8 in $\lambda = -1, 1$ when $\tau \geq 0.6$. Regarding the (tentative) power with both sides hypothesis ($\mathrm{H}_1 : \theta \neq 0$) in figure 4.4, the (tentative) power of PC(t-test) or PC(WSRT) was less than other measures (DTS or SPC), because of the skewness of the distributions.

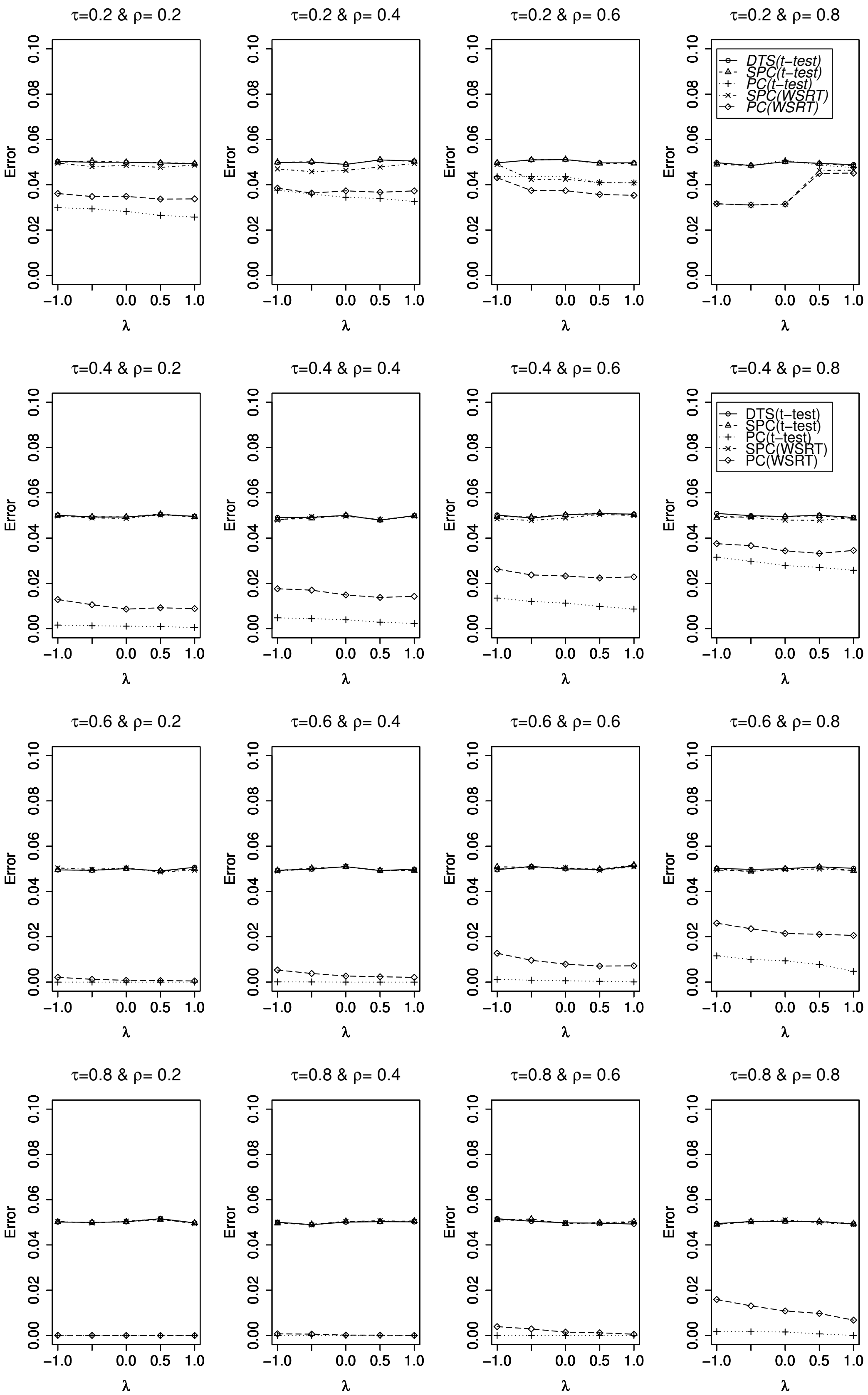


Figure 4.1: The relationship b/w type I error and λ (One sample & Set the 0.05 in one side)

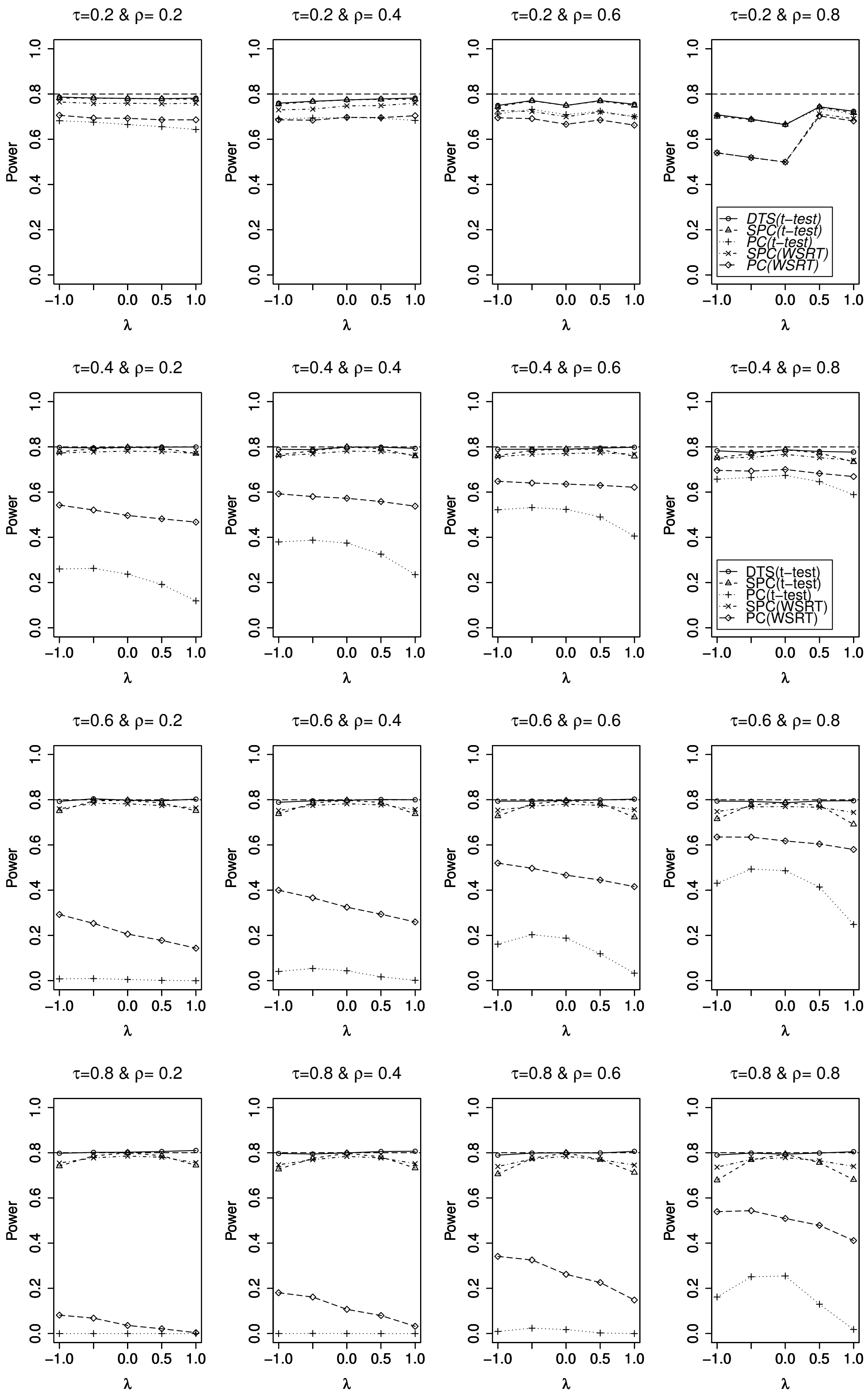


Figure 4.2: The relationship b/w power and λ (One sample & Set the 0.05 in one side)

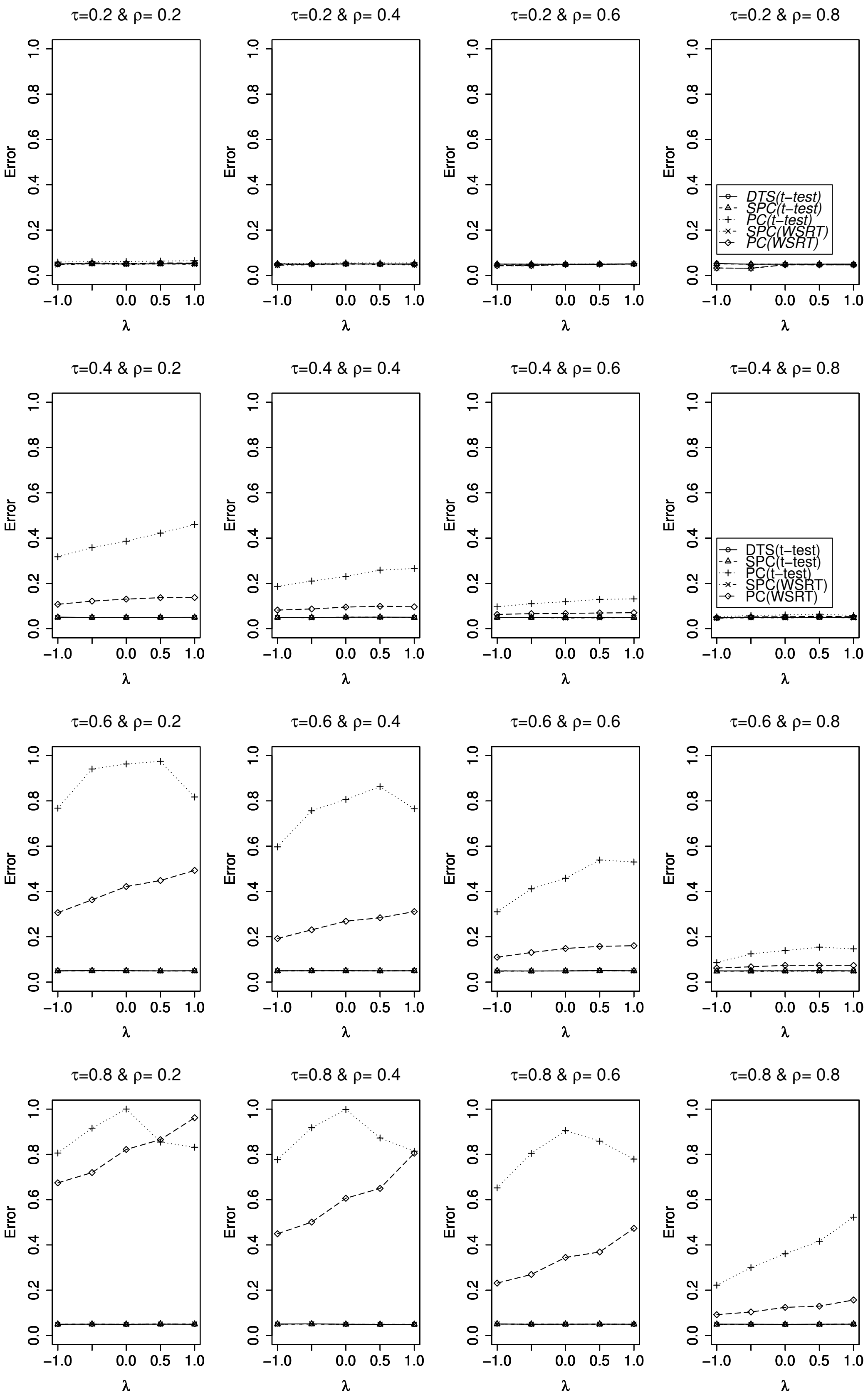


Figure 4.3: The relationship b/w type I error and λ (One sample & Set the 0.05 in both sides)

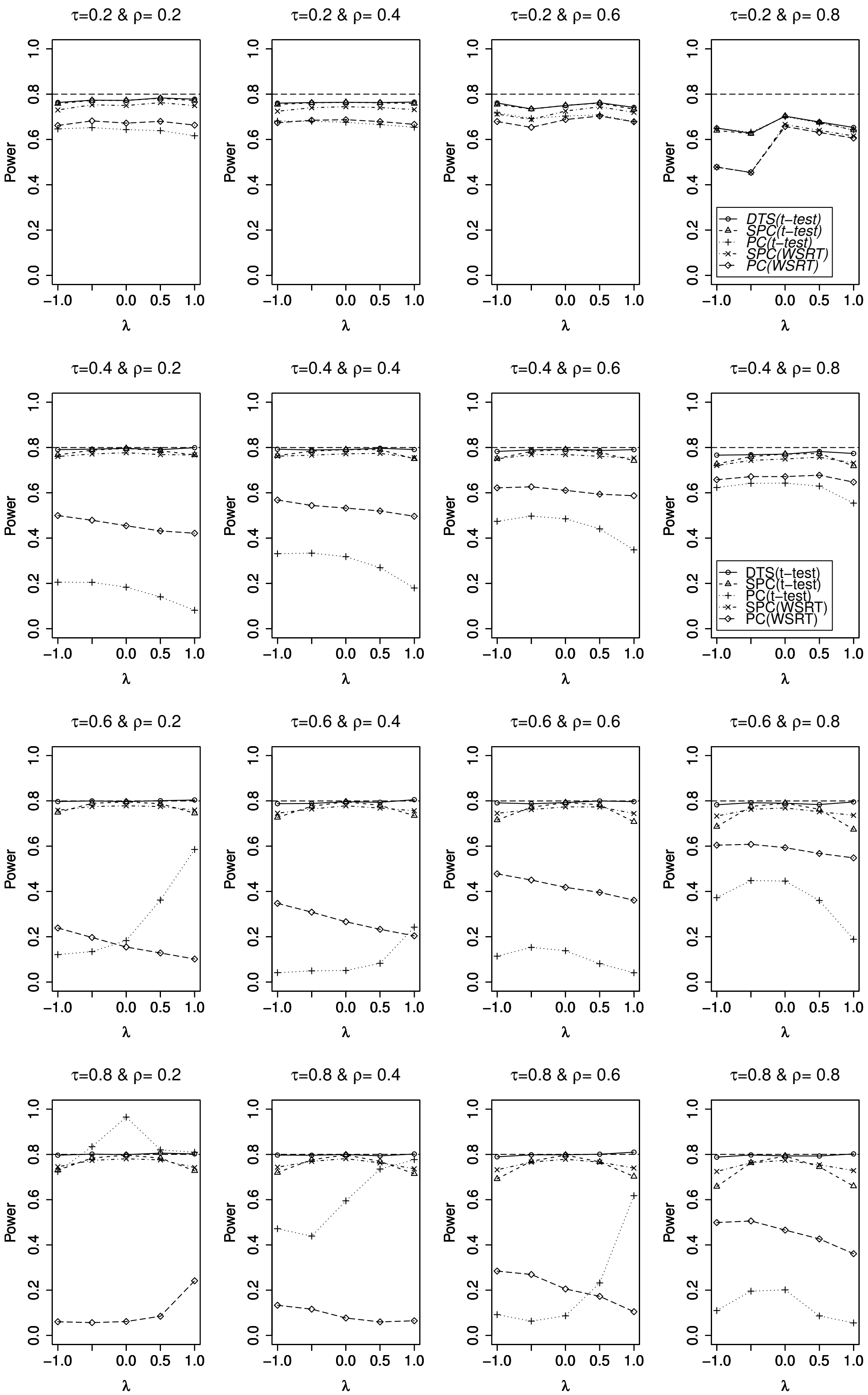


Figure 4.4: The relationship b/w power and λ (One sample & Set the 0.05 in both sides)

4.2 Simulation 2: Two samples comparison

4.2.1 Design of simulation 2

In this section, we evaluate the effect on statistical test of two samples which are treatment and control groups, when the distribution of PC or SPC has skewness. Pre- and post-data are assumed as BPN and the PC and SPC are calculated from pre- and post-data. And we also consider the situation that some effect is to reduce the post-data from pre-data as well as one sample comparison in previous section.

Hypothesis of the statistical test. In this simulation, the following hypotheses with 0.05 of significance level are set for three measures PC, SPC, DTS. Two-samples t-test is applied to PC, SPC and DTS and Wilcoxon Rank Sum Test (WRST) is also applied to PC. The PC and SPC has the same order because these two measures are functions of pre- and post-data and can show a relational expression($PC = 2 \times SPC/(1 - SPC)$). Therefore, the statistical results of PC(WRST) are the same as SPC(WRST). The hypotheses of interest are based on the one-side and are as follows,

$$\begin{aligned} & \mathrm{H}_0 : \theta_T = \theta_C, \\ \text{vs} \quad & \mathrm{H}_1 : \theta_T < \theta_C, \end{aligned}$$

where θ_T or θ_C are the expected value or median of treatment or control for PC, SPC(expected value only) and DTS(expected value only) in two samples test. In addition, we also evaluate the hypotheses based on the two sides,

$$\begin{aligned} & \mathrm{H}_0 : \theta_T = \theta_C, \\ \text{vs} \quad & \mathrm{H}_1 : \theta_T \neq \theta_C. \end{aligned}$$

Parameter setting. We set the 5 levels λ $(= -1, -0.5, 0, 0.5, 1)$, 4 levels τ $(= 0.2, 0.4, 0.6, 0.8)$, 4 levels ρ $(= 0.2, 0.4, 0.6, 0.8)$ and R(=0, 10 %), and then calculate the μ_1, μ_2, σ_1 and σ_2 based on the reparametrization method in each treatment group(Maruo, *et al.*, 2011; Maruo & Goto, 2012). Sample size is calculated based on the DTS(t-test), because DTS is normal in many cases and has highest power. Minimum sample size to exceed the power of 0.8 for DTS(t-test) sets in both hypotheses, and this size is double of one simulation 1.

We calculate the proportion of significance per total numbers of simulations about DTS, PC and SPC, when two-samples t-test or WRST (PC only) are applied. Total numbers of simulation is 100,000 times. The (tentative) type I error rate is defined as the proportion of significance when $R = 0$, and the (tentative) power is defined as the proportion of significance when $R = 10$. We can evaluate the loss of information based on the difference from the power of PC or SPC to 0.8, because sample size sets near the value of 0.8 for power of DTS.

4.2.2 Results of simulation 2

Figure 4.5 shows the (tentative) type I error rate for one side hypothesis. The (tentative) type I error rate of DTS(t-test), SPC(t-test), PC(WRST) were nearly equal to 0.05 in all parameter combinations. PC(t-test) was also nearly equal to 0.05 when $\tau = 0.2$ and $\tau = 0.4$. However, PC(t-test) was less than 0.05, when τ=0.6 and $\lambda = -1, +1$ or when τ=0.8 and $\lambda = -1, -0.5, +0.5, +1$.

On the other hand, figure 4.7 shows the (tentative) type I error rate for both side hypothesis. The trends of all measures were the same as the trends for one side hypothesis of figure 4.5. This was different from one sample results with both sides hypothesis and it was not shown that there was $\theta_T > \theta_C$.

Next, we show the results of (tentative) power of two sample test with one side hypothesis in figure 4.6. DTS(t-test) had the highest (tentative) power and the (tentative) power was nearly equal to 0.8. The (tentative) power of PC(WRST) and SPC(t-test) was almost same and PC(t-test) had the lowest (tentative) power. The (tentative) power of PC(WRST), SPC(t-test) and PC(t-test) decreased with increasing the absolute values of λ (λ=-1 or +1) when ρ and τ were constant. The reason why the (tentative) power of PC(t-test) became small was considered based on increasing the standard error of difference with increasing the distribution of skewness of PC. The (tentative) power of SPC(t-test) and PC(WRST) was nearly equal to 0.8 without regard to ρ and λ when τ=0.2 and 0.4. These two measures, SPC(t-test) and PC(WRST), had less (tentative) power than 0.8 when $\lambda = -1$ and +1. And the (tentative) power about two sample test with both sides in figure 4.8 were same trend as the (tentative) power with one side in figure 4.6.

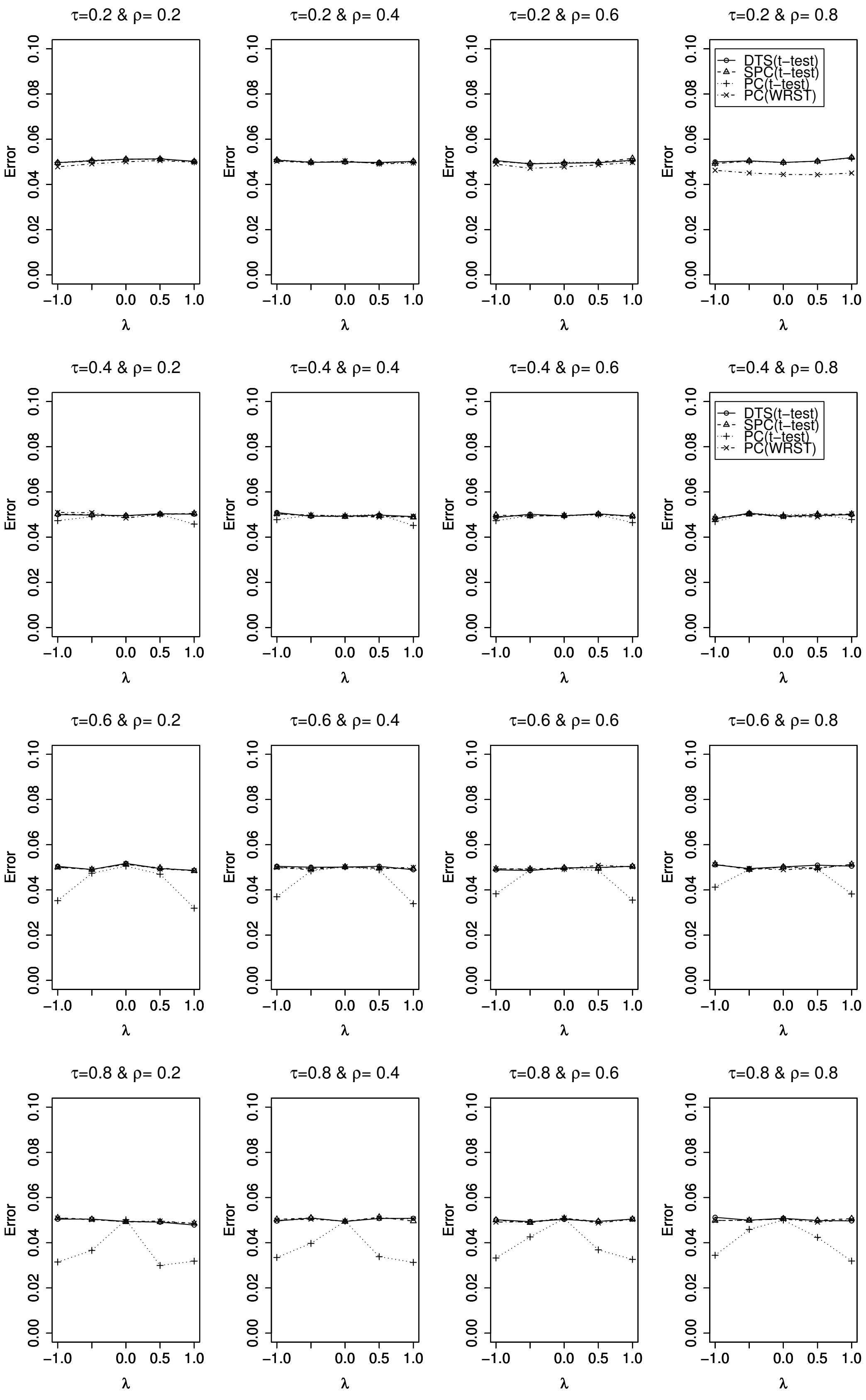


Figure 4.5: The relationship b/w type I error and λ (Two samples & Set the 0.05 in one sides)

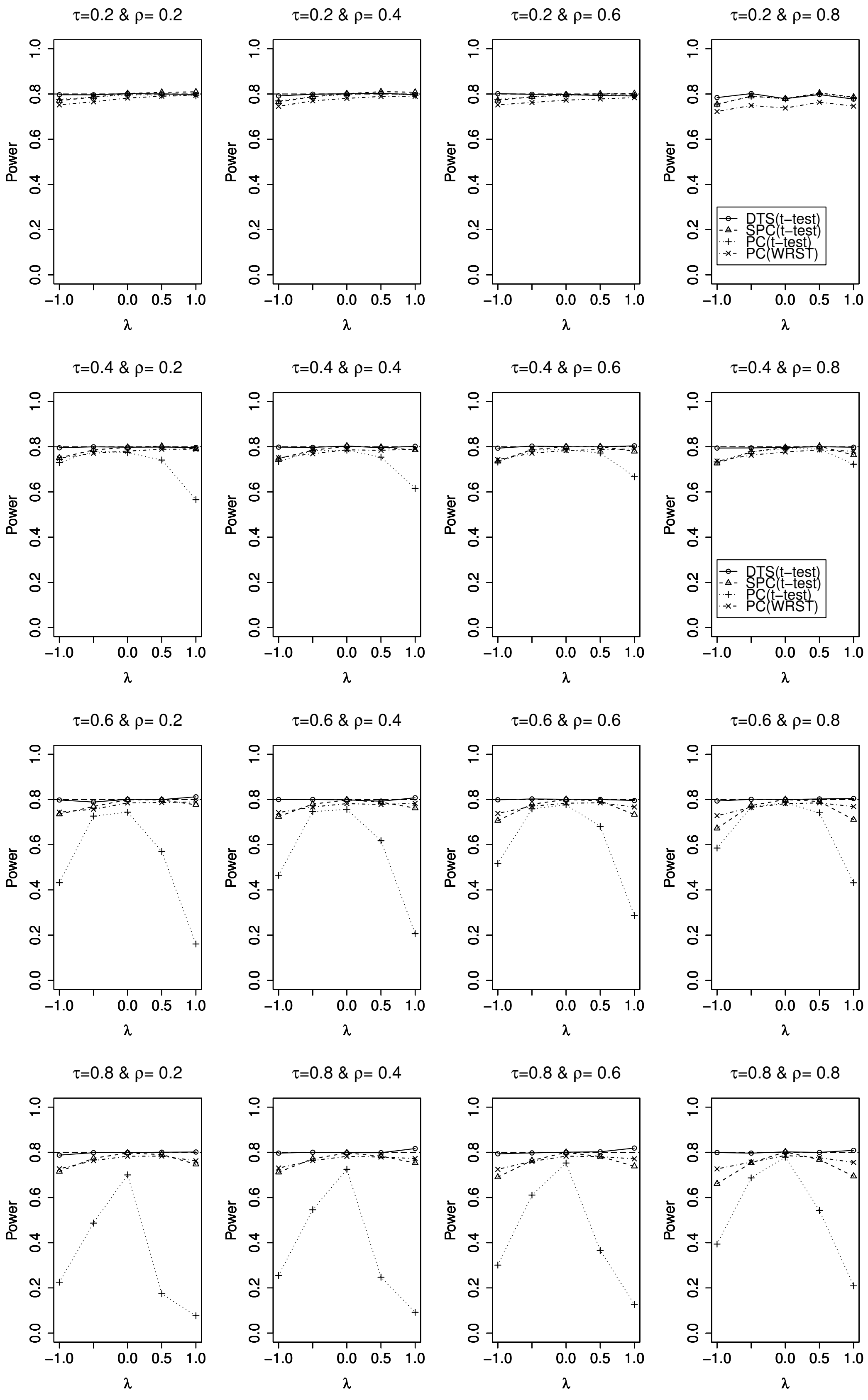


Figure 4.6: The relationship b/w Power and λ (Two samples & Set the 0.05 in one sides)

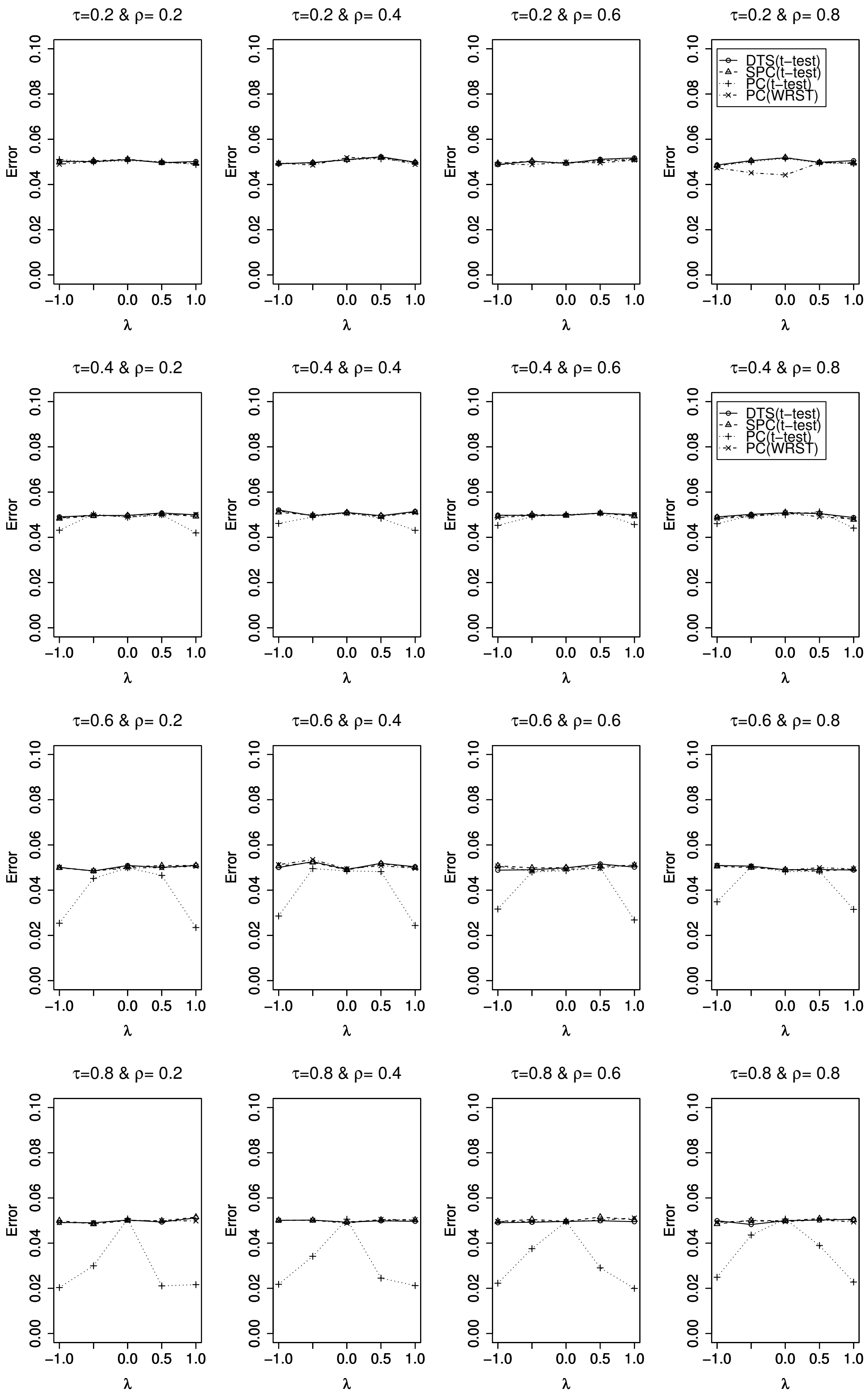


Figure 4.7: The relationship b/w type I error and λ (Two samples & Set the 0.05 in both sides)

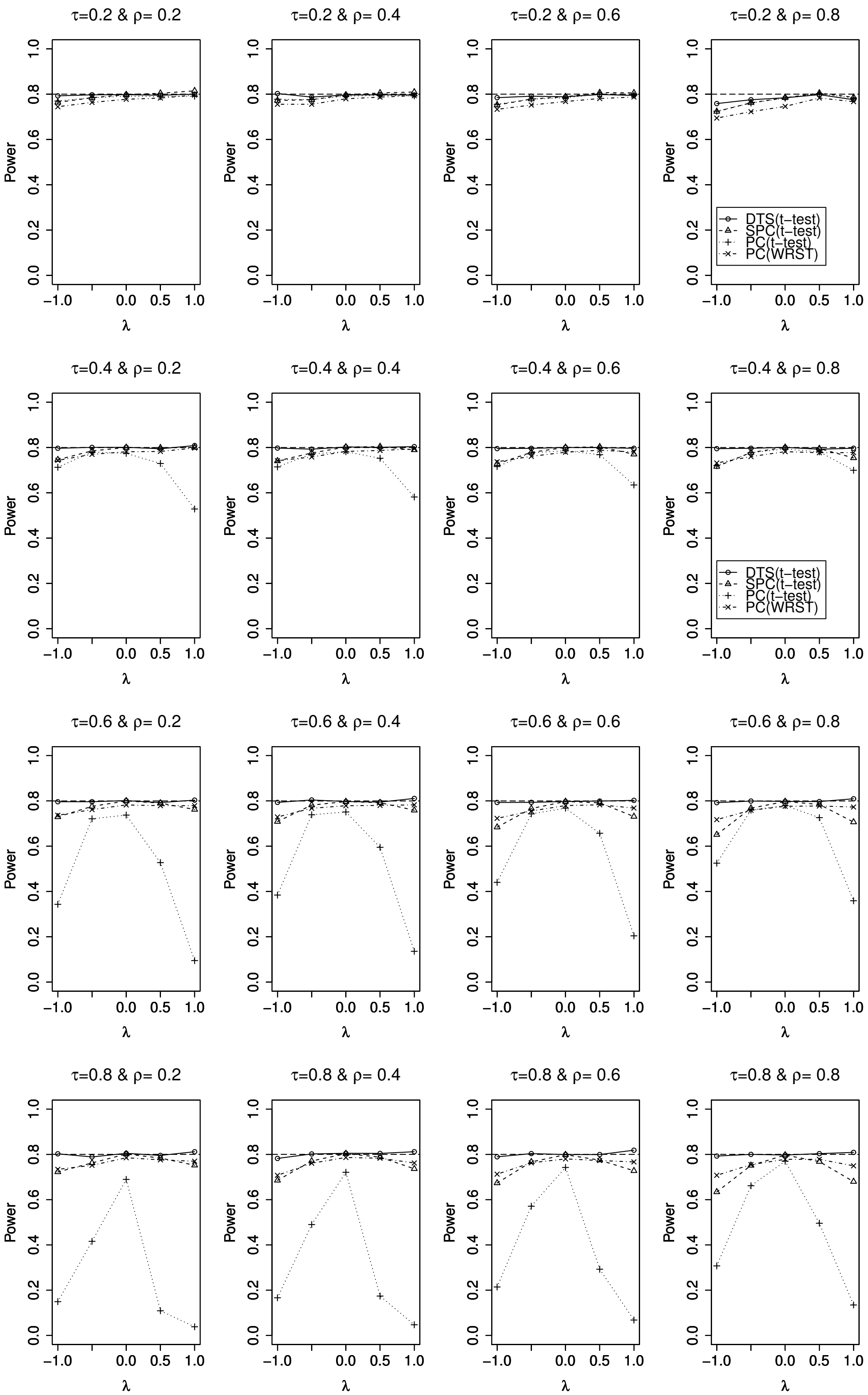


Figure 4.8: The relationship b/w Power and λ (Two samples & Set the 0.05 in both sides)

4.3 Case example

4.3.1 The application of symmetrized percent change to epilepsy data

In this section, we introduce the case example of SPC application as an example of the data analysis in phase III study of Gabapentin which is treated as an add-on therapy for refractory epilepsy. In this trial, SPC was named "Response Ratio (RR)".

The main objective of epilepsy treatment is to reduce the seizure frequency of each patient. Therefore, we evaluate the treatment effect to compare the seizure frequency of pre- and post-data in the clinical development of an antiepileptic drug. The evaluation of efficacy is based on the percent change from baseline in this field (French, 2001), because seizures are large variability in both intra- and inter-subjects and are skewness distribution for pre- and post-data. Additionally, when we evaluate seizure frequency (count data), the count data sometimes increase dramatically, such as from 10 counts of pre-data to more than 100 of post-data, if there are no treatment effects. For example, if a patient has 10 seizures of pre-data and 110 of post-data, then PC is 1000 %, and this value is too large. On the other hand, then SPC is 0.909 and is not too large. From this example, we consider SPC does not have less skew distribution than PC, because SPC does not give the too large value and shows robustness to outliers.

The 12-weeks, placebo controlled, double-blind study was conducted to evaluate the efficacy and safety of Gabapentin (Yagi & Sase, 2007 : Yamauchi *et al.*, 2006). This study set three treatment arms, 1200mg/day, 1800mg/day and placebo, and target population is the patients who had more than eight seizures in baseline period (pre-data) of 12 weeks. The 209 patients (86 patients in 1200mg/day, 41 patients in 1800mg/day and 82 patients in placebo) were included in the study, and per protocol set (PPS) was defined as the primary efficacy population. Table 4.4 shows the efficacy results of SPC about the comparison between 1200mg/day and placebo for PPS. The p-value was 0.0032 and statistical significance was shown for the comparison between 1200mg/day and placebo about 0.05 significance level.

Table 4.4: Efficacy Results of Gabapentin(From Goto *et al.,* 2007).

$SPC(RR)$	Placebo $n = 75$	Gabapentin (1,200mg/day) $n = 80$
Mean	-0.037	-0.144
SD	0.214	0.230
95%CI	$[-0.086, 0.012]$	$[-0.195, -0.093]$
Dif. b/w two groups	-0.107	
95%CI	$[-0.176, -0.038]$	
p value(t-test)	0.0032	

Next, we evaluate the shape of the sample distribution about the seizure data in the above study. Figure 4.9 shows the histogram of pre- and post-seizure data in 1200mg/day and placebo and shows the data driven power normal distribution(Goto *et al.,* , 2007; Goto *et al.,* 1979; Goto *et al.,* 1983). The estimate value of shape parameter (λ) is follows.

	1,200mg/day	Placebo
Pre-data	-0.38	-0.37
Post-data	-0.14	-0.23

These estimated values of λ are negative near 0, and this result shows that the distribution of the seizure data in this study can approximate the log-normal distribution.

Figure 4.10 shows the histogram of SPC data in 1200mg/day and placebo and shows the distribution given by expression (3.4). In addition, we show the estimated value of λ as follows. From these results, it was shown that the distribution of SPC was nearly equal to a normal distribution.

	1,200mg/day	Placebo
SPC	1.45	0.94

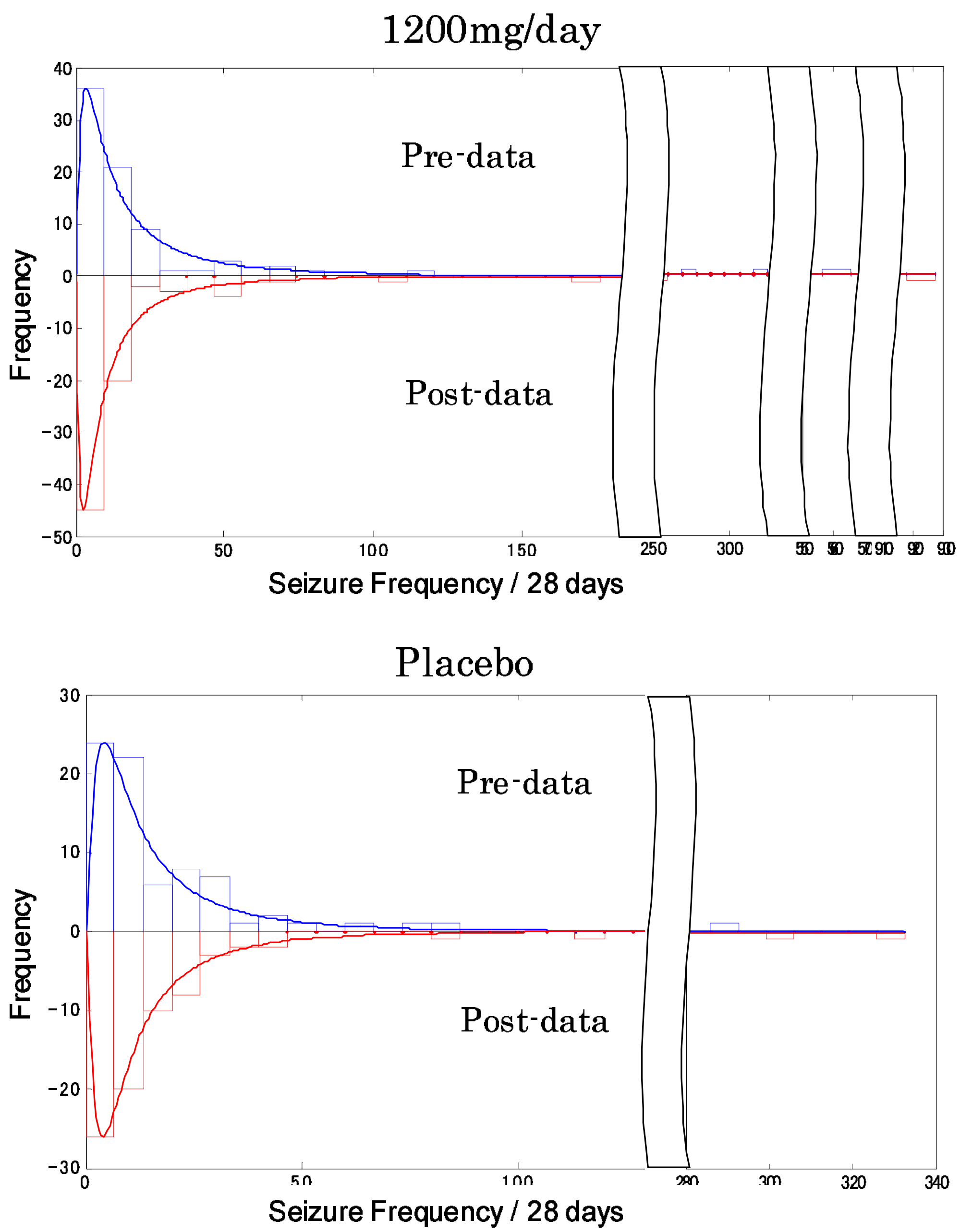


Figure 4.9: The histogram and applied power-normal distribution to seizure frequency (log-normal distribution): From Goto *et al.,* (2007)

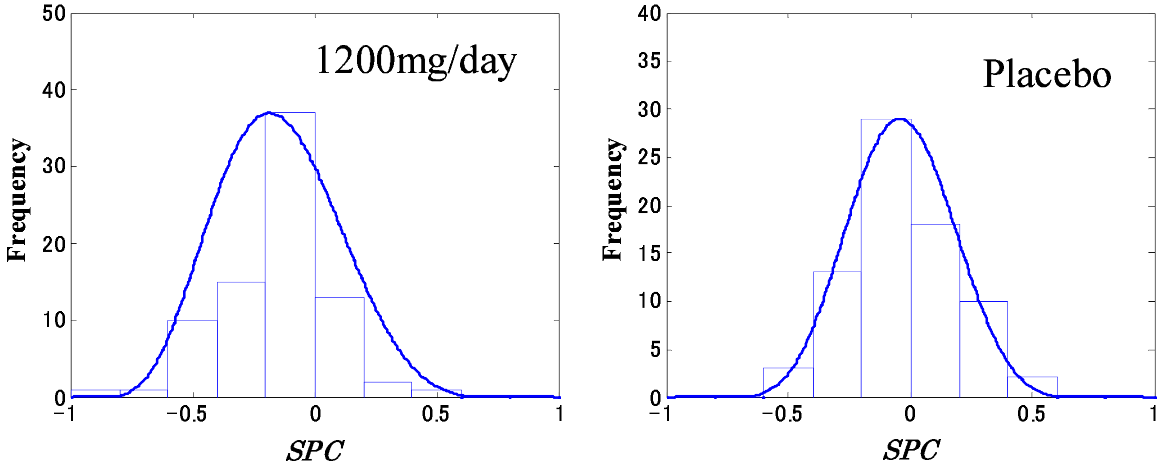


Figure 4.10: The histogram and applied distribution to SPC (given by expression (3.4)): From Goto *et al.*, (2007)

Section 5

Conclusion

5.1 Results and productive findings of this study

Both PC and SPC are used as ratio measures in a clinical trial in which a treatment effect is evaluated. However, PC was shown to have some difficulty to apply the statistical methodology based on the parametric methodology(Asakura *et al.*,, 2011; Pharm-Gia *et al.*, 2006; Senn & Julious, 2009), and SPC was not clear in the statistical properties. In this paper, we investigated statistical properties of PC and SPC in which declaration of pdf, evaluation of skewness and evaluation to statistical power are included. And we propose how to apply the two measures in various situations in later paragraphs. In addition, we declared the relationship between SPC and coefficient of variation (CV).

Statistical properties of PC. The distribution of PC was positively skewed when post-data was much larger than pre-data. This condition arises in the combinations of the following points.

1. The scale like parameter (τ) of pre- and post-data became large.
2. The correlation parameter (ρ) between pre- and post-data became small.
3. The distribution of pre- and post-data became far from bivariate log-normal distribution, which means that the distribution becomes more positively skewed than log-normal distribution (λ becomes close to -1) or that the distribution is close to normal and becomes negative skew (λ becomes close to $+1$).

The τ gave the largest contribution to a skewness of distribution for PC. The second largest contribution was ρ and third was λ. It is difficult to identify the condition to symmetrize the distribution, because the cause of distribution skewness exists more than one component and is from the combination of the components. Therefore, we recommend to confirm each distribution skewness in each situation taking the above results into consideration before applying the statistical methodology.

From results of comparison between pre- and post-data based on one sample statistical test, the (tentative) type I error of PC(t-test) and PC(WSRT) became extremely low from the pre-defined significance level and the (tentative) power decreased more 0.1 (10 %) than DTS, which means that the (tentative) power was less than 0.7 in all conditions. From results of two group comparisons based on the two samples statistical test, the (tentative) type I error of PC(t-test) became nearly equal to or slightly less than the pre-defined significance level and the (tentative) power decreased extremely when the scale like parameter (τ) became large or the distribution became far from log-normal (absolute value of λ is +1). On the other hand, the (tentative) type I error of PC(WSRT) became nearly equal to pre-defined significance level and the (tentative) power was only slightly smaller than DTS.

Statistical properties of SPC. The distribution of SPC kept symmetry without regard to shape, scale like and correlation parameters of the distribution for pre- and post-data. Therefore, we can consider the application of statistical analysis based on the parametric methodology. From results of comparison between pre- and post-data based on one-sample test, the (tentative) type I error of SPC(t-test) became nearly equal to pre-defined significance level and the (tentative) power of SPC(t-test) slightly decreased when scale like parameter was large (especially $\tau \geq 0.6$) and the distribution of pre- and post-data was far from log-normal. However, the (tentative) power of SPC(t-test) was larger than PC(t-test) or PC(WSRT). From results of two groups comparison based on the two samples statistical test, the (tentative) type I error of SPC(t-test) became nearly equal to pre-defined significance level and the (tentative) power was also nearly equal to or slightly less than DTS. In addition, the (tentative) power of SPC(t-test) was also same as PC(WRST) in all conditions.

Proposal how to apply two measures. We can use the PC(WRST) when the objective is to evaluate the two groups comparison based on the ratio measures. Because PC is easy to interpret and PC(WRST) can keep high power. However, it is necessary to investigate the possibility of application carefully, if we know the factors to affect the treatment effect, if we estimate the effect based on the statistical methodology such as analysis of covariance(ANCOVA) and if we apply the PC(t-test) for groups comparison. We need to make sure preliminarily whether or not the assumption to apply the PC(t-test) are satisfied. If assumptions are not satisfied or if assumptions cannot be confirmed, then we can analyze the data to apply the SPC and can interpret the results after transforming the robust percent change (RPC) proposed by Berry(1989).

In addition, we consider the statistical analysis should be done based on the SPC, when the objective is to compare the one-group comparison. Because the type I error keeps significance level and the power became large. The SPC is necessary to re-transformation for interpretation of the results. However, we consider that SPC is one of the favorable options for ratio measures, because we can use SPC in various shape of distributions. For example, when we analyze the change between pre- and post data of laboratory items, difference or percent change are only applied based on the past experience. We think laboratory items such as triglyceride (TG), which has positive skew distribution and becomes primary or secondary efficacy endpoint, should be applied SPC.

When we select a measure of effect in a clinical trial, a difference or a percent change are only applied based on the past experience without investigating the statistical properties so much. The important thing when selecting the measures of effect is to define the goal of statistical analysis definitely, and is to evaluate statistical properties, such as a skewness of distribution or a power in addition to evaluation from the clinical points of view.

Relationship between SPC and CV. There is the relationship between SPC and CV which is that the numerator of CV is replaced by the numerator of SPC with absolute value. Therefore, the distribution of CV is the distribution folded back negative value of SPC to positive.

5.2 Subjects for future investigation

As we mentioned in the previous section, SPC needs to be transformed into appropriate measures such as RPC proposed by Berry (1989) for interpretation. And Berry & Ayers (2006) also mentioned that the investigator would report an estimated RPC with an appropriately calculated standard error or confidence interval. However, there is no research about the standard error or the confidence interval. Therefore, it would be desirable to propose these in future research.

Appendix A

Reparametrization.

We show the reparametrization method to apply in section 2 of this paper(Maruo & Goto, 2008; Maruo, *et al.* 2010; Maruo & Goto, 2012). The τ is defined as the scale parameter. When $\lambda \neq 0$, μ and σ cannot be obtained from $\lambda, \xi_{0.5}$, and τ explicitly. Thus, they have to be calculated based on the grid search method. The calculation process for μ and σ is given as follows:

S1. Give $\lambda, \xi_{0.5}$, and τ, and set $K^{(\tau)} = \{-100, -99.9, \ldots, 99.9, 100\}$. $K_i^{(\tau)}$ is the ith factor of $K^{(\tau)}$ $(i = 1, \ldots, 2000)$.

S2. Calculate $\mu_i^{(\tau)}$ and $\sigma_i^{(\tau)}$ for all $K_i^{(\tau)}$ based on (A.1). Then evaluate $\delta_i^{(\tau)} = \xi_{0.75i}^{(\tau)} - \xi_{0.25i}^{(\tau)} - \xi_{0.5}\tau$ and replace $\delta_i^{(\tau)}$ that can not be evaluated for some reason (e.g., obtained as NaN or infinity because of calculation precision of computers) and seven values on both sides of it by sufficiently large values, where $\xi_{pi}^{(\tau)}$ is the percentile of the power-normal distribution with parameters: $\lambda, \mu_i^{(\tau)}$, and $\sigma_i^{(\tau)}$.

$$\mu = \begin{cases} \left(1 + \dfrac{z_{0.5^*}}{K}\right)^{-1} \left(\dfrac{\xi_{0.5}^{\lambda} - 1}{\lambda - \dfrac{z_{0.5^*}}{\lambda K}}\right), & K \neq 0, \\ -\dfrac{1}{\lambda}, & K = 0, \end{cases} \qquad \sigma = \begin{cases} \dfrac{1 + \lambda\mu}{\lambda K}, & K \neq 0, \\ \dfrac{\xi_{0.5}^{\lambda}}{\lambda z_{0.5^*}}, & K = 0, \end{cases} \tag{A.1}$$

S3. Set $i_{min} = \arg\min_i \mathrm{abs}(\delta_i^{(\tau)})$. If $\delta_i^{(\tau)} < 0$, divide $\{K_{i_{min}}^{(\tau)}, K_{i_{min+1}}^{(\tau)}\}$, else divide $\{K_{i_{min}-1}^{(\tau)}, K_{i_{min}}^{(\tau)}\}$ into required accuracy of intervals (e.g., 0.0001) and replace $K^{(\tau)}$ by this set. Repeat S1 and S2.

S4. Set $i_{min} = \arg\min_i \mathrm{abs}(\delta_i^{(\tau)})$. Calculate μ and σ from $K_{i_{min}}^{(\tau)}$.

This calculation process can be performed by any programming language capable of parallel computation. The range $[-100 \leq K \leq 100]$ covers almost any realistic situations, but cannot be calculated because K tends to become too large in the neighborhood of $\lambda = 0$. When $\lambda = 0$,

μ and σ can be calculated explicitly:

$$\mu = \log(\xi_{0.5}),\ \sigma = \frac{\log\{(\tau + \sqrt{\tau^2 + 4})/2\}}{z_{0.75}}.$$

Fig.A.1 illustrates the relations between $A(K)$ and λ for $\xi_{0.5} = 100$ and $\tau = 0.1, 0.3$, and 0.5. Simulations where the truncation is ignored should be run for $\tau < 0.5$. In addition, this relationship is invariant for $\xi_{0.5}$ though we set $\xi_{0.5} = 100$ in this figure.

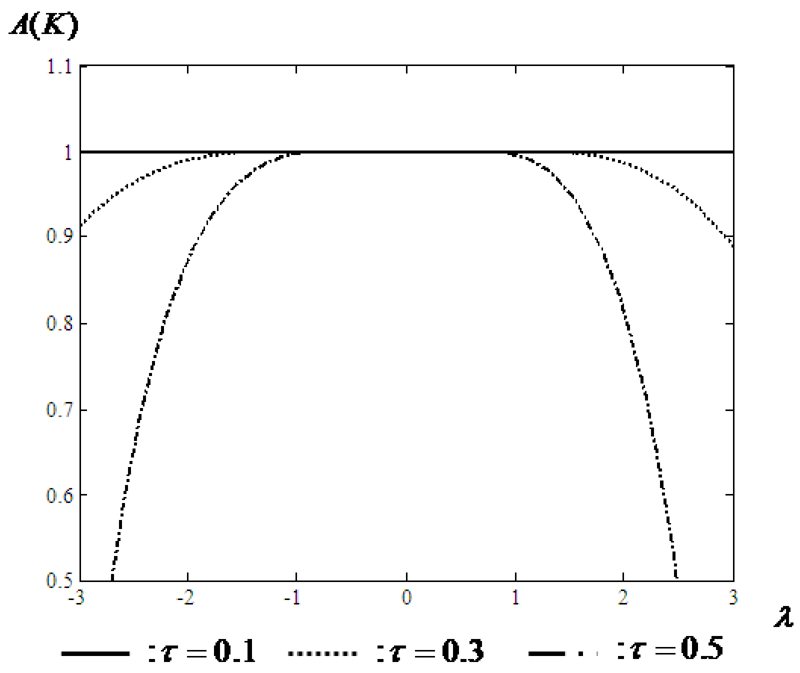


Figure A.1: The relationship b/w A(K) and λ for $\xi_{0.5} = 100$, and $\tau = 0.1, 0.3$ and 0.5.

Reference

[1] Adachi, H., Imaizumi, T., Murakami, M. and Abe, M (2009):A phase III, randmized, parallel-group comparative study of Caduet (an amlodipine/atorvastatin combination drug) in patients with concurrent hypertension and hyper-LDL-cholesteremia. *Journal of New Remedies & Clinics*, **58,** 2(1496)-16(1510).

[2] Asakura, K., Uesaka, H., Sugimoto, T. and Hamasaki, T.(2011):A Note on Analysis of Ratio of Two Correlated Normal Variables. *Japanese Journal of Applied Statistics*, **40(1)**, 53-71 (in Japanese).

[3] Berry, D. A.(1989) *Statistical Methodology in the Pharmaceutical Sciences.*Marcel Dekker, New York.

[4] Berry, D. A. and Ayers, G. D.(2006): Symmetrized percent change for treatment comparisons. *The American Statistician*, **60 :** 27-31.

[5] Bonate, P. L.(2000):*Analysis of Pretest-Posttest Designs.* Boca Raton: Chapman and Hall.

[6] Box, G.E.P. and Cox, D.R. (1964): An analysis of transformations (with discussion). *J. Roy. Statist. Soc.,*, **B26(2)**, 211-246.

[7] Brouwers, P. and Mohr, E.(1989): A metric for the evaluation of change in clinical trials. *Clinical Neuropharmacology*, **12 :** 129-133.

[8] Douglas, M.F., Marian, K., C.T. Viswanathan, Jacquelin, O'S., S. PeterK., Ajai, C., Russell, W., Anthony, J, D., and Daniel T.(2009). Workshop Report and Follow-Up-AAPS Workshop on Current Topics in GLP Bioanalysis:Assay Reproducibility for Incurred Samples-Implications of Crystal City Recommendations. *American Association of Pharmaceutical Scientists*, **Vol.11, No.2**, 238-241.

[9] Endrikat, J., Graeser T., Mellinger U., Ertan, K. and Holz, C.(2007):A multicenter, prospective, randomized, double-blind, placebo-controlled study to investigate the efficacy of a continuous-combined hormone therapy preparation containing 1mg estradiol valerate/2 mg dienogest on hot flushes in postmenopausal women. *Maturitas*, **58**, 201-207.

[10] French, J. A.(2001). Proof of efficacy trials : endpoints. *Epilepsy Research*, **45**, 53-56.

[11] Goto, M., Matsubara, Y. and Tsuchiya, Y. (1983): Power-normal distribution and its applications. *Rep. Stat. Appl. Res., JUSE*, **30**, 8-28.

[12] Goto, M. and Hamasaki, T. (2002): The bivariate power normal distribution. *Bulletin of Informatics and Cybernetics*, **34(1)**, 29-49.

[13] Goto, M., Yamabe. T., Maruo, K. and Kawai, N.(2007):Statistical Data Analysis based on Response Ratio. *Japanese Journal of Clinical Psychopharmacology*, **10**, 667-676 (in Japanese).

[14] Hamasaki, T. and Goto, M.(2002): On inference of Parameters in the Bivariate Power-Normal Distribution. *The Japanese Journal of Behaviormetrika*, **29(2)**, 199-222(in Japanese).

[15] Homma, Y., Paick, J.S., Lee J.G. and Kawabe K. on behalf of the Japanese and Korean Tolterodine Study Group (2003): Clinical efficacy and tolerability of extended-release tolterodine and immediate-release oxybutynin in Japanese and Korean patients with an overactive bladder: a randomized, placebo-controlled trial. *BJU Int.*, **92(7)**, 741-747.

[16] Janney, C., A., Cauley, J., A., Cawthon, P. M. and Kriska, A. M.(2010): Longitudinal Physical Activity Changes in Older Men in the Osteoporotic Fractures in Men Study. *Journal of the American Geriatrics Society.*, **58(6)**, 1128-1133.

[17] Kaiser, L.(1989): Adjusting for Baseline: Change or Percentage Change? *Statistics in Medicine*, **8**, 1183-1190.

[18] Kitazawa, Y., and KP2035 Study group(2009): Phase III double-blind study of latanoprost/timolol combination (KP2035) in patients with primary open-angle glaucoma or ocular hypertension. Japanese Journal of Clinical Ophthalmology, **63(5)**, 807-315 (in Japanese).

[19] Koti, K.M.(2001): On a primary efficacy endpoint. *Drug Information Journal*, **35**, 157-162.

[20] Lane, N. E., Schnitzer, T. J., Birbara, C. A., Mokhtarani, M., Shelton, D. L., Smith, M. D. and Brown, M. T.(2010):Tanezumab for the Treatment of Pain from Osteoarthritis of the Knee. *The new England journal of medicine*, **363(16)**, 1521-1531.

[21] Mario L. Rocci, Jr., Viswanath Devanarayan, David B. Haughey, and Paula Jardieu.(2007). Confirmatory Reanalysis of Incurred Bioanalytical Samples. *American Association of Pharmaceutical Scientists*, **Vol.9, No.3**, E336-E343.

[22] Maruo, K. and Goto, M. (2012): Percentile estimation based on the power-normal distribution. *Computational Statistics* (in press).

[23] Maruo, K. & Goto, M. (2008). On estimation of parameters in power-normal distribution. *Proceedings of IASC 2008* (*Joint Meeting of the 4th World Conference of the IASC and the 6th Conference of the Asian Regional Section of the IASC on Computational Statistics & Data Analysis*), 1130-1139.

[24] Maruo, K., Shirahata, S. and Goto, M.(2011): Underlying assumptions of the power-normal distribution. *Behaviormetrika*, **38, No. 1**, 85-95.

[25] Maruo, K., Shirahata, S., Goto, M. and Komazawa, T.(2008): Statistical investigation of reference intervals of clinical laboratory data. *The Japanese Journal of Behaviormetrika*, **35(1)**, 73-89(in Japanese).

[26] Pham-Gia, T., Turkkan, N. and Marchand, E.(2006): Density of the Ratio of Two Normal Random Variables and Applications. *Communications in Statistics.Theory and Methods.*, **35**, 1569-1591.

[27] Satoh, J., Yagihashi, S., Baba, M., Suzuki M., Arakawa A., Yoshiyama, T. and Shoji S. (2010): Treatment efficacy and safety of pregabalin for treating neuropathic pain associated with diabetic peripheral neuropathy: a 14 week, randomized, double-blind, placebo-controlled trial. *Diabetic Medicine*, **28 (1)**, 109-116.

[28] Senn S. and Julious, S. (2009): Measurement in clinical trials: A neglected issue for statisticians?. *Statistics in Medicine*, **28**, 3189-3209.

[29] Steel, R. G. D., and Torrie, J. H. (1980): Principles and Procedures of Statistics (2nd ed.), New York: McGraw-Hill.

[30] Törnqvist L, Vartia P, and Vartia YO. How Should relative Changes Be Measured ? *The American Statistical Association* 1985; 39(1):43-6.

[31] Uesaka, H. and Goto, M.(1980):Analysis of laboratory data based on the power normal distribution. *Japanese Journal of Applied Statistics*, **9**, 23-33 (in Japanese).

[32] Yagi, K., and Sase, S.(2007). Clinical efficacy of Gabapentin. *Japanese Journal of Clinical Psychopharmacology*, **10**：641 − 649(in Japanese).

[33] Yamauchi, T., Kaneko, S., Yagi, K. and Sase, S.(2006):Treatment of partial seizures with gabapentin: Double-blind, placebo-controlled, parallel-group study. *Psychiatry and Clinical Neurosciences*, **60**, 507-515.

[34] Yamabe, T., Maruo, K., Shirahata, S. and Goto, M.(2012):Statistical properties of two ratio measures based on pre- and post observed values which are assumed as bivariate power normal distribution. *Japanese Journal of Applied Statistics* in press. (in Japanese).

[35] 椿 広計，藤田利治，佐藤俊哉（2004）：これからの臨床試験 医薬品の科学的評価—原理と方法, 朝倉書店.

List of publication

[1] Goto, M., Yamabe. T., Maruo, K. and Kawai, N.(2007): Statistical Data Analysis based on Response Ratio. *Japanese Journal of Clinical Psychopharmacology*, **10**, 667-676 (in Japanese).

[2] Moroi, Y., Yamabe, T., Shibata, O. and Abe, Y.(2000): Apparatus for measuring the evaporation rate of water across an air/water interface. *Langmuir*, **16 (25)**, 9697 -9698.

[3] Nagashima, H., Suzuki, M., Araki, S., Yamabe, T., Muto, C. and PF-04383119 study group.(2011): Preliminary assessment of the safety, and efficacy of tanezumab in Japanese patients with moderate to severe osteoarthritis of the knee: a randomized, double-blind, dose escalation, placebo-controlled study. *Osteoarthritis and Cartilage*, **19(12)**, 1405-1412.

[4] Nagashima, H., Suzuki, M., Araki, S., Yamabe, T., Shoji, S. and PF-04383119 study group.(2010): Preliminary assessment of the safety, and efficacy of tanezumab in Japanese patients with moderate to severe osteoarthritis of the knee: a randomized, double-blind, dose escalation, placebo-controlled study. *Proceedings of Asia Pacific League of Association for Rheumatology (APLAR)*, Hong Kong, China.

[5] Osawa, M., Shirasaka, Y., Ohtsuka, Y., Imai, K., Mimaki, M., Sasaki, M., Tohyama, J., Akasaka, N., Iyoda, K., Yamabe, T., and Machii, K.(2011): Efficacy and safety of gabapentin adjunctive therapy in Japanese pediatric refractory partial epilepsy. *Japanese Journal of Clinical Psychopharmacology*, **14(7)**, 1205-1222 (in Japanese).

[6] Yamabe, T. and Goto, M.(2010): Statistical properties of two ratio measures based on pre- and post observed values which are assumed as bivariate power normal distribution. *Proceedings of the 24th Symposium of Japanese Society of Computational Statistics*, Osaka, Japan (in Japanese).

[7] Yamabe, T. and Moroi, Y.(1999): Micelle formation of anionic surfactant with divalent counterion of separate electric charge. *Journal of Colloid and Interface Science*, **215(1)**, 58-63.

[8] Yamabe, T., Moroi, Y., Abe, Y. and Takahashi, T.(2000): Micelle formation and surface adsorption of N-(1,1-Dihydroperfluoroalkyl)-N,N,N-trimethylammonium chloride. *Langmuir*, **16 (25)**, 9754 -9758.

[9] Yamabe, T., Isogawa, N., Maruo, K. and Goto, M.(2011): Statistical properties of symmetrized percent change and percent change based on the bivariate power normal distribution. *Proceedings of The 58th World Statistics Congress of the International Statistical Institute (ISI) in 2011*, Dublin, Ireland.

[10] Yamabe, T., Maruo, K., Shirahata, S. and Goto, M.(2012): Statistical properties of two ratio measures based on pre- and post observed values which are assumed as bivariate power normal distribution. *Japanese Journal of Applied Statistics* in press. (in Japanese).