2009

Equivariance and generalized inference in location-scale family

Fuqi Chen
University of Windsor

Equivariance And Generalized Inference In Location-Scale Family

by

Fuqi Chen

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Mathematics and Statistics
in Partial Fulfillment of the Requirements for
the Degree of Master of Science at the
University of Windsor

Windsor, Ontario, Canada

2009

Author's Declaration of Originality

I hereby certify that I am the sole author of this thesis and that no part of this thesis has been published or submitted for publication.

I certify that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

Abstract

In this thesis, we revisit some statistical problems, where classical inference can not provide small-sample optimal solution. These problems motivated Tsui and Weerahandi (1989), and Weerahandi (1993) to introduce the concepts of generalized inference which consist in constructing generalized test variable (GTV) and generalized pivotal quantity (GPQ). However, in general location-scale family, the existing literatures do not provide any systematic method for deriving these quantities.

To overcome this problem, the equivariance principle is applied to construct GTV and GPQ in location-scale family. Namely, we construct the GPQ and GTV for the parameters of interest in one-sample and two-sample families cases. Particularly, we study inference problem concerning the difference between two location parameters.

The simulation studies show that the suggested methods preserve the nominal level, and provides satisfactory power in small and moderate sample sizes. Finally, some real data sets are analyzed in order to illustrate the application of the suggested procedures.

Acknowledgements

I would like to express my profound gratitude to my supervisor Dr. Nkurunziza. He did not only lead me to Statistics, but also to a wonderful world. During my graduate life, he never hesitated to help me whenever I need help throughout my study. He also advises many valuable suggestions, which lead this work to be successful. I am very grateful to him for the Research Assistantship he has provided to me for this research. I also thank University of Windsor for the Graduate Assistantship. These financial supports help me to finish my master program easier.

In addition, my great gratitude also goes to Dr. Ghrib, Dr. Hussein and Dr. Hlynka, the committee members during the oral defense. Thanks for having read a draft of this thesis and having made their precious comments and suggestions. As for any remaining errors, the responsibility rests entirely upon myself.

Last but not least, I thank my parents. They have given me constant love, encouragement and support, and have always been eager to help me in various ways that they could. Their attitudes to life have guided me in the right direction.

Finally, I would like to thank everyone who directly or indirectly offered the help to this thesis.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

1.1. Introduction

The study of generalized confidence interval was motivated by the fact that the small sample optimal confidence intervals (CI) in statistical problems involving nuisance parameters may not be available. For example, exact confidence intervals based on minimal sufficient statistics are not available (see Weerahandi, 1993) in the problem of constructing confidence interval for the difference in means of two exponential distributions, or two normal distributions with different variances. For such problems, when the sample sizes are small, there do not exist optimal solutions based on the classical pivotal method. To overcome these problems, Tsui and Weerahandi (1989) introduced a concept called generalized P-value (GPV) and later on Weerahandi (1993) developed the generalized confidence interval (GCI). In the above papers, similar to the construction of the classical p-value and confidence interval, the GCI are established by constructing a related quantity called generalized pivotal quantity (GPQ) and GPV by generalized test variable (GTV).

In hypothesis testing, the concepts of GPV and GCI are used as on extension of the classical P-value and confidence interval, respectively, and have performed well in obtaining P-values and confidence intervals for those cases where the classical procedures do not give satisfactory results. For example, Weerahandi (1993) applied the generalized confidence intervals to the difference in two exponential means and two normal means. In addition, Bebu and Mathew (2007) developed a generalized pivotal quantity for comparing the means and variances of a bivariate log-normal distribution.

However, there are some limitations of the methods proposed in the quoted papers. In fact, the authors in these papers deal only with some special distributions, and the provided inference methods are not applicable for all families of parametric models. So far, there is no systematic method of constructing GPQ applicable to all families of parametric models. Motivated by these limitations, the purpose of the presented thesis is to develop the appropriate method of constructing the GPQ for the general location-scale family. Further, based on the GPQ, we establish the GCI and GPV of the location and scale parameters.

The thesis is organized as follows. In Chapter 1, we present the mathematical statistics background, whereby we clarify the concepts used to establish the framework. To reach the purpose, we start with the univariate location-scale family, then extend the approaches to the bivariate case. Univariate problems are discussed in Chapter 2, we consider some well known location-scale families such as normal, gamma distribution, where the GPQ are well known and easy to be constructed. Then, based on those examples, we show a method about how to construct the GPQ.

Based on these ideas we construct the GPQ for location scale parameters, by extending the methods suggested in Lawless (1972) who constructed CI for the location and scale parameters of the Cauchy and logistic distributions, by using some classical pivotal quantities conditional on some ancillary statistics. It should be noted that, in Lawless (1972), the conditional pivotal quantities are based on the maximum likelihood estimator (MLE), which can be found in Cauchy and logistic families. However, MLE may not exist in some special cases (see Pitman, 1979). Therefore, in extending the Lawless (1972) method to a generalized form in term of the general location-scale family, we consider the case that MLE does not exist. We replace the MLE by the minimum risk equivariant estimator (MRE), which will be discussed in Chapter 3.

Furthermore, based on the univariate case, we develop the GPQ for the bivariate location-scale family. For the bivariate case, Sprott (2000, Chapter 7) provided an approach of constructing classical conditional pivotal quantities for some special cases such as normal distribution. We extend the Sprott's work to the generalized form over the entire location-scale families. The approach provided in Sprott (2000, Chapter 7) is based on MLE only, without considering the case where the MLE does not exist (see Chapter 3). In this work, we propose a more general approach that is based on the MRE instead of the MLE. In addition, for the inference problem of the difference between the two location parameters, where the ratio of two scale parameters is unknown, we use another approach which is extended from the univariate case. These concepts will be discussed in Chapter 4.

In chapter 6, we evaluate our GPQ methods by using simulation methods. In most of examples the numerical results are close to the ideal levels, which show that our methods are consistent.

1.2. The mathematical statistics background

In this section, we define the concepts used in this thesis. In fact, most of the statistical definitions are based on some related mathematical concepts. Hence, it is convenience to provide first some related mathematical concepts to define the statistical concepts. These concepts are outlined here for the convenience of the reader. Nevertheless, for more details, the interested reader is referred to Billingsley (1995), Casella and Berger (2001), Lehmann and Casella (1998) and Schervish (1997) among others.

$\mathfrak{S}$-algebra, measure, dominating measure and measurable space. Let Ω be a sample space, i.e. a set of all possible outcomes from a random experiment. The concept of $\mathfrak{S}$-algebra is important in mathematical analysis and probability theory. Formally, the definition of $\mathfrak{S}$-algebra is given as follows.

Definition 1.1. A class of subsets of Ω, denoted by $\mathcal{B}$, is called a $\mathfrak{S}$-algebra (or $\mathfrak{S}$-field), if it satisfies the following properties:

a. $\emptyset \in \mathcal{B}$.

b. If $A \in \mathcal{B}$, then $A^c \in \mathcal{B}$.

c. If $A_1, A_2, ... \in \mathcal{B}$, then $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{B}$.

Based on $\mathfrak{S}$-algebra, measure and measurable space are defined as follows.

Definition 1.2. Let $\mathcal{B}$ be a $\mathfrak{S}$-algebra associated to the sample space Ω. Then, a nonnegative function λ defined on $\mathcal{B}$ is said to be a measure function if it satisfies the following properties:

a. $\lambda(\emptyset) = 0$.

b. Assume $A_1, A_2, ... \in \mathcal{B}$. If $\{A_i\}_{i=1}^{\infty}$ mutually disjoint, then

$$\lambda\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \lambda(A_i).$$

If λ is a measure, $(\Omega, \mathcal{B}, \lambda)$ is called measure space, and $(\Omega, \mathcal{B})$ is called measurable space.

Definition 1.3. A measure ν is said to be *absolutely continuous* with respect to λ if $\lambda(A) = 0$ implies $\nu(A) = 0$. Further, λ is so-called the *dominating measure*.

We consider a particular case of measure function that satisfies $\lambda(\Omega) = 1$. Then λ is said to be a probability function and $(\Omega, \mathcal{B}, \lambda)$ is said to be a probability space.

Measurable function. Another concept related to the definition of statistic is measurable function, which is defined in the following way.

Definition 1.4. Let Ω_1 and Ω_2 be nonempty sets. In addition, let $\mathcal{B}_1$ and $\mathcal{B}_2$ be $\mathfrak{S}$-algebras of subsets of Ω_1 and Ω_2 respectively. Then, a function $f : \Omega_1 \to \Omega_2$ is said to be measurable if

$$E \in \mathcal{B}_2 \Rightarrow f^{-1}(E) \equiv \{x \in \Omega_1 | f(x) \in E\} \in \mathcal{B}_1.$$

Statistic and estimator. When we apply the statistical distributions to model populations, we usually deal with a family of distributions rather than a single distribution. This family is indexed by one or more parameters, which allow us to vary certain characteristics of the distribution while the functional form remains fixed. For example, when we use the normal distribution to model a particular population, since we can not precisely specify the mean, we need to deal with a parametric family of normal distributions with mean μ, where μ is an unspecified parameter, $-\infty < \mu < \infty$. In this case, the normal distribution involving unknown μ is called normal family.

Based on the definitions given above, let $(\mathbb{R}^n, \mathcal{B})$ be a measurable space where $\mathcal{B}$ denote the $\mathfrak{S}$-algebra of $\mathbb{R}^n$. Further, let $(\mathcal{X}, \mathcal{A})$ be a measurable space. Then, a statistic is a measurable transformation T from the sample space $(\mathcal{X}, \mathcal{A})$ into a measurable space $(\mathbb{R}^n, \mathcal{B})$. In other words, a statistic is the result of applying a function (statistical algorithm) to a set of data.

A statistic is distinct from an unknown parameter, which is not computable from a sample. A key use of statistics is as estimators in statistical inference, to estimate parameters of a distribution based on a sample. For instance, the sample mean is a statistic, while the population mean is a parameter.

One can distinguish two main types of estimators: point estimators and interval estimators. Following Casella and Berger (2001, p. 311, 417), the definitions of these two types of estimators are respectively given as follows.

DEFINITION 1.5. A point estimator is any function $W(X_1, ..X_n)$ of a random sample; that is, any statistic is a point estimator.

To compare the difference between an estimate and an estimator, it is noticed that an estimator is a function of the sample, while an estimate is the realized value of an estimator (that is, a number) obtained when a sample is actually taken. For example, when a sample is given, an estimator is a function of the random variables $X_1, ..., X_n$, while an estimate is a function of the realized values $x_1, ..., x_n$.

DEFINITION 1.6. An interval estimate of a real-valued parameter θ is any pair of functions, $L(x)$ and $U(x)$, of a sample that satisfy $L(x) \leqslant U(x)$ for all $x \in \mathcal{X}$. If $X = x$ is observed, the inference $L(x) \leqslant \theta \leqslant U(x)$ is made. The random interval $[L(X), U(X)]$ is called an interval estimator.

Sufficient statistics and complete statistics. Sufficient statistics arise in nearly every aspect of statistical inference. For example, to start a statistical analysis such as parameter estimation, we usually select a random observable variable X, whose distribution depends on the parameter of interest. In this case, it usually turns out that there are some so-called sufficient statistics, which may capture all of the information about the parameter of interest. Any additional statistics besides these, carries no information about the parameter of interest. In this case, X can be reduced to or replaced by the sufficient statistics, without losing any information about the parameter of interest.

As described in Casella and Berger (2001, p. 272, 417), a sufficient statistic is formally defined as:

DEFINITION 1.7. A statistic $T(X)$ is a sufficient statistic for θ if the conditional distribution of the random sample X given the value of $T(X)$ does not depend on θ.

In addition, completeness, which is a property of a family of probability distributions, is closely related to statistical sufficiency and often occurs in conjunction with it. As discussed in Lehmann and Casella (1998 p. 42), since the complete sufficient statistics are particularly effective in reducing the data, there are many applications concerning this concept. For example, Basu's Theorem, used to prove Proposition 4.1 in this thesis, requires the related statistic to be sufficient and complete. The concept of completeness is defined as follows.

Definition 1.8. Let $\{f(t|\theta), \theta \in \Theta\}$ be a family of probability density functions (pdfs) or probability mass functions (pmfs) for a statistic $T(X)$. The family of probability distributions is called complete if for any measurable function g with $E_\theta g(T) = 0$ for all θ implies that $P_\theta(g(T) = 0) = 1$ for all θ. Equivalently, $T(X)$ is called complete statistic.

Ancillary Statistics. Ancillary statistics, is one of Fisher's most fundamental contributions to statistical inference (see David, 2003). In this thesis, the way of establishing the distributions of pivotal quantities is based on ancillary statistics. In general, the definition of ancillary statistic is given as follows.

Definition 1.9. A statistic $S(X)$ whose distribution does not depend on the parameter θ is called an ancillary statistic for θ.

In other words, an ancillary statistic contain no information about θ. In addition, it was pointed out by Fisher (1925) that although an ancillary statistic by itself fails to provide any information about the parameter, yet in conjunction with another statistic, typically the maximum likelihood estimator (MLE), it could provide valuable information about the parameter. For more detail related to above definitions, we refer to Casella and Berger (2001, p. 282).

Pivotal quantity. Pivotal quantity, is a quantity involving the data and the unknown parameter of interest, and this is mainly used for constructing confidence intervals. In this thesis, we consider the construction of generalized confidence interval based on generalized pivotal quantity, which will be discussed in Chapter 3. To this end, it is important to present the related concepts.

DEFINITION 1.10. Let θ be the parameter of interest and let $X \equiv (X_1, X_2, ..., X_n)$ be a random sample. A random variable $Z(X, \theta) = Z(X_1, ..., X_n, \theta)$ is a pivotal quantity for θ if it is a function of X and θ and the distribution of $Z(X, \theta)$ is independent of the parameter θ. That is, if $X \sim f(x|\theta)$, then $Z(X, \theta)$ has the same distribution no matter what the value of θ is (Casella and Berger 2001, p. 427).

Loss function and risk. In statistics, when we estimate θ by $T(x)$, we would like to evaluate the distance between $T(x)$ and the exact value of θ. A function of such distance is called loss function. A loss function is used to represent the loss associated with an estimate being "wrong" (different from either a desired or a true value) as a function of a measure of the difference between the estimated value and the true or desired value.

Definition 1.11. Let $X_1, ..., X_n$ be iid random sample from the population whose probability density function (pdf) or probability mass function (pmf) is $f(x|\theta)$, $\theta \in \Theta$. Further, let λ be the dominating measure and let $\mathcal{A}$ denote the set of allowable decisions (known as action space). Then, a nonnegative function $L(\theta, a)$ defined over $\Theta \times \mathcal{A}$, is called a loss function.

Furthermore, a loss function satisfies the definition of a random variable so one can establish a cumulative distribution function and an expected value. However, more commonly, the loss function is expressed as a function of some other random variable. The expected loss $R^*(\theta, \delta)$, also known as risk, is defined by

$$R^*(\theta, \delta) = E[L(\theta, \delta)|\theta] = \int_{\mathcal{X}} L(\theta, \delta(x)) f(x|\theta) d\lambda(x),$$

where $\lambda(x)$ is the related dominating measure.

Equivariance. The idea of equivariant estimator is based on the theory of invariant estimation. Thus, to describe the equivariant estimation, we should begin from the definition of invariant family of distributions and invariant estimator.

Also, to introduce the concept of invariant estimation, we need to define some other concepts such as group of transformation, and invariant problem. These definitions will be used as the backgrounds of invariance.

Definition 1.12. A nonempty set H together with a binary operation $\circ$ is called a group if it satisfies the following conditions:

1.(Closure): For all a, b in H, the result of $a \circ b$ is also in H.

2.(Associativity): For all a, b and c in H, we have $(a \circ b) \circ c = a \circ (b \circ c)$.

3.(Identity): There exists an identity element e in H such that for all a in H, we have $e \circ a = a \circ e = a$.

4.(Inverse): For each a in H, there exists an element a^{-1} in H such that $a \circ a^{-1} = a^{-1} \circ a = e$, where e is the identity element.

Definition 1.13. A set of functions $\{h(x) : h \in H\}$ from the sample space $\mathcal{X}$ onto $\mathcal{X}$ is called a group of transformations of $\mathcal{X}$ if

1.(Inverse): For every $h \in H$, there is a $h' \in H$ such that $h'(h(x)) = x$ for all $x \in \mathcal{X}$.

2.(Composition): For every $h \in H$ and $h' \in H$, there exists $h'' \in H$ such that $h'(h(x)) = h''(x)$ for all $x \in \mathcal{X}$.

3.(Identity): The identity, defined by $e(x) = x$ for all $x \in \mathcal{X}$, is an element of H.

Based on the concepts of group and group of transformation, the definitions of invariant distribution and invariant loss function are defined respectively as follows.

Definition 1.14. A family of distributions $\{f(x|\theta) : \theta \in \Theta\}$ is said to be invariant under the group H if for every $\theta \in \Theta$, and $h \in H$, there is a unique $\theta' \in \Theta$, such that $X \sim f(x|\theta)$ implies $Y = h(X) \sim f(y|\theta')$.

For a fixed $h \in H$, the correspondence that takes $\theta \to \theta'$ defines a function, which we denote by $\bar{h}(\theta) = \theta'$. Then the invariant loss function is defined as follows.

DEFINITION 1.15. Let $\{f(x|\theta) : \theta \in \Theta\}$ be invariant under group H and let $L(\theta, a)$ be a loss function on $\Theta \times \mathcal{A}$, where $\mathcal{A}$, the set of possible decisions, coincides with H. We say that the loss function is invariant under H if, for every $h \in H$ and $a \in \mathcal{A}$, there exists an $a^* \in \mathcal{A}$ such that for all $\theta \in \Theta$, $L(\theta, a) = L(\bar{h}(\theta), a^*)$, where $\bar{H} = \{\bar{h} : \bar{h} \in H\}$ is a group of transformations from Θ to itself.

Further, let $\tilde{H} = \{\tilde{h} : \tilde{h} \in H\}$ be a group of transformation from $\mathcal{A}$ to itself, and let $\tilde{h}(a) = a^*$, where $\tilde{h} \in \tilde{H}$. Then the invariant estimation problem and invariant estimator are respectively defined as follows.

DEFINITION 1.16. Let H be a group of transformations. An estimation problem is invariant under H if the family of distributions and the loss function are invariant.

DEFINITION 1.17. For an estimation problem that is invariant under H, a point estimator $S(X)$ of θ is an invariant estimator under the group H if for every $x \in \mathcal{X}$,

$\theta \in \Theta$, and $h \in H$, $S(h(x)) = S(x)$.

With above definitions, the definitions of equivariant estimator and minimum risk equivariant estimator, as defined in Lehmann (1998, p.161), are provided in the following way.

Definition 1.18. Let $\mathcal{A}$ be the set of possible decisions that coincides with H, and let $\tilde{h}$ be any one to one transformation from $\mathcal{A}$ to itself. Then in an invariant estimation problem, an estimator $S(X)$ is said to be equivariant if it satisfies

$$S(h(X)) = \tilde{h}(S(X)) \tag{1}$$

for all $h \in H$.

Definition 1.19. In an invariant estimation problem, if an equivariant estimator exists which minimizes the risk, it is called the minimum risk equivariant estimator (MRE).

For more details related to equivariant estimator, we refer to Lehmann and Casella (1998, Chapter 3).

Location and scale families. In probability theory, especially in the field of statistics, one of the well-known groups of transformations is location and scale family. There are three types of models included in location and scale families: location families, scale families, and location-scale families. Each of the families is constructed by specifying a single pdf, say $g(x)$, called the *standard probability density function(pdf)* for the family. Then all other pdfs in the family are generated by transforming the standard pdf in a prescribed way.

DEFINITION 1.20. Let $g(x)$ be any pdf. Then the family of pdfs $f(x|\theta) = g(x-\mu)$, indexed by the parameter $\theta = \mu$, $-\infty < \mu < \infty$, is called the *location family* with standard pdf $g(x)$ and μ is called the *location parameter*.

DEFINITION 1.21. Let $g(x)$ be any pdf. Then for any $\sigma > 0$, the family of pdfs $f(x|\theta) = (1/\sigma)g(x/\sigma)$, indexed by the parameter $\theta = \sigma$, is called the *scale family* with standard pdf $g(x)$ and σ is called the *scale parameter*.

DEFINITION 1.22. Let $g(x)$ be any pdf. Then for any $-\infty < \mu < \infty$ and $\sigma > 0$, the family of pdfs $f(x|\theta) = (1/\sigma)g(\frac{x-\mu}{\sigma})$, indexed by the parameter $\theta = (\mu, \sigma)$, is called the *location and scale family* with standard pdf $g(x)$ and μ and σ are called the *location parameter* and *scale parameter*, respectively.

To summarize the above definitions, note that the construction of location and scale families is related to the following theorem.

THEOREM 1.1. *Let $g(x)$ be any pdf and let μ and $\sigma > 0$ be any given constants. Then the function*

$$f(x|\mu, \sigma) = \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right)$$

is a pdf.

The proof of this theorem is given in the appendix.

CHAPTER 2

Generalized inference in univariate location-scale family

In this chapter, we discuss the concept of generalized pivotal quantity (GPQ), which is an approach for generating the generalized confidence interval. Also, we introduce the concept of generalized p-value (GPV) based on generalized test variable (GTV), which can be considered as a function of GPQ. Therefore, in order to compute the GCI and GPQ for a given inference problem for univariate location-scale family, it is important to construct the related GPQ first. In this case, we develop an approach for finding the required GPQ, which is based on the conditional pivotal quantity. Finally, we present the GPQ of some special location-scale families as examples.

2.1. Concepts of generalized inference

The concept of generalized P-value was first introduced by Tsui and Weerahandi (1989) whereas the generalized confidence interval was introduced by Weerahandi (1993). In the quoted papers, given the following two definitions are

DEFINITION 2.1. Let $X_1, ..., X_n$ be the iid observable random variables with pdf $f(x|\theta)$, where $\theta = (\theta_1, \theta_2)$ is a vector of unknown parameters. Here θ_1 is the parameter of interest and θ_2 is a vector of nuisance parameters. Let $\mathcal{X}$ denote the sample space of possible values of X, where $X = X_1, ..., X_n$, and let Θ denote the parameter space of θ and Θ_1 be the parameter space of θ_1. In addition, we denote x ($x \in \mathcal{X}$) as an observation from X. Let $R = R(X, x, \theta)$ be a function of X, x, θ, where $\theta = (\theta_1, \theta_2)$. Then, the function R is said to be a generalized pivotal quantity if it satisfies the following conditions .

1. Given x, the distribution of R is free from unknown parameters;

2. the observed pivotal, defined as $R_{obs} = R(x, x, \theta)$, does not depend on the nuisance parameter.

Then, for a given generalized pivotal quantity R, and a confidence coefficient γ, a $100\gamma\%$ generalized confidence interval for θ_1, say Θ_{θ_1}, as defined in Weerahandi (1993), is given by

$$\Theta_{\theta_1} = \{\theta_1 \in \Theta_1 | R_{obs} \in CI_{\theta_1}\}, \tag{2}$$

where the subset CI_{θ_1} is given by

$$Pr(R \in CI_{\theta_1}) = \gamma.$$

Especially, when $R_{obs} = R_(x, x, \theta) = \theta_1$, and $CI_{\theta_1} \subseteq \Theta_1$,

$$\Theta_{\theta_1} = \{\theta_1 \in \Theta_1 | \theta_1 \in CI_{\theta_1}\} = CI_{\theta_1}. \tag{3}$$

Besides, as introduced by Tsui and Weerahandi (1989), generalized test variable (GTV), which is used to compute the generalized p-value (GPV), is defined in the following way.

Definition 2.2. Let x be the observed value of the random vector X, and let $\theta = (\theta_1, \theta_2)$, where θ_1 is the parameter of interest, and θ_2 is a vector of nuisance parameters. Then, the generalized test variable, is defined as a function of (X, x, θ), say $T(X, x, \theta)$, which satisfies the following requirements.

(1). $T(x, x, \theta) = t$ is free of θ.

(2). For fixed x and θ, the distribution of $T(X, x, \theta)$ is free of the nuisance parameter θ_2.

(3). For fixed x and θ_2, $P[T(X, x, \theta) \geqslant t|\theta_1]$ is non-decreasing in θ_1.

For the first requirement, Tsui and Weerahandi (1989) points out that it can be considered as a redundant requirement, because if the function $T(X, x, \theta)$ we construct does not satisfy the first requirement, then we can define a new GTV $T'(X, x, \theta) = T(X, x, \theta) - T(x, x, \theta)$, which satisfies the first requirement. In addition, for the third requirement, Krishnamoorthy, Mathew, and Ramachandran (2007) gives a more general form:

(3). For fixed x and θ_2, $P[T(X, x, \theta) \geqslant t|\theta_1]$ is stochastically monotone in θ_1 (i.e. stochastically increasing or decreasing in θ_1).

In this thesis, we use the definition of Krishnamoorthy et al. (2007) instead of Tsui and Weerahandi (1989).

Now, suppose that we are interested in testing the hypotheses

$$H_0 : \theta_1 \geqslant \theta_0 \qquad v.s. \qquad H_1 : \theta_1 < \theta_0, \tag{4}$$

where θ_1 is the parameter of interest and θ_0 is a specified constant. Then, for fixed x and θ_2, as discussed in Krishnamoorthy et al. (2007), one can construct the GTV by $T_1(X, x, \theta) = R(X, x, \theta) - R(x, x, \theta)$, where $R(X, x, \theta)$ is the GPQ discussed in Definition 2.1 and $R(x, x, \theta)$ is the observed value of $R(X, x, \theta)$. In this case, it can be verified that $P[T(X, x, \theta) \geqslant t]$ is decreasing in θ_1. Then, the generalized p-value for (4) is given by

$$p = \sup_{H_0} P[T_1(X, x, \theta) \geqslant 0] = \sup_{H_0} P[R(X, x, \theta) - R(x, x, \theta) \geqslant 0]. \tag{5}$$

Especially, when $R_{obs} = \theta_1$,

$$p = P(R(X, x, \theta) \geqslant \theta_0). \tag{6}$$

From what we discussed above, it can be seen that the GTV is just a function of the GPQ, $R(X, x, \theta)$, since θ_0 is a known constant. Therefore, if we can construct the $R(X, x, \theta)$, it is easy to construct the GTV and compute the GPV by using (5). Because the distribution of $R(X, x, \theta)$ is free of any unknown parameters, the generalized p-value at θ_0 can be obtained by using a numerical method or estimated by using Monte Carlo simulation.

2.2. Generalized pivotal quantity for some location-scale families

As discussed in Subsection 2.1, most of difficulty in finding GCI and GPV for location-scale family is in constructing the related GPQ. To clarify the idea of the procedures of constructing GPQ for location-scale family, we first look at some well-known location-scale families as examples. In the sequel, we denote $\bar{X}$ and S_X^2 as the sample mean and the sample variance respectively.

1. Normal distribution. Let $X_1, X_2, ...X_n$ be iid with $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $i = 1, ..., n$, it can be verified that $\bar{X}, S_X^2$ are the uniformly minimum variance unbiased estimators (UMVUE) of μ, σ^2, respectively. Since

$$(\bar{X} - \mu)/(S_X/\sqrt{n}) \sim \mathcal{T}_{n-1} \qquad \text{and} \qquad (n-1)S_X^2/\sigma^2 \sim \mathcal{X}_{n-1}^2,$$

the functions $(\bar{X} - \mu)/(S_X/\sqrt{n})$ and $(n-1)S_X^2/\sigma^2$ are the pivotal quantities. Therefore, the generalized pivotal quantities are given by

$$\begin{aligned} R_\sigma &= s_{obs}^2((n-1)S_X^2/\sigma^2)^{-1}, \\ R_\mu &= \bar{x}_{obs} - (\bar{X} - \mu)/(S_X/\sqrt{n})\sqrt{s_{obs}^2}. \end{aligned}$$

2. Exponential distribution. Let $X_1, X_2, ...X_n$ be iid with $X_i \sim exp(\sigma)$, $i = 1, 2, ..., n$. Then it can be verified that $\bar{X}$ is the UMVUE of σ. Therefore, the classical pivotal quantity is given by $\bar{X}/\sigma$ and GPQ is $\bar{x}_{obs}(\bar{X}/\sigma)^{-1}$.

3. Gamma distribution. Let $X_1, X_2, ...X_n$ be iid with $X_i \sim Gamma(\alpha, \sigma)$, α is known.

It can be verified that, $\bar{X} \sim Gamma(n\alpha, \sigma/n)$. Therefore, $2n\bar{X}/\sigma \sim Gamma(n\alpha, 2)$. This implies that $2n\bar{X}/\sigma$ is the classical pivotal quantity of σ. Then, the GPQ of σ is $2n\bar{x}_{obs}(2n\bar{X}/\sigma)^{-1}$.

4. Cauchy distribution. Let $X_1; ...; X_n$ be iid sample from a Cauchy distribution with location parameter μ, then $\bar{X} - \mu$ follows the Cauchy distribution with location parameter 0. Further, it is shown in Hass, Bain and Antle (1970) that the pivotal quantities of Cauchy distribution with location parameter μ and scale parameter σ are $(\hat{\mu} - \mu)\sqrt{n}/\sigma$ and $\hat{\sigma}/\sigma$, then the generalized pivotal quantities are given by

$$\hat{\sigma}_{obs}(\hat{\sigma}/\sigma)^{-1}, \quad \text{and} \quad \hat{\mu}_{obs} - (\hat{\mu} - \mu)/\sigma \times \hat{\sigma}_{obs}(\hat{\sigma}/\sigma)^{-1},$$

where the $\hat{\mu}$ and $\hat{\sigma}$ are the maximum likelihood estimator (MLE) of μ, and σ, respectively. The procedures of finding these estimators are also given in Haas et al. (1970).

Based on the above examples, we provide a framework of constructing GPQ and GPV, and this will be discussed in the next section.

2.3. Generalized pivotal quantity in location-scale family

In Section 2.2, we have considered some well-known distributions and we have shown the approach of how the related generalized pivotal quantities are constructed. As one can see, most of the GPQ are related to the following pivotal quantities: $\hat{\mu} - \mu$, $\hat{\sigma}/\sigma$ and $(\hat{\mu} - \mu)/\sigma$, where $\hat{\mu}$ and $\hat{\sigma}$ are some estimators. In this case, it is reasonable to assume that for any location-scale family, there exist a general method of constructing the related GPQ by using the pivotal quantities mentioned above.

In fact, the applications of above pivotal quantities in location-scale family case are already discussed. For example, Lawless (1972) used those pivotal quantities to develop some conditional confidence interval procedures for the location and scale parameters of the Cauchy and logistic distributions. It should be noted that these procedures can also be applied to the general location-scale family where the MLE exists. In this case, we can construct the GPQ by extending the idea of his work.

However, since the purpose of Lawless (1972) is to deal with Cauchy and logistic distributions, where the MLE exists, the conditional pivotal quantities discussed in his paper is based on the MLE only. In this thesis, we are interested in extending these conditional pivotal quantities to the general location-scale families including the case where the MLE does not exist (see Chapter 3). Briefly, as one of the contributions in this thesis, we use the minimum risk equivariant estimator as a replacement of the MLE. This will be discussed in Chapter 3.

2.3.1. Conditional pivotal function in location-scale family.

Let $X_1, X_2, ...X_n$ be iid with pdf $f(x_i) = \sigma^{-1}g((x_i - \mu)/\sigma)$. Lawless (1972) presented the conditional pivotal function of location-scale family with 3 situations, that are: Location family, scale family and location-scale family.

2.3.1.1. *Location family.* In this subsection, we consider the case where μ is an unknown parameter while σ is supposed to be known. Thus, since σ is known, without loss of generality, we can let σ be equal to one for convenience. Then, the joint distribution of $X_1, X_2, ...X_n$ are given by

$$f(x_1, x_2, ...x_n|\mu) = \prod_{i=1}^{n} g(x_i - \mu).$$

The following lemma is useful in establishing the pivotal quantities for the location scale parameters. Recall that two random variables X and Y are said to be functionally independent if the only function $\phi(k_1, k_2)$ such that $\phi(X, Y) = 0$ is $\phi = 0$.

Lemma 2.1. *Let X and Y be independent nondegenerated random variables. Then X and Y are functionally independent.*

Proof. Let $\mathfrak{S}(X)$ and $\mathfrak{S}(Y)$ denote the sigma-fields generated by X and Y respectively. Then, X and Y are independent iff $\mathfrak{S}(X)$ and $\mathfrak{S}(Y)$ are independent. Further, let ϕ be a measurable function such that $\phi(X, Y) = 0$. This implies that ϕ is both $\mathfrak{S}(X)$-measurable and $\mathfrak{S}(Y)$-measurable. Then, since $\mathfrak{S}(X)$ and $\mathfrak{S}(Y)$ are independent, ϕ must be a constant function, and since $\phi(X,Y)=0$, we get $\phi=0$, and that completes the proof. □

PROPOSITION 2.1. *Let $T(X_1,...,X_n)$ be a nonconstant measurable function of $(X_1,...,X_n)$. Then $T(X_1,...,X_n)$ and $X_1,...,X_n$ are not functionally independent.*

PROOF. Let

$$\phi(k_1,...,k_{n+1}) = k_1 - T(k_2,...,k_{n+1}).$$

We have

$$\phi(T, X_1,...X_n) = T - T(X_1,...,X_n) = 0.$$

However, $\phi(k_1,...,k_{n+1})$ is a function other than $\phi=0$. This implies $T(X_1,...,X_n)$ and $X_1,...,X_n$ are not functionally independent.

□

PROPOSITION 2.2. *Let $T(X_1,...,X_n)$ be a measurable function of $(X_1,...,X_n)$. Then, $T(X_1,...,X_n)$ and $X_1,...,X_{n-1}$ are functionally independent.*

PROOF. **1**. If T is a constant, this does not depend on $X_1,...,X_n$:

Suppose T and $X_1,...,X_{n-1}$ are not functionally independent. There exists a function $\phi_1(k_1,...,k_n)$ other than $\phi_1=0$, but $\phi_1(T,X_1,...,X_{n-1})=0$.

Since T is a constant, $\phi_1(T,X_1,...,X_{n-1})$ can be considered as a function of $X_1,...,X_{n-1}$.

Set $\phi_2(k_2,...,k_n)=\phi_1(T,k_2,...,k_n)$. Then

$$\phi_2(X_1,...,X_{n-1}) = \phi_1(T,X_1,...,X_{n-1}) = 0.$$

This implies that there exists a function $\phi_2(k_1, ..., k_{n-1})$ other than $\phi_2 = 0$. However, since $\phi_2(X_1, ..., X_{n-1}) = 0$, we can verify that $X_1, ..., X_{n-1}$ are not functionally independent. This contrasts the fact that, by Lemma 2.1, $X_1, ..., X_{n-1}$ are functionally independent. Therefore, T and $X_1, ..., X_{n-1}$ are functionally independent.

2. If T is a nonconstant and measurable function of $X_1, ..., X_n$, suppose that T at least depends on one random variable X_j. By changing the order of $X_1, ..., X_n$, let X_n be the X_j. Then T can be expressed as a function of X_n and X^*, where X^* is a subset of $\{X_1, ..., X_{n-1}\}$.

Let $T = \tilde{T}(X^*, X_n)$. Suppose T and $X_1, ..., X_{n-1}$ are not functionally independent. There exists an $\phi_1(k_1, ..., k_n)$ other than $\phi_1 = 0$, but $\phi_1(T, X_1, ..., X_{n-1}) = 0$.

Since $T = \tilde{T}(X^*, X_n)$,

$$\phi_1(T, X_1, ..., X_{n-1}) = \phi_1(\tilde{T}(X^*, X_n), X_1, ..., X_{n-1}),$$

which is a function of $X_1, ..., X_{n-1}, X_n$. Set $\phi_2(k_1, ..., k_n) = \phi_1(\tilde{T}, k1, ..., k_{n-1})$, such that

$$\phi_2(X_1, ..., X_n) = \phi_1(\tilde{T}(X^*, X_n), X_1, ..., X_{n-1}).$$

This implies that there exist a function $\phi_2(k_1, ..., k_n)$ other than $\phi_2 = 0$. However, since $\phi_2(X_1, ..., X_n) = 0$, one can verify that $X_1, ..., X_n$ are not functionally independent, which contradicts the fact that, by Lemma 2.1, $X_1, ..., X_n$ are functionally independent.

Therefore, $T(X_1, ..., X_n)$ and $X_1, ..., X_{n-1}$ are functionally independent. □

Let $\hat{\mu}$ be the MLE or equivariant estimator of μ. In the sequel, we denote

$$a_i = X_i - \hat{\mu}, i = 1, ...n. \tag{7}$$

Note that, for the case of scale family, a_i is replaced by $b_i = X_i/\hat{\sigma}$, $i = 1, 2, ..., n$.

By using Proposition 2.1 and 2.2, we establish the following corollary and propositions:

COROLLARY 2.1. *Let $\hat{\mu}$ be a nonconstant estimator of μ. Then,*

(i): *$\hat{\mu}$ and $X_1, ..., X_n$ are not functionally independent.*

(ii): *$\hat{\mu}$ and $X_1, ..., X_{n-1}$ are functionally independent.*

PROOF. The statement (i) follows directly from Proposition 2.1, and the statement (ii) follows directly from Proposition 2.2. □

PROPOSITION 2.3. *Assume that relation (7) holds. Then $a_1, ..., a_n$ are not functionally independent.*

PROOF. Since $\hat{\mu}$ can be expressed as a function of $x_1, ..., x_n$, once we fix the value of $x_1, ..., x_{n-1}$, then the value of x_n will be fixed.

This implies that once we fix the values of $a_1 = x_1 - \hat{\mu}, ..., a_{n-1} = x_{n-1} - \hat{\mu}$, the value of $a_n = x_n - \hat{\mu}$ will be fixed.

Therefore, since $\hat{\mu}$ is equivariant, a_n can be expressed as a function of $a_1, ..., a_{n-1}$, we can set $a_n = T(a_1, ..., a_{n-1})$. There exists a function

$$\phi(k_1, ..., k_n) = k_n - T(k_1, ..., k_{n-1})$$

other than $\phi = 0$, but $\phi(a_1, ...a_{n-1}, a_n) = 0$.

This implies that $a_i = X_i - \hat{\mu}, i = 1, ..., n$ are not functionally independent. □

PROPOSITION 2.4. *Assume that relation (7) holds. Then $a_1, ..., a_{n-1}$ are functionally independent.*

PROOF. Suppose $a_1, ..., a_{n-1}$ are not functionally independent. There exists a function $\phi(k_1, ..., k_{n-1})$ other than $\phi = 0$, but $\phi(a_1, ...a_{n-1}) = 0$;

Since $a_i = X_i - \hat{\mu}, i = 1, ..., n-1$, we can verify that

$$\phi(a_1, ...a_{n-1}) = \phi(X_1 - \hat{\mu}, ..., X_{n-1} - \hat{\mu}),$$

which is a function of $\hat{\mu}$ and $X_1, ..., X_{n-1}$. Set

$$\phi_1(X_1, ..., X_{n-1}, \hat{\mu}) = \phi(X_1 - \hat{\mu}, ..., X_{n-1} - \hat{\mu}) = \phi(a_1, ...a_{n-1}).$$

Therefore, there exists a function $\phi_1(k_1, ..., k_n)$ other than $\phi_1 = 0$, with

$\phi_1(X_1, ..., X_{n-1}, \hat{\mu}) = 0$.

This implies that $\hat{\mu}$ and $X_1, ..., X_{n-1}$ are not functionally independent. But by Corollary 2.1, $\hat{\mu}$ and $X_1, ..., X_{n-1}$ are functionally independent, and that is a contradiction. Therefore, $a_i = X_i - \hat{\mu}, i = 1, ..., n-1$ are functionally independent, which completes the proof. □

PROPOSITION 2.5. *Assume that relation (7) holds. Then $a_1, ..., a_{n-1}$ are ancillary statistics for any location-scale family.*

PROOF. Define a_n by $X_n = a_n + \hat{\mu}$. Then, since $\hat{\mu}$ is an equivariant estimator of μ, a_n can be expressed as a function of $a_1, ..., a_{n-1}$ and $\hat{\mu}$. Set $a_n = T(a_1, ..., a_{n-1}, \hat{\mu})$.

Then

$$X_i = a_i + \hat{\mu}, i = 1, ..., n-1; X_n = a_n + \hat{\mu}.$$

Let $\tilde{X} = (X_1, ..., X_n)$ and let $\tilde{x} = (x_1, ..., x_n)$. We have $f(\tilde{x}) = \prod_{i=1}^{n} g(x_i - \mu)$.

Since $a_1, ..., a_{n-1}$ and $\hat{\mu}$ are functionally independent,

$$f(\tilde{a}, \hat{\mu}) = |J| \prod_{i=1}^{n} g(a_i + \hat{\mu} - \mu),$$

where $\tilde{a} = (a_1, ..., a_{n-1})$ and $|J|$ is the Jacobian matrix.

Then,

$$f(\tilde{a}) = \int_{-\infty}^{\infty} |J| \prod_{i=1}^{n} g(a_i + \hat{\mu} - \mu) d\hat{\mu} = |J| \int_{-\infty}^{\infty} \prod_{i=1}^{n} g(a_i + z) dz,$$

where $z = \hat{\mu} - \mu$ and

$$J = \begin{pmatrix} \frac{\partial X_1}{\partial a_1} & \cdots & \frac{\partial X_1}{\partial a_{n-1}} & \frac{\partial X_1}{\partial \hat{\mu}} \\ \vdots & \cdots & \vdots & \vdots \\ \frac{\partial X_n}{\partial a_1} & \cdots & \frac{\partial X_n}{\partial a_{n-1}} & \frac{\partial X_n}{\partial \hat{\mu}} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 0 & 1 \\ \vdots & \cdots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Since

$$f(\tilde{a}) = |J| \int_{-\infty}^{\infty} \prod_{i=1}^{n} g(a_i + z) dz,$$

by integrating out z eliminates μ, which implies that $f(\tilde{a})$ does not depend on μ. Therefore, $a_1, ..., a_{n-1}$ are ancillary statistics, which completes the proof. □

Since $a_i = X_i - \hat{\mu}$, we have $X_i = a_i + \hat{\mu}$. By applying the Jocobi's transformation, the joint pdf of $a_1, ..., a_{n-1}, \hat{\mu}$ are given by:

$$f(a_1, ..., a_{n-1}, \hat{\mu}) = |J| \prod_{i=1}^{n} g(a_i + \hat{\mu} - \mu).$$

Therefore,

$$f(\hat{\mu}|a_1,...,a_{n-1}) = \frac{f(a_1,...,a_n,\hat{\mu})}{f(a_1,...a_n)} = \frac{|J|\prod_{i=1}^{n} g(a_i+\hat{\mu}-\mu)}{|J|\int_{-\infty}^{\infty}\prod_{i=1}^{n} g(a_i+z)dz}.$$

Hence,

$$f(\hat{\mu}|a_1,...,a_{n-1}) = \frac{\prod_{i=1}^{n} g(a_i+\hat{\mu}-\mu)}{\int_{-\infty}^{\infty}\prod_{i=1}^{n} g(a_i+z)dz},$$

where the $|J|$ and z are defined in the proof of Proposition 2.5.

Therefore, it is clear that $z_1 = \hat{\mu}-\mu$ is the pivotal quantity, with conditional density:

$$f(z_1|a_1,...,a_{n-1}) = \frac{\prod_{i=1}^{n} g(a_i+z_1)}{\int_{-\infty}^{\infty}\prod_{i=1}^{n} g(a_i+z_1)dz_1}.$$

2.3.1.2. *Scale family.* In this subsection, we consider the case where σ is an unknown parameter, while μ is supposed to be known. Since the μ is known, without loss of generality, we can let μ equal to zero for convenience. Then, the joint distribution of $X_1, X_2, ...X_n$ are given by:

$$f(x_1,x_2,...,x_n|\sigma) = \sigma^{-n}\prod_{i=1}^{n} g(x_i/\sigma).$$

In this subsection, we assume that MLE of σ exist. The case where the MLE does not exist, is discussed in the Chapter 3. Let $\hat{\sigma}$ be the MLE of σ if it exists. Then, the quantities

$$b_i = X_i/\hat{\sigma}, i = 1, ...n \tag{8}$$

will also satisfy the Propositions 2.4 and 2.5 by replacing (7) with (8).

Then, similar to the location family, we find that conditional on $b_1, ..., b_n$, $Z_2 = \hat{\sigma}/\sigma$ is the pivotal quantity, with conditional density:

$$f(z_2|b_1, ..., b_{n-1}) = \frac{z_2^{n-1}\prod_{i=1}^{n} g(b_i z_2)}{\int_0^\infty z_2^{n-1}\prod_{i=1}^{n} g(b_i z_2) dz_2}.$$

2.3.1.3. *Location-scale family.* In this subsection, we consider the more general case where μ and σ are both unknown.

The joint pdf of $X_1, ..., X_n$ is

$$f(x_1, ..., x_n) = \sigma^{-n}\prod_{i=1}^{n} g\left(\frac{x_i - \mu}{\sigma}\right).$$

Here we assume that the MLEs for (μ, σ) exist. Thus, let $\hat{\mu}, \hat{\sigma}$ be the MLE of μ and σ, respectively. Once again, the case where MLE does not exist will be discussed in Chapter 3. Then the quantities $a_i = (x_i - \hat{\mu})/\hat{\sigma}, i = 1, ..., n-2$ are functionally independent ancillary statistics. Similar to the previous sections, we can verify that

$$Z_3 = (\hat{\mu} - \mu)/\hat{\sigma}, \quad Z_4 = (\hat{\sigma})/\sigma$$

are pivotal quantities, with joint conditional density:

$$f(z_3, z_4|a_1, ..., a_{n-2}) = \frac{z_4^{n-1}\prod_{i=1}^{n} g((z_3 + a_i)z_4)}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^{n} g((z_3 + a_i)z_4) dz_3 dz_4}. \quad (9)$$

In summary, the conditional pivotal functions of the 3 situations are given by:

1: μ unknown, σ known: $a_i = \frac{X_i - \hat{\mu}}{\sigma}$, where the $\hat{\mu}$ is the estimator of μ. The pivotal function of μ is $Z_1 = \hat{\mu} - \mu$.

2: μ known, σ unknown: $a_i = \frac{X_i - \mu}{\hat{\sigma}}$, where the $\hat{\sigma}$ is the estimator of σ. The pivotal function of σ is $Z_2 = \hat{\sigma}/\sigma$.

3: μ, σ both unknown: $a_i = \frac{X_i - \hat{\mu}}{\hat{\sigma}}$, the pivotal functions are $Z_3 = (\hat{\mu} - \mu)/\hat{\sigma}$, and $Z_4 = \hat{\sigma}/\sigma$.

2.3.2. Generalized pivotal quantity in location-scale family.

Based on the above classical pivotal functions, the generalized pivotal quantities are given by:

1: μ unknown, σ known: the generalized pivotal quantity of μ is

$$R_1 = \hat{\mu}_{obs} - (\hat{\mu} - \mu), \tag{10}$$

where the $\hat{\mu}$ is the MLE of μ.

2: μ known, σ unknown: the generalized pivotal quantity of σ is

$$R_2 = \hat{\sigma}_{obs}(\hat{\sigma}/\sigma)^{-1}, \tag{11}$$

where the $\hat{\sigma}$ is the MLE of σ.

3: μ, σ both unknown: the generalized pivotal quantities are

$$R_3 = \hat{\mu}_{obs} - \hat{\sigma}_{obs}(\hat{\mu} - \mu)/\hat{\sigma}, \qquad R_4 = \hat{\sigma}_{obs}(\hat{\sigma}/\sigma)^{-1}. \tag{12}$$

Again, it should be recalled that, here, we consider the case where the MLE of (μ, σ) exist. For the case where MLE does not exist, the GPQ in (10), (11), and (12) are applied by replacing MLE with equivariant estimators, which will be discussed in the next chapter.

2.4. Generalized confidence interval and P-value

With the same notations given in the previous sections, let $X_1, X_2, ...X_n$ be iid with pdf $f(x_i) = \sigma^{-1} g((x_i - \mu)/\sigma)$. Further, let $Z_3 = (\hat{\mu} - \mu)/\hat{\sigma}$, and let $Z_4 = \hat{\sigma}/\sigma$. As discussed above, the GPQ for location and scale parameters are respectively given by

$$R_3 = \hat{\mu}_{obs} - \hat{\sigma}_{obs} Z_3, \qquad R_4 = \hat{\sigma}_{obs} Z_4^{-1}, \tag{13}$$

where the distribution of Z_3 and Z_4 are discussed in Subsection 3.1.3.

For R_3 and R_4, it can be verified that their observed pivotal are μ and σ, respectively. In this case, as discussed in Section 2.1 of Chapter 2, the $100\gamma\%$ GCI of μ, say CI_μ is given by

$$Pr(R_3 \in CI_\mu) = \gamma, \tag{14}$$

and CI_σ, the $100\gamma\%$ GCI of σ, is

$$Pr(R_4 \in CI_\sigma) = \gamma. \tag{15}$$

Further, consider the testing problem

$$H_0 : \mu \geqslant \mu_0 \qquad v.s. \qquad H_1 : \mu < \mu_0, \tag{16}$$

based on what we discussed in subsection 2.1, the generalized p-value is given by

$$p_\mu \quad = \quad P(R_3 \geqslant \mu_0) = P(\hat{\mu}_{obs} - \hat{\sigma}_{obs} Z_3 \geqslant \mu_0). \tag{17}$$

Similarly, in the testing problem

$$H_0 : \sigma \geqslant \sigma_0 \qquad v.s. \qquad H_1 : \sigma < \sigma_0, \tag{18}$$

the generalized p-value is

$$p_\sigma \;=\; P(R_4 \geqslant \sigma_0) = P(\hat{\sigma}_{obs} Z_4^{-1} \geqslant \sigma_0). \tag{19}$$

CHAPTER 3

Equivariant method

It is well known that the maximum likelihood estimators (MLE) of the location and scale parameters may not exist. For example, as shown in Gupta and Székely(1994), if

$$g(x) = c(x \log^2 x)^{-1}$$

where $0 < x \leqslant k < 1$, k is any constant that satisfies $0 < k < 1$ and $c = -1/\log(k)$ is a constant, then the MLE's for the location and scale parameters for the location-scale family $\sigma^{-l} g((x-\mu)/\sigma)$ does not exist.

Another similar example is given in Pitman (1979), which shows that, if

$$g(x) = \frac{1}{2(1+|x|)(1+\log(1+|x|))^2},$$

where $-\infty < x < \infty$. Then the MLE's for the location and scale parameters for the probability density function $\sigma^{-l} g((x-\mu)/\sigma)$ does not exist. In this case, there is a need for finding other good estimators as replacement.

As discussed in Chapter 2, one solution for the case where the MLE does not exist is to use the equivariant estimator instead of MLE. In this section, we study the efficiency of equivariance method by discussing the concepts and applications of equivariant estimators particularly in location-scale family. In this case, first we present some concepts about equivariant estimator in location-scale family.

3.1. Equivariant point estimator

3.1.1. Equivariant estimator of the location parameter.

Let $X = (X_1, ..., X_n)$, where $X_1, ..., X_n$ are iid with joint pdf

$$f(x|\mu) = \prod_{i=1}^{n} g(x_i - \mu), -\infty < \mu < \infty,$$

here g is known and μ is an unknown location parameter. Here we consider the estimation problem of μ under the loss function $L(\mu, a)$ that satisfies $L(\mu, a) = L(\mu+c, a+c)$, where $a \in A$ and c is any constant. Also, consider the group of transformations

$$H = \{X' = X + c = (X_1 + c, ..., X_n + c), \qquad \mu' = \mu + c, \qquad a' = a + c\}. \tag{20}$$

Under the transformations (20), one can verify that:

$$f(x'|\mu') = f(x|\mu), \qquad L(\mu', a') = L(\mu, a)$$

for all $c, \mu \in R$, $x \in X$, and $a \in A$ and this implies that $f(x|\mu)$ and $L(\mu, a)$ are invariant under the group of transformations (20). Therefore, the problem of estimating μ is said to be an invariant estimation problem under the translation group.

In this case, an estimator $\hat{\mu}(X)$ is an equivariant estimator for μ under (20) if it satisfies

$$\hat{\mu}(X + c) = \hat{\mu}(X) + c, \tag{21}$$

for all c.

For a given location family, there may exist more than one estimator that satisfies (21). Thus, it is desirable to choose an equivariant estimator whose risk is minimal, so-called minimum risk equivariant estimator (MRE).

Under the square error loss function, Pitman (1939) developed a set of minimum risk equivariant estimators, known as Pitman estimators, for the location families.

In this thesis, we denote $\hat{\mu}_P$ the Pitman estimator for μ. Also, we denote $\hat{\mu}_M$ the MLE for μ. Similarly, the notation $\hat{\sigma}_P$ and $\hat{\sigma}_M$ are used in order to denote respectively the Pitman estimator and MLE for σ. As quoted from Schervish (1997, Chapter 6), the related expressions of MRE are presented as follows.

THEOREM 3.1. *Let $X_1, ..., X_n$ be the iid random sample from location family with pdf $f(x|\theta) = g(x - \mu)$, where μ is the unknown parameter. Then, under the loss function $L(\mu, a) = (\mu - a)^2$, the MRE of μ is given by:*

$$\hat{\mu}_P(x) = \frac{\displaystyle\int_{-\infty}^{\infty} t \prod_{i=1}^{n} g(x_i - t)dt}{\displaystyle\int_{-\infty}^{\infty} \prod_{i=1}^{n} g(x_i - t)dt}. \tag{22}$$

PROOF. Let $h(x) = h(x_1, ..., x_n) = (x_1 + \mu, ..., x_n + \mu) = x + \mu$, and $z = t - \mu$, we can verify that:

$$\hat{\mu}_P(h(x)) \;=\; \hat{\mu}_P(x_1 + \mu, ..., x_n + \mu) = \frac{\displaystyle\int_{-\infty}^{\infty} t \prod_{i=1}^{n} g(x_i + \mu - t)dt}{\displaystyle\int_{-\infty}^{\infty} \prod_{i=1}^{n} g(x_i + \mu - t)dt}.$$

Then

$$\hat{\mu}_P(h(x)) = \frac{\displaystyle\int_{-\infty}^{\infty}(z+\mu)\prod_{i=1}^{n}g(x_i - z)dz}{\displaystyle\int_{-\infty}^{\infty}\prod_{i=1}^{n}g(x_i - z)dz} = \hat{\mu}(x) + \mu = \bar{h}(\hat{\mu}(x)),$$

that is, $\hat{\mu}_P(h(X)) = \bar{h}(\hat{\mu}_P(X))$, which implies that $\hat{\mu}_P(X)$ is an equivariant estimator for μ.

In addition, it can be verified that (22) minimizes the mean squared-error function (see Schervish, 1997, p. 348, or Lehmann and Casella, 1998, p. 154). This implies that (22) is the MRE for location parameter μ. □

Note that the MRE given in (22) is referred in literature as Pitman estimator for location parameters (see Pitman, 1939). Here we present an example which illustrates the application of relationship (22). Also, by this example, we illustrate a relationship between Pitman estimator and MLE in the location family case. More precisely, in Example 3.1, Pitman estimator and MLE are the same.

Example 3.1. Let $X_1, ..., X_n$ iid with $X_i \sim \mathcal{N}(\mu, 1)$, where $-\infty < \mu < \infty$ is an unknown parameter. It can be verified that the MLE for μ is $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. In addition, the following proposition can be verified.

PROPOSITION 3.1. *Let $X_1, ..., X_n$ be a random sample as given in Example 4.1. Then, the Pitman estimator for μ is $\bar{X}$.*

PROOF. From (22), we can verify that:

$$\hat{\mu}_P(x) = \frac{\int_{-\infty}^{\infty} t \prod_{i=1}^{n} g(x_i - t)dt}{\int_{-\infty}^{\infty} \prod_{i=1}^{n} g(x_i - t)dt} = \frac{\int_{-\infty}^{\infty} \frac{t}{(2\pi)^{n/2}} \exp[-\frac{1}{2}\sum_{i=1}^{n}(x_i - t)^2]dt}{\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2}} \exp[-\frac{1}{2}\sum_{i=1}^{n}(x_i - t)^2]dt}$$

$$= \frac{\int_{-\infty}^{\infty} \frac{t}{(2\pi)^{n/2}} \exp[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \bar{x})^2 - \frac{n}{2}(\bar{x} - t)^2]dt}{\int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2}} \exp[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \bar{x})^2 - \frac{n}{2}(\bar{x} - t)^2]dt}.$$

Therefore,

$$\hat{\mu}_P(x) = \frac{\exp[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \bar{x})^2] \int_{-\infty}^{\infty} \frac{t}{(2\pi)^{n/2}} \exp[-\frac{n}{2}(\bar{x} - t)^2]dt}{\exp[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \bar{x})^2] \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{n/2}} \exp[-\frac{n}{2}(\bar{x} - t)^2]dt} = \frac{\bar{x}}{1} = \bar{x}.$$

□

In summary, for the normal location family case, the Pitman estimator and MLE are same for the location parameter. In general, the following propositions make a connection between equivariant estimator and MLE.

PROPOSITION 3.2. *Let μ be the location parameter and suppose that $\hat{\mu}_M$, the MLE of μ, exists. Then, $\hat{\mu}_M$ is equivariant.*

The details of the proof can be found in Lehmann and Casella (1998, p.150). In other words, in location family, the MRE performs always better than MLE. It should be noted that similar result holds for the scale family case, which will be discussed in the next subsection.

3.1.2. Equivariant estimator of the scale parameter.

Let $X = (X_1, ..., X_n)$, where $X_1, ..., X_n$ are iid with joint pdf

$$f(x|\sigma) = \frac{1}{\sigma^n}\prod_{i=1}^{n} g(x_i/\sigma), 0 < \sigma < \infty,$$

here g is known and σ is an unknown scale parameter. Consider the estimation problem of σ under the loss function $L(\sigma, a)$, which satisfies $L(\sigma, a) = L(c\sigma, ca)$, where $a \in A$ and c is any constant. For example, $L(\sigma, a) = (a-\sigma)^2/\sigma^2$ is a legitimate such loss function.

Further, consider the group of transformations:

$$H = \{X' = cX = (cX_1, ..., cX_n), \qquad \sigma' = c\sigma, \qquad a' = ca\}. \tag{23}$$

We can verify that

$$f(x'|\sigma') = f(x|\sigma), \qquad L(\sigma', a') = L(\sigma, a)$$

for all c, σ, x, and $a \in A$.

Therefore, this estimation problem is invariant and any estimator $\hat{\sigma}(X)$ that satisfies

$$\hat{\sigma}(X') = c\hat{\sigma}(X),$$

for all $c > 0$, is said to be an equivariant estimator for σ under the group of transformations in (23).

Once again, we would like to choose an equivariant estimator whose risk is minimal, that is the MRE.

THEOREM 3.2 (Pitman, 1939). *Let $X_1, ..., X_n$ be independent random sample from scale family with pdf $f(x|\theta) = \sigma^{-1}g(x/\sigma)$, where σ is the unknown parameter. Then, under the loss function $L(\sigma, a) = (a-\sigma)^2/\sigma^2$, the MRE, so-called Pitman estimator of σ, is given by:*

$$\hat{\sigma}_P(X) = \frac{\displaystyle\int_0^\infty t^{-n-2}\prod_{i=1}^{n} g(x_i/t)dt}{\displaystyle\int_0^\infty t^{-n-3}\prod_{i=1}^{n} g(x_i/t)dt} \tag{24}$$

PROOF. Similar to the proof of Pitman estimator for location, we can verify that $\hat{\sigma}_P(h(X)) = \bar{h}(\hat{\sigma}_P(X))$, where $h(X) = X/\sigma$.

This implies that $\hat{\sigma}_P$ is an equivariant estimator for σ. Furthermore, it can be verified that (24) minimizes the risk under the loss function $L(\sigma, a) = (\sigma - a)^2/\sigma^2$. For more details, we refer to Schervish (1997, p. 352). This implies that the estimator $\hat{\sigma}_P(X)$ is the MRE for scale parameter σ. □

In addition, the connection between MRE and MLE in scale family is provided by the following proposition.

PROPOSITION 3.3. *Let σ be the scale parameter and suppose that $\hat{\sigma}_M$, the MLE of σ exists. Then, $\hat{\sigma}_M$ is also equivariant.*

For details, we refer to Lehmann and Casella (1998, p. 168). In addition, recall that MRE is the equivariant estimator that minimizes the risk. In other words, in scale family case, if MRE exists, it is always better than MLE (with respect to the loss function in Theorem 3.2). This relationship is shown in the following example, in which we illustrate the application of relation (24) for the normal scale family case.

Example 3.2. Let $X_1, ..., X_n$ iid with $X_i \sim \mathcal{N}(0, \sigma^2)$, where $0 < \sigma < \infty$ is an unknown parameter. It can be verified that the MLE for σ is $\sqrt{\frac{1}{n}\sum_{i=1}^{n} X_i^2}$. In addition, the Pitman estimator for σ is given by:

$$\hat{\sigma}_P(x) = \frac{\int_0^\infty t^{-n-2}\prod_{i=1}^{n} g(x_i/t)dt}{\int_0^\infty t^{-n-3}\prod_{i=1}^{n} g(x_i/t)dt} = \frac{\int_0^\infty \frac{t^{-n-2}}{(2\pi)^{n/2}}\exp[-\frac{1}{2t^2}\sum_{i=1}^{n} x_i^2]dt}{\int_0^\infty \frac{t^{-n-3}}{(2\pi)^{n/2}}\exp[-\frac{1}{2t^2}\sum_{i=1}^{n} x_i^2]dt}.$$

Let $u = \frac{s}{t^2}$, where $s = \sum_{i=1}^{n} x_i^2$. We have,

$$\hat{\sigma}_P(x) = \frac{\int_0^\infty s^{-(n+1)/2}u^{(n-1)/2}\exp\left[-\frac{u}{2}\right]du}{\int_0^\infty s^{-(n+2)/2}u^{n/2}\exp\left[-\frac{u}{2}\right]du}.$$

After some computations, we get

$$\hat{\sigma}_P(x) = \frac{\sqrt{s}\Gamma(\frac{n+1}{2})\int_0^\infty \frac{(u/2)^{(n+1)/2-1}}{\Gamma(\frac{n+1}{2})}\exp\left[-\frac{u}{2}\right]du}{\sqrt{2}\Gamma(\frac{n+2}{2})\int_0^\infty \frac{u^{(n+2)/2-1}}{\Gamma(\frac{n+2}{2})}\exp\left[-\frac{u}{2}\right]du} = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n+2}{2})}\sqrt{\frac{s}{2}}. \tag{25}$$

This implies that the Pitman estimator is unequal to the MLE for scale parameter of the normal distribution. However, in agreement with Proposition 3.4, MLE of σ is an equivariant estimator of σ, but the MRE should be better than MLE. To verify this result, we use the following proposition and corollary which prove that the mean square error (MSE) of MRE is less than that of MLE.

PROPOSITION 3.4. *Let $X_1, ..., X_n$ be iid from a scale family whose scale parameter is σ, and suppose that MLE of σ, $\hat{\sigma}_M(X)$ exists. Further, let $\hat{\sigma}_P(X)$ be the MRE of σ as given by Theorem 3.2. Then, under the loss function of Theorem 3.2, the MSE of $\hat{\sigma}_P(X)$ is less than the MSE of $\hat{\sigma}_M(X)$.*

PROOF. From Proposition 3.3, we evaluated that $\hat{\sigma}_M(X)$, the MLE of σ is equivariant. In addition, since $\hat{\sigma}_P(X)$, the Pitman estimator for σ, is the MRE, with the loss function $L(\sigma, a) = (\sigma - a)^2/\sigma^2$, we have,

$$R(\sigma, \hat{\sigma}_P(X)) \leqslant R(\sigma, \hat{\sigma}_M(X)), \tag{26}$$

where

$$R(\sigma, \delta^*(X)) = \int_{\mathcal{X}} \frac{(\sigma - \delta^*(x))^2 f(x|\sigma)}{\sigma^2} dx$$

is the risk function of $\delta^*(X)$. In addition, since $\sigma > 0$, from (26) we have

$$\int_{\mathcal{X}} \frac{(\sigma - \hat{\sigma}_P(x))^2 f(x|\sigma)}{\sigma^2} dx \leqslant \int_{\mathcal{X}} \frac{(\sigma - \hat{\sigma}_M(x))^2 f(x|\sigma)}{\sigma^2} dx,$$

and then,

$$\int_{\mathcal{X}} (\sigma - \hat{\sigma}_P(x))^2 f(x|\sigma) dx \leqslant \int_{\mathcal{X}} (\sigma - \hat{\sigma}_M(x))^2 f(x|\sigma) dx.$$

That gives

$$E\left[(\sigma - \hat{\sigma}_P(x))^2\right] \leqslant E\left[(\sigma - \hat{\sigma}_M(x))^2\right].$$

Therefore,

$$MSE(\hat{\sigma}_P(X)) \leqslant MSE(\hat{\sigma}_M(X)),$$

i.e., the MSE of $\hat{\sigma}_P(X)$ is less than the MSE of $\hat{\sigma}_M(X)$. □

Based on Proposition 3.4, we establish the following corollary.

Corollary 3.1. *Let $X_1, ..., X_n$ be iid with $X_i \sim \mathcal{N}(0, \sigma^2), i = 1, 2, ..., n$. Then, the MSE of the MRE $\hat{\sigma}_P(X)$, is less than the MSE of the MLE $\hat{\sigma}_M(X)$.*

In order to illustrate numerically this theoretical result, in Chapter 5 we use the simulation method to evaluate the efficiency of Pitman estimator. Briefly, as given in Chapter 5, the simulation study and numerical results confirm the above theoretical results.

3.1.3. Equivariant estimator of the location and the scale parameters.

Let $X = (X_1, ..., X_n)$, where $X_1, ..., X_n$ are iid with joint pdf

$$f(x|, \mu, \sigma) = \frac{1}{\sigma^n} \prod_{i=1}^{n} g\left(\frac{x_i - \mu}{\sigma}\right), 0 < \sigma < \infty, -\infty < \mu < \infty,$$

where g is known and (μ, σ) are unknown location and scale parameters. First, consider the estimation problem of σ under the loss function $L(\sigma, a)$, which satisfies $L(\sigma, a) = L(c\sigma, ca)$, where $a \in A$ and c is any constant.

By using arguments similar to those in the previous sections, one can verify that the problem remains invariant under the group of transformations

$$H = \{X' = b + cX = (b + cX_1, ..., b + cX_n), \; \sigma' = c\sigma, \quad a' = ca, \quad \mu' = b + c\mu\}. \quad (27)$$

Then, it can be verified that any estimator $\hat{\sigma}(X)$ that satisfies

$$\hat{\sigma}(b + cX) = c\hat{\sigma}(X)$$

is said to be an equivariant estimator for σ under the group of transformations in (27). Secondly, we consider the estimation problem of μ. The transformations in (27) relating to the sample space and parameter space remain the same, but the transformations of the decision space now become $a' = b + ca$. Similarly, one can verify that the problem remains invariant if the associated loss function is given by

$$L(\mu, \sigma, a) = \rho\left(\frac{a - \mu}{\sigma}\right), \tag{28}$$

where ρ is any function. For example, $\rho\left((a - \mu)/\sigma\right) = (a - \mu)^2/\sigma^2$.

Then, any estimator $\hat{\mu}(X)$ is said to be equivariant for μ if it satisfies

$$\hat{\mu}(b + cX) = b + c\hat{\mu}(X).$$

Of course, for reasons similar to those in the previous subsections, it is of interest to choose an equivariant estimator whose risk is minimal. The following theorem provides the MRE formula for μ and σ, respectively.

THEOREM 3.3. *Let $X_1, X_2, ..., X_n$ be iid random sample from location-scale family with pdf $f(x|\theta) = 1/\sigma g((x - \mu)/\sigma)$, where μ and σ are the unknown parameters. Then, under the loss function (28), the MRE (Pitman estimator) of μ is given by:*

$$\hat{\mu}_P(x) = \frac{\displaystyle\int_0^\infty \int_{-\infty}^\infty \frac{u}{v^{n+3}} \prod_{i=1}^n g\left(\frac{x_i - u}{v}\right) dudv}{\displaystyle\int_0^\infty \int_{-\infty}^\infty \frac{1}{v^{n+3}} \prod_{i=1}^n g\left(\frac{x_i - u}{v}\right) dudv}. \quad (29)$$

In addition, the MRE (Pitman estimator) of σ is:

$$\hat{\sigma}_P(x) = \frac{\displaystyle\int_0^\infty v^{-n-2} \int_{-\infty}^\infty \prod_{i=1}^n g\left(\frac{x_i - u}{v}\right) dudv}{\displaystyle\int_0^\infty v^{-n-3} \int_{-\infty}^\infty \prod_{i=1}^n g\left(\frac{x_i - u}{v}\right) dudv}. \quad (30)$$

The proof is similar to that given in Theorem 3.1 and 3.2. For more details, the reader is referred to Schervish (1997, Chapter 6).

Example 3.3. Let $X_1, ..., X_n$ be iid with $X_i \sim \mathcal{N}(\mu, \sigma)$, where $-\infty < \mu < \infty$ and $0 < \sigma < \infty$ are both unknown. By (29), one can verify that the Pitman estimator of μ is $\bar{X}$, which is the same as MLE. Also, by (30) the Pitman estimator of σ is given by

$$\hat{\sigma}_P(X) = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n+1}{2})} \sqrt{\frac{1}{2} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

Again, here we should note that this Pitman estimator is different from MLE, which is given by

$$\hat{\sigma}_M(X) = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

In addition, from Proposition 3.4, it can be verified that the MSE of Pitman estimator is less than the MSE of MLE. This theoretical result is confirmed by simulated results presented in Chapter 5.

From Example 3.1, 3.2 and 3.3, we find that the Pitman estimator is better than the

MLE since the MSE of Pitman estimator is less than the MSE of the MLE where it exists. Furthermore, even in the case where MLE does not exist, Pitman estimation method can also be applied. To illustrate this point of view, we apply the Pitman estimation to what we mentioned in the beginning of this chapter (see Chapter 5), the 2 examples in which the MLE of location and scale parameters did not exist. This highlights the efficiency of Pitman estimator in location and scale family. The simulation studies and numerical results are presented in Chapter 5.

3.2. GPQ based on equivariant estimator

In Chapter 2, we discuss the approaches for constructing different MLE based GPQ, which can be used to compute GCI and GPV. However, since it is already pointed out in this section that the MRE performs better than MLE, we would like to evaluate the approaches of GPQ based on MRE, that is, for (10), (11), (12), (17) and (19), we replace the MLE by Pitman estimator, respectively. The evaluations are made by applying these approaches to some particular problems.

Since the GPV and GCI in location-scale family is more important than the other two types of families, which do not include any nuisance parameter, here we only study the GCI and GPV in location-scale family case.

3.2.1. Evaluation of performances.

For the following examples, we only discuss the theoretical approach. Further, the numerical simulation results are presented in Chapter 5.

Example 3.4. Normal distribution with unknown location and scale parameters. Let $X_1, X_2, ...X_n$ iid $\mathcal{N}(\mu, \sigma^2)$, where $-\infty < \mu < \infty$, and $0 < \sigma < \infty$ are unknown parameters. In addition, let $a_i = (x_i - \hat{\mu})/\hat{\sigma}$, where $\hat{\mu} = \bar{X}$ and $\hat{\sigma}$ is chosen from 2 different estimators: MLE ($\hat{\sigma}_M(X)$) and Pitman estimator, ($\hat{\sigma}_P(X)$), respectively. It was shown in Chapter 2 that conditional on $a = (a_1, a_2, ...a_{n-2})$, $Z_3 = (\hat{\mu} - \mu)/\hat{\sigma}$, $Z_4 = \hat{\sigma}/\sigma$ are the pivotal quantities. Further, from (9), the joint conditional pdf is:

$$f(z_3, z_4|a) = \frac{z_4^{n-1}\prod_{i=1}^{n} g((z_3 + a_i)z_4)}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^{n} g((z_3 + a_i)z_4) dz_3 dz_4}. \tag{31}$$

More precisely, the following proposition helps to simplify the pdf of Z_3 and Z_4. The proof of this proposition is done by applying the Basu's Theorem (see Casella and Berger, 2001, p. 287 or Lehmann and Casella, 1998, p. 42). Alternatively, another proof based on direct calculation and transformation is given in the appendix.

PROPOSITION 3.5. *Let* $\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$ *and* $S^2 = \sum_{i=1}^{n}(a_i - \bar{a})^2$. *If* $X_1, ...X_n$ *are iid* $\mathcal{N}(\mu, \sigma^2)$, *then,*

(i):

$$\frac{\sqrt{n(n-1)}Z_3}{S}\Big| a_1, a_2, ...a_n \sim \mathcal{T}_{n-1}.$$

(ii):

$$S^2 Z_4^2 | a \sim \mathcal{X}_{n-1}^2.$$

Proof. It can be verified that

$$\frac{\sqrt{n(n-1)}Z_3}{S} = \frac{\sqrt{n}(\bar{X}-\mu)}{S_X}, \tag{32}$$

where $S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$. Further, it is well known that

$$\frac{\sqrt{n}(\bar{X}-\mu)}{S_X} \sim \mathcal{T}_{n-1} \tag{33}$$

In addition, it can be verified that $(\bar{X}, S_X^2)$ is a complete sufficient statistic for (μ, σ^2); while $(a_1, ..., a_n)$ is an ancillary statistic for (μ, σ^2). Then, by Basu's Theorem, $\frac{\sqrt{n}(\bar{X}-\mu)}{S_X}$ and $(a_1, ..., a_n)$ are independent. Therefore, from (32) and (33),

$$\frac{\sqrt{n(n-1)}Z_3}{S} | a \sim \mathcal{T}_{n-1}$$

(ii). Similar to (i), one can prove that $S^2 Z_4^2 | a \sim \mathcal{X}_{n-1}^2$. □

Therefore, the generalized pivotal quantities of μ and σ are given by:

$$\begin{aligned} R_\mu &= \hat{\mu}_{obs} - \hat{\sigma}_{obs}\left(\frac{\hat{\mu}-\mu}{\hat{\sigma}}\right) \\ &= \hat{\mu}_{obs} - \hat{\sigma}_{obs}\frac{S_{obs}}{\sqrt{n(n-1)}}\left(\frac{\sqrt{n(n-1)}\left(\frac{\hat{\mu}-\mu}{\hat{\sigma}}\right)}{S}\right) \\ &= \hat{\mu}_{obs} - \hat{\sigma}_{obs}\left(\frac{S_{obs}}{\sqrt{n(n-1)}}\left(\mathcal{T}_{n-1}\right)\right), \end{aligned}$$

and

$$R_\sigma = \hat{\sigma}_{obs}(\hat{\sigma}/\sigma)^{-1} = S_{obs}\hat{\sigma}_{obs}(S^2\hat{\sigma}^2/\sigma^2)^{-\frac{1}{2}} = \frac{S_{obs}\hat{\sigma}_{obs}}{\sqrt{\mathcal{X}_{n-1}^2}}.$$

Example 3.5. Cauchy distribution with unknown location and scale parameters. For Cauchy distribution and Logistic distribution, which is discussed in the next example, Lawless (1972) already evaluated the efficiency of MLE. Therefore, here we only study the case of Pitman estimator.

Let $X_1, ...X_n$ iid $Cauchy(\mu, \sigma^2)$, where $-\infty < \mu < \infty$, and $0 < \sigma < \infty$ are unknown parameters. Accordingly, let $a_i = (x_i - \hat{\mu})/\hat{\sigma}$, where $\hat{\mu}$ and $\hat{\sigma}$ are respectively given by $\hat{\mu}_P(X)$, and $\hat{\sigma}_P(X)$, the Pitman estimator for location and scale parameters. It is shown in Chapter 2 that

$$R_3 = \hat{\mu}_{obs} - \hat{\sigma}_{obs} Z_3, \qquad R_4 = \hat{\sigma}_{obs} Z_4^{-1}, \tag{34}$$

where $Z_3 = (\hat{\mu} - \mu)/\hat{\sigma}$, and $Z_4 = \hat{\sigma}/\sigma$, are the generalized pivotal quantities. By using (9), the joint conditional density is given by:

$$f(z_3, z_4|a) = \frac{z_4^{n-1} \prod_{i=1}^{n} g((z_3 + a_i) z_4)}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1} \prod_{i=1}^{n} g((z_3 + a_i) z_4) dz_3 dz_4},$$

where $g(x) = \dfrac{1}{\pi} \dfrac{1}{1 + x^2}$. Therefore, the conditional probability density function of Z_3 given a is:

$$f(z_3|a) = \frac{\int_0^\infty z_4^{n-1} \prod_{i=1}^{n} g((z_3 + a_i) z_4) dz_4}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1} \prod_{i=1}^{n} g((z_3 + a_i) z_4) dz_3 dz_4},$$

and then,

$$f(z_3|a) = \frac{\displaystyle\int_0^\infty z_4^{n-1}\prod_{i=1}^n \frac{1}{\pi}\frac{1}{1+((z_3+a_i)z_4)^2}dz_4}{\displaystyle\int_0^\infty\int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^n \frac{1}{\pi}\frac{1}{1+((z_3+a_i)z_4)^2}dz_3dz_4}. \tag{35}$$

Similarly, the conditional pdf of Z_4 given a is:

$$f(z_4|a) = \frac{\displaystyle\int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^n g((z_3+a_i)z_4)dz_3}{\displaystyle\int_0^\infty\int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^n g((z_3+a_i)z_4)dz_3dz_4},$$

that is,

$$f(z_4|a) = \frac{\displaystyle\int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^n \frac{1}{\pi}\frac{1}{1+((z_3+a_i)z_4)^2}dz_3}{\displaystyle\int_0^\infty\int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^n \frac{1}{\pi}\frac{1}{1+((z_3+a_i)z_4)^2}dz_3dz_4}. \tag{36}$$

Example 3.6. Logistic distribution with unknown location and scale parameters. Let $X_1, ...X_n$ iid $Logistic(\mu, \sigma^2)$, where $-\infty < \mu < \infty$, and $0 < \sigma < \infty$ are unknown parameters. Then, similar as Example 3.5, let $a_i = (x_i - \hat{\mu})/\hat{\sigma}$, where $\hat{\mu}$ and $\hat{\sigma}$ are respectively given by $\hat{\mu}_P(X)$, and $\hat{\sigma}_P(X)$, the Pitman estimator for location and scale parameters. Then,

$$R_3 = \hat{\mu}_{obs} - \hat{\sigma}_{obs}Z_3, \qquad R_4 = \hat{\sigma}_{obs}Z_4^{-1}, \tag{37}$$

where $Z_3 = (\hat{\mu} - \mu)/\hat{\sigma}$, and $Z_4 = \hat{\sigma}/\sigma$, are the generalized pivotal quantities. From (9), the joint conditional density of (Z_3, Z_4) is given by:

$$f(z_3, z_4|a) = \frac{z_4^{n-1}\prod_{i=1}^{n} g((z_3+a_i)z_4)}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^{n} g((z_3+a_i)z_4)dz_3dz_4},$$

where $g(x) = \frac{e^x}{(1+e^x)^2}$. Then,

$$f(z_3|a) = \frac{\int_0^\infty z_4^{n-1}\prod_{i=1}^{n}\frac{1}{(1+\exp((z_3+a_i)z_4))^2}dz_4}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^{n}\frac{1}{(1+\exp((z_3+a_i)z_4))^2}dz_3dz_4}. \tag{38}$$

and

$$f(z_4|a) = \frac{\int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^{n}\frac{1}{(1+\exp((z_3+a_i)z_4))^2}dz_3}{\int_0^\infty \int_{-\infty}^\infty z_4^{n-1}\prod_{i=1}^{n}\frac{1}{(1+\exp((z_3+a_i)z_4))^2}dz_3dz_4}. \tag{39}$$

From the previous examples and their simulated results, which are given in Chapter 5, one can see that the GPQ approaches based on Pitman estimator perform very well. In addition, for the case where MLE does not exist, the approaches based on Pitman estimators still provide satisfactory results. The related numerical examples are presented in Chapter 5.

CHAPTER 4

Generalized inference in bivariate location-scale family

In the previous chapters, we discussed the GPQ method in location-scale family for one-sample case. It is noticed that the GPQ was derived from conditional pivotal quantities. Then based on these results we present the pivotal quantities in location-scale family. However, it should be noted that the motivation of GPQ is to solve some complex inference problems involving nuisance parameters, which may not be solved by using classical inference methods. In location family or scale family case, there is no any nuisance parameter and hence, the GPQ should provide similar result as that provided by the classical inference method. From this point of view, in this chapter we only study the problems related to location-scale family. Nevertheless, the problems related to location family only or scale family only can be verified in a similar way.

4.1. Description of the problems

Let $X = (X_1, ..., X_n)$, $Y = (Y_1, ..., Y_m)$ be iid with the distributions of X_i and Y_j given by

$$\begin{aligned} f_X(x_i|\mu_1, \sigma_1) &= \frac{1}{\sigma_1} g_1\left(\frac{x_i - \mu_1}{\sigma_1}\right), \qquad i = 1, .., n; \\ f_Y(y_j|\mu_2, \sigma_2) &= \frac{1}{\sigma_2} g_2\left(\frac{y_j - \mu_2}{\sigma_2}\right), \qquad j = 1, ..., m; \end{aligned}$$

where $\mu_1, \mu_2, \sigma_1, \sigma_2$ are all unknown and g_1, g_2 are the given probability density functions. Then, based on the above conditions, our interest is to make inference about the following parameters, respectively.

1. The ratio of the scale parameters $\rho = \frac{\sigma_2}{\sigma_1}$.
2. The difference between the location parameters $\delta = \mu_1 - \mu_2$, with known ρ.
3. The difference between the location parameters $\delta = \mu_1 - \mu_2$, with unknown ρ.

Then, extending the GPQ methods of the univariate case, we develop the following 2 approaches for the above bivariate problems. Briefly, for the inference problems of ρ and δ with known ρ, we use the first approach, which is extended from Sprott (2000). In addition, for the inference problem of δ with unknown ρ, here we provide an alternative approach, which will be discussed in this chapter.

4.2. Generalized pivotal quantity in bivariate case (first approach)

Since $X = (X_1, ...X_n)$ and $Y = (Y_1, ..., Y_m)$ are independent, the joint pdf of $X_1, ..., X_n, Y_1, ..., Y_m$ is

$$f_{XY}(x_1, ..., x_n, y_1, ..., y_m) = \frac{1}{\sigma_1^n \sigma_2^m} \prod_{i=1}^{n} g_1 \left(\frac{x_i - \mu_1}{\sigma_1} \right) \prod_{j=1}^{m} g_2 \left(\frac{y_j - \mu_2}{\sigma_2} \right).$$

Let $z_3 = (t_1, t_2)$, where $t_1 = (\hat{\mu_1} - \mu_1)/\hat{\sigma_1}$, $t_2 = (\hat{\mu_2} - \mu_2)/\hat{\sigma_2}$, and $z_4 = (t_3, t_4)$, where $t_3 = \hat{\sigma_1}/\sigma_1$, $t_4 = \hat{\sigma_2}/\sigma_2$. In addition, let $(a, b) = \{a, b_1, ..., b_{m-2}\}$, where

$$a_i = \frac{X_i - \hat{\mu_1}}{\hat{\sigma_1}}, \qquad \text{and} \qquad b_j = \frac{Y_j - \hat{\mu_2}}{\hat{\sigma_2}},$$

$i = 1, ..., n,\ j = 1, ..., m$. Then, as in the previous sections, it can be verified that the join pdf of Z_3, Z_4 conditional on (a, b),

$$f(z_3, z_4|a, b) \quad = \quad f(t_1, t_2, t_3, t_4|a, b),$$

is given by

$$f(z_3, z_4|a, b) = Ct_3^{n-1}t_4^{m-1}\prod_{i=1}^{n} g_1((t_1 + a_i)t_3)\prod_{j=1}^{m} g_2((t_2 + b_j)t_4), \tag{40}$$

where

$$C = \left(\int_0^\infty \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty t_3^{n-1}t_4^{m-1}\prod_{i=1}^{n} g_1((t_1 + a_i)t_3)\prod_{j=1}^{m} g_2((t_2 + b_j)t_4)dt_1dt_2dt_3dt_4\right)^{-1}.$$

Commonly, $(\hat{\mu}_i, \hat{\sigma}_j)$ are the MLEs. However, it should be noted that in some special bivariate cases, the MLEs do not exist. In this case, with the same procedures as discussed in Chapter 3 and 4, we can use the Pitman estimator instead. Further, it can be verified that in bivariate case, the MSE of Pitman estimator is less than the MSE of the MLE. Then, here we only consider the Pitman estimator if the estimation is required. For example, we let $\hat{\mu}_i$, $\hat{\sigma}_j$ denote the Pitman estimators of μ_i and σ_j, respectively.

In addition, for a discussion of the above pivotal quantities based on MLE, we refer to Sprott (2000, Chapter 7), in which the author provides similar results, but based on the MLE only, without considering the case when the MLE does not exist.

Furthermore, Sprott (2000, Chapter 7) uses (40) based on MLE to construct the classical conditional pivotal quantities of some special cases such as the normal distribution. We extend the approach of Sprott (2000, Chapter 7) to the generalized form over the class of location-scale families. That is, in any bivariate location-scale family, the GPQ constructed based on (40) can be applied for the problems considered here.

4.2.1. Inference problem for the ratio of scale parameters $\rho = \frac{\sigma_2}{\sigma_1}$.

For this problem, let

$$v = \frac{t_4}{t_3} = \frac{\hat{\sigma_2}/\sigma_2}{\hat{\sigma_1}/\sigma_1} = \frac{\hat{\sigma_2}/\hat{\sigma_1}}{\rho}. \tag{41}$$

By using Jacobian method, one can transform (40) to the joint pdf of T_1, T_2 and V. That is,

$$\begin{aligned} f_1(t_1, t_2, t_3, v|a, b) &= f(t_1, t_2, t_3, vt_3|a, b)|t_3| \\ &= Cv^{m-1} t_3^{n+m-1} \prod_{i=1}^{n} g_1((t_1 + a_i)t_3) \prod_{j=1}^{m} g_2((t_2 + b_j)vt_4). \end{aligned} \tag{42}$$

Then, the pdf of v can be computed by taking the marginal pdf. That is,

$$f_V(v|a, b) = CI_V(v) = \frac{I_V(v)}{\int_0^\infty I_V(v)dv}, \tag{43}$$

where

$$I_V(v) = v^{m-1} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty t_3^{n+m-1} \prod_{i=1}^{n} g_1((t_1 + a_i)t_3) \prod_{j=1}^{m} g_2((t_2 + b_j)vt_3) dt_1 dt_2 dt_3,$$

and

$$C = \left(\int_0^\infty I_V(v)dv \right)^{-1}.$$

From the above equation, one can see that the pdf of v does not dependent on ρ. This implies that v can be considered as a pivotal quantity of ρ.

From the pivotal quantity v, the constructed generalized pivotal quantity is

$$R_{\rho 1} = (\hat{\sigma}_{2obs}/\hat{\sigma}_{1obs}) \left(\frac{\hat{\sigma}_2/\hat{\sigma}_1}{\rho}\right)^{-1} = \frac{\hat{\sigma}_{2obs}}{\hat{\sigma}_{1obs}} v^{-1}. \tag{44}$$

Based on (44), consider the testing problem

$$H_0 : \rho \geqslant \rho_0 \qquad v.s. \qquad H_1 : \rho < \rho_0. \tag{45}$$

Similar as the results presented in Chapter 2, the generalized p-value is given by

$$p_\rho \;\; = \;\; P(R_{\rho 1} \geqslant \rho_0) = P\left(\frac{\hat{\sigma}_{2obs}}{\hat{\sigma}_{1obs}} v^{-1} \geqslant \rho_0\right). \tag{46}$$

In the following subsections, we consider inference problem for the difference between location parameters. To this end, let $\delta = \mu_1 - \mu_2$. For clarity sake, we present first case, where the parameter ρ is known and secondly we deal with the case where ρ is unknown.

4.2.2. Inference for the difference of location parameters $\delta = \mu_1 - \mu_2$ (ρ known).

Consider the inference problem concerning the difference between location parameters, $\delta = \mu_1 - \mu_2$, with known ρ. In this case, since δ contains two parameters of interest μ_1 and μ_2, which are included in t_1 and t_2, respectively, it is convenient to make a connection between the pivotal quantity of δ and t_1, t_2. Let

$$d = t_1 - t_2 \rho v = t_1 - t_2 \frac{\hat{\sigma}_2}{\hat{\sigma}_1} = \frac{\hat{\mu}_1 - \hat{\mu}_2 - \delta}{\hat{\sigma}_1}. \tag{47}$$

In addition, let

$$u = t_1 \rho v + t_2. \tag{48}$$

Then,

$$t_1(u,d) \quad = \quad (\rho v u + d)\left(1+\rho^2 v^2\right)^{-1} \text{ and } \quad t_2(u,d) = (u - \rho v d)\left(1+\rho^2 v^2\right)^{-1}.$$

This makes a transformation. In this case, since ρ is known, the joint pdf of u, d, v can be obtained from (42). That is,

$$f_2(u,d,t_3,v|ab) \quad = \quad f_1(t_1(u,d), t_2(u,d), t_3, v|ab)|J|,$$

that gives

$$\begin{aligned} f_2(u,d,t_3,v|ab) = C|J|v^{m-1}t_3^{n+m-1}\prod_{i=1}^{n} g_1\left(\left((\rho v u + d)\left(1+\rho v\right)^{-1} + a_i\right)t_3\right) \\ \times \prod_{j=1}^{m} g_2\left(\left((u-\rho v d)\left(1+\rho^2 v^2\right)^{-1} + b_j\right) v t_3\right) \end{aligned} \tag{49}$$

where J is the Jacobian matrix. One can verify that

$$|J| = \left(1+\rho^2 v^2\right)^{-1}.$$

An alternative construction of d, as suggested by Sprott (2000, Chapter 7), is given by

$$d \quad = \quad \frac{\hat{\mu_1} - \hat{\mu_2} - \delta}{\sqrt{\hat{\sigma_1}^2 + \hat{\sigma_2}^2}} = \frac{\dfrac{\hat{\mu_1} - \mu_1}{\hat{\sigma_1}} - \dfrac{\hat{\mu_2} - \mu_2}{\hat{\sigma_2}}\rho v}{\sqrt{1+\rho^2 v^2}} = \frac{t_1 - t_2 \rho v}{\sqrt{1+\rho^2 v^2}}, \tag{50}$$

and

$$u \quad = \quad \frac{\dfrac{\hat{\mu_1} - \mu_1}{\hat{\sigma_1}^2} + \dfrac{\hat{\mu_2} - \mu_2}{\hat{\sigma_2}^2}}{\sqrt{\dfrac{1}{\hat{\sigma_1}^2} + \dfrac{1}{\hat{\sigma_2}^2}}} = \frac{\dfrac{\hat{\mu_1} - \mu_1}{\hat{\sigma_1}}\rho v + \dfrac{\hat{\mu_2} - \mu_2}{\hat{\sigma_2}}}{\sqrt{1+\rho^2 v^2}} = \frac{t_1 \rho v + t_2}{\sqrt{1+\rho^2 v^2}}. \tag{51}$$

These also make a transformation, where

$$\begin{aligned} t_1(u,d) \quad &= \quad (\rho v u + d)\left(\sqrt{1+\rho^2 v^2}\right)^{-1}, \\ t_2(u,d) \quad &= \quad (u - \rho v d)\left(\sqrt{1+\rho^2 v^2}\right)^{-1}, \end{aligned}$$

and the Jacobian matrix satisfies $|J| = 1$. In this case, the joint pdf of u, d, v is given by

$$f_2(u,d,t_3,v|a,b) \quad = \quad f_1(t_1(u,d), t_2(u,d), t_3, v|a,b)|J|$$

and so,

$$f_2(u,d,t_3,v|a,b) = Cv^{m-1}t_3^{n+m-1}\prod_{i=1}^{n} g_1\left(\left((\rho vu+d)\left(\sqrt{1+\rho^2v^2}\right)^{-1}+a_i\right)t_3\right) \times\prod_{j=1}^{m} g_2\left(\left((u-\rho vd)\left(\sqrt{1+\rho^2v^2}\right)^{-1}+b_j\right)vt_3\right). \quad (52)$$

Since the Jacobian is 1, with u and d given by (51) and (50), respectively, in the sequel, we use (51), and (50), instead of (48) and (47) respectively.

Since ρ is known, (41) does not contain any unknown parameter. Furthermore, notice that $\hat{\sigma_1}$ and $\hat{\sigma_2}$ only depend on a and b, respectively. In this case, for given a and b, v is fixed. Therefore, instead of the joint pdf (52), here one should use the conditional pdf

$$\begin{aligned} f_3(u,d,t_3|v,a,b) &= \frac{f_2(u,d,t_3,v|a,b)}{f_V(v|a,b)} \\ &= f_2(u,d,t_3,v|a,b)\left(\int_{-\infty}^{\infty}\int_{0}^{\infty}\int_{-\infty}^{\infty} f_2(u,x,t_3,v|a,b)dudt_3dx\right)^{-1}, \end{aligned}$$

that gives

$$f_3(u,d,t_3|v,a,b) = \frac{f_2(u,d,t_3,v|a,b)}{\displaystyle\int_{-\infty}^{\infty} I_D(x)dx},$$

where

$$I_D(d) = v^{m-1}\int_0^{\infty}\int_{-\infty}^{\infty} t_3^{n+m-1}\prod_{i=1}^{n} g_1\left(\left((\rho vu+d)\left(\sqrt{1+\rho^2v^2}\right)^{-1}+a_i\right)t_3\right) \times\prod_{j=1}^{m} g_2\left(\left((u-\rho vd)\left(\sqrt{1+\rho^2v^2}\right)^{-1}+b_j\right)vt_3\right) dudt_3.$$

Then, by integrating with respect to u and t_3, respectively, the conditional pdf of d is found to be

$$f(d|v,a,b) = \frac{I_D(d)}{\int_{-\infty}^{\infty} I_D(x)dx}, \tag{53}$$

which does not contain any unknown parameter. This implies that d is a pivotal quantity of δ.

Similarly, the GPQ of δ is given by

$$R_{\delta 1} = \hat{\mu_1}_{obs} - \hat{\mu_2}_{obs} - \sqrt{\hat{\sigma_1}^2_{obs} + \hat{\sigma_2}^2_{obs}} \left(\frac{\hat{\mu_1} - \hat{\mu_2} - \delta}{\sqrt{\hat{\sigma_1}^2 + \hat{\sigma_2}^2}} \right),$$

and finally,

$$R_{\delta 1} = \hat{\mu_1}_{obs} - \hat{\mu_2}_{obs} - d\sqrt{\hat{\sigma_1}^2_{obs} + \hat{\sigma_2}^2_{obs}}. \tag{54}$$

Then, the generalized p-value for the testing problem

$$H_0 : \delta \geqslant \delta_0 \qquad v.s. \qquad H_1 : \delta < \delta_0, \tag{55}$$

is given by

$$p_\delta = P(R_{\delta 1} \geqslant \delta_0). \tag{56}$$

However, if ρ is unknown, as one can see, the pdf (52) includes ρ. This implies that d is not any longer a pivotal quantity since the related pdf contains unknown parameter ρ. Therefore, the first approach is inappropriate for the bivariate case where ρ is unknown. In this case, we introduce another approach, which is discussed in the following subsection.

4.3. Generalized pivotal quantity in bivariate case (second approach)

Let $Z_{3i} = (\hat{\mu}_i - \mu_i)/\hat{\sigma}_i$, $Z_{4i} = \hat{\sigma}_i/\sigma_i$, $i = 1, 2$. Since $(X_1, ...X_n)$ and $(Y_1, ..., Y_m)$ are independent, as discussed in Chapter 2, one can develop the following generalized pivotal quantities for μ_1, μ_2, σ_1, σ_2, respectively.

$$\begin{aligned} R_{31} &= \hat{\mu}_{1obs} - \sigma_{1obs} Z_{31}, \\ R_{32} &= \hat{\mu}_{2obs} - \sigma_{2obs} Z_{32}, \\ R_{41} &= \hat{\sigma}_{1obs} Z_{41}^{-1}, \\ R_{42} &= \hat{\sigma}_{2obs} Z_{42}^{-1}, \end{aligned}$$

where the distributions of Z_{3i}, Z_{4i} are discussed in Chapter 2. In this case, let

$$R_{\delta 2} = R_{31} - R_{41} = \hat{\mu}_{1obs} - \hat{\mu}_{2obs} - \sigma_{1obs} Z_{31} + \sigma_{2obs} Z_{32}. \tag{57}$$

It can be verified that the distribution of $R_{\delta 2}$ is free of any unknown parameters. Further, when $\hat{\mu}_{iobs} = \hat{\mu}_i$, and $\hat{\sigma}_{iobs} = \hat{\sigma}_i$, $i = 1, 2$, $R_{\delta 2}$ reduces to δ. These imply that $R_{\delta 2}$ is a generalized pivotal quantity for δ.

In addition,by using the similar way, one can verify that

$$R_{\rho 2} = R_{42}/R_{41} = \frac{\hat{\sigma}_{2obs} Z_{42}^{-1}}{\hat{\sigma}_{1obs} Z_{41}^{-1}} = \frac{\hat{\sigma}_{2obs} Z_{41}}{\hat{\sigma}_{1obs} Z_{42}}. \tag{58}$$

is another generalized pivotal quantity for ρ.

For (58) and (57), since each of them contains 2 independent distributions, sometimes it is hard to compute the confidence bounds and GPV by using the classical cumulative probability approach. In this case, it is convenient to use Monte Carlo estimation approach instead. For the related algorithm and discussions, the interested readers can refer to Krishnamoorthy, K. and Mathew, Thomas. (2003).

Based on the above GPQs, the GPV for testing (55) is

$$p_\delta \quad = \quad P(R_{\delta 2} \geqslant \delta_0). \tag{59}$$

Besides, in addition to (46), one can develop an alternative GPV for testing (45), which is given by

$$p_\rho \quad = \quad P(R_{\rho 2} \geqslant \rho_0). \tag{60}$$

4.4. Example 4.1

To illustrate the above procedures, here we carry out some examples in which the distributions of X_i and Y_j are normal. In other words, we assume

$$\begin{aligned} f_X(x_i|\mu_1,\sigma_1) &= \frac{1}{\sigma_1} g_1\left(\frac{x_i-\mu_1}{\sigma_1}\right) = \frac{1}{\sigma_1\sqrt{2\pi}} \exp\left[-\frac{(x_i-\mu_1)^2}{2\sigma_1^2}\right], \\ f_Y(y_j|\mu_2,\sigma_2) &= \frac{1}{\sigma_2} g_2\left(\frac{y_j-\mu_2}{\sigma_2}\right) = \frac{1}{\sigma_2\sqrt{2\pi}} \exp\left[-\frac{(y_j-\mu_2)^2}{2\sigma_2^2}\right], \end{aligned} \tag{61}$$

where $i = 1, .., n$ and $j = 1, ..., m$. Under the above assumption, we can apply the methods provided in this chapter to construct the GPQ for each of the bivariate problems of interest. Furthermore, we also provide methods for computing GCI and GPV, based on the GPQ constructed.

Construction of GPQ.

1. First approach. Under the above assumptions, we have the following proposition.

PROPOSITION 4.1. *If g_1 and g_2 satisfy (61), respectively. Then,*

(i):

$$\frac{(n-1)\sum b_j^2 v^2}{(m-1)\sum a_i^2} \sim \mathcal{F}(m-1, n-1),$$

where $\mathcal{F}(m-1, n-1)$ stands for Fisher distribution with $m-1$ and $n-1$ degrees of freedom.

(ii): *When ρ is known,*

$$\left(\frac{mn(m+n-2)v^2(1+\rho^2 v^2)}{(n\rho^2 v^2 + mv^2)(\sum a_i^2 + v^2 \sum b_j^2)}\right)^{\frac{1}{2}} d \sim \mathcal{T}_{m+n-2}.$$

PROOF. Let g_1 and g_2 satisfy (61). Then, for (43),

$$I_V(v) = C_V v^{m-1},$$

where

$$\begin{aligned} C_V &= \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{t_3^{m+n-1}}{\sqrt{2\pi}^{m+n}} \exp\left[-\frac{t_3^2}{2}\sum_{i=1}^n (t_1+a_i)^2 - \frac{v^2 t_3^2}{2}\sum_{j=1}^m (t_2+b_j)^2\right] dt_1 dt_2 dt_3 \\ &= \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty C_1 t_3^{m+n-1} \exp\left[-\frac{t_3^2}{2}\sum_{i=1}^n (t_1+a_i)^2 - \frac{v^2 t_3^2}{2}\sum_{j=1}^m (t_2+b_j)^2\right] dt_1 dt_2 dt_3, \end{aligned}$$

and $C_1 = \left(\sqrt{2\pi}\right)^{-m-n}$ is a component which does not contain v, t_3, t_2, t_1. Let

$$I_{V1}(v) = \frac{I_V(v)}{C_1}.$$

Then,

$$\begin{aligned} f(v|a,b) &= I_V(v)\left(\int_0^\infty I_V(v)dv\right)^{-1} = I_{V1}(v)C_1\left(\int_0^\infty I_{V1}(v)C_1 dv\right)^{-1} \\ &= I_{V1}(v)\left(\int_0^\infty I_{V1}(v)dv\right)^{-1}. \end{aligned}$$

Similarly, let $t_{11} = \sqrt{n} t_1 t_3$ and $t_{21} = \sqrt{m} v t_2 t_3$. It can be verified that

$$
\begin{aligned}
I_{V1}(v) &= v^{m-1} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty t_3^{m+n-1} \\
&\quad \times \exp\left(-\frac{t_3^2}{2} \sum_{i=1}^n (t_1 + a_i)^2 - \frac{v^2 t_3^2}{2} \sum_{j=1}^m (t_2 + b_j)^2 \right) dt_1 dt_2 dt_3 \\
&= v^{m-1} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty t_3^{m+n-1} \\
&\quad \times \exp\left(-\frac{t_3^2}{2} \left(\sum_{i=1}^n a_i^2 + v^2 \sum_{j=1}^m b_j^2 \right) - \frac{n t_1^2 t_3^2}{2} - \frac{m v^2 t_2^2 t_3^2}{2} \right) dt_1 dt_2 dt_3,
\end{aligned}
$$

and then,

$$
\begin{aligned}
I_{V1}(v) = v^{m-1} \int_0^\infty \int_{-\infty}^\infty C_2 t_3^{m+n-2} \exp\left(-\frac{t_3^2}{2} \left(\sum_{i=1}^n a_i^2 + v^2 \sum_{j=1}^m b_j^2 \right) - \frac{m v^2 t_2^2 t_3^2}{2} \right) dt_2 dt_3 \\
\times \left[\int_{-\infty}^\infty \frac{\exp(-t_{11}^2/2)}{\sqrt{2\pi}} dt_{11} \right].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
I_{V1}(v) = v^{m-2} \int_0^\infty C_3 t_3^{m+n-3} \exp\left(-\frac{t_3^2}{2} \left(\sum_{i=1}^n a_i^2 + v^2 \sum_{j=1}^m b_j^2 \right) \right) dt_3 \\
\times \left[\int_{-\infty}^\infty \frac{\exp(-t_{21}^2/2)}{\sqrt{2\pi}} dt_{21} \right].
\end{aligned}
$$

Then, by integrating with respect to t_{21}, one can verify that

$$
\begin{aligned}
I_{V1} = C_4 v^{m-2} \left(\sum_{i=1}^n a_i^2 + v^2 \sum_{j=1}^m b_j^2 \right)^{-\frac{m+n-2}{2}} \left(\frac{1}{2} \left(\sum_{i=1}^n a_i^2 + v^2 \sum_{j=1}^m b_j^2 \right) \right)^{\frac{m+n-2}{2}} \\
\times \int_0^\infty \frac{(t_3^2)^{\frac{m+n-4}{2}}}{\Gamma(\frac{m+n-2}{2})} \exp\left(-\frac{t_3^2}{2} \left(\sum_{i=1}^n a_i^2 + v^2 \sum_{j=1}^m b_j^2 \right) \right) dt_3^2,
\end{aligned}
$$

which gives

$$
I_{V1} = C_4 v^{m-2} \left(\frac{1}{2} \sum_{i=1}^n a_i^2 + \frac{1}{2} v^2 \sum_{j=1}^m b_j^2 \right)^{-\frac{m+n-2}{2}},
$$

where C_2, C_3 and C_4 are given in the following way:

$$\begin{aligned} C_2 &= \sqrt{n}^{-1}\sqrt{2\pi}C_1, \\ C_3 &= \sqrt{m}^{-1}\sqrt{2\pi}C_2, \\ C_4 &= 2^{\frac{m+n-4}{2}}\Gamma\left(\frac{m+n-2}{2}\right)C_3. \end{aligned}$$

It can be seen that the quantities C_2, C_3 and C_4 do not contain v, t_3, t_2, t_1. Let

$$\bar{a^2} = \sum_{i=1}^{n} a_i^2/(n-1), \quad \text{and} \quad \bar{b^2} = \sum_{i=1}^{m} b_j^2/(m-1),$$

In this case,

$$I_{V1}(v) = C_4 v^{m-2}\left((n-1)\bar{a^2} + (m-1)\bar{b^2}v^2\right)^{-\frac{m+n-2}{2}},$$

and therefore, by letting $I_{V2}(v) = \frac{I_{V1}(v)}{C_4}$, it can be verified that

$$f(v|a,b) = I_{V1}(v)\left(\int_0^\infty I_{V1}(v)dv\right)^{-1} = I_{V2}(v)\left(\int_v I_{V2}(v)dv\right)^{-1}.$$

Let $v_1 = \bar{b^2}v^2/\bar{a^2}$, after some computations, we find

$$f_{v_1}(v_1) = \frac{v_1^{\frac{m-1}{2}-1}(n-1+(m-1)v_1)^{-\frac{m+n-2}{2}}}{\int_0^\infty v_1^{\frac{m-1}{2}-1}(n-1+(m-1)v_1)^{-\frac{m+n-2}{2}}dv_1},$$

and then,

$$f_{v_1}(v_1) = \frac{1}{\mathcal{B}(\frac{m-1}{2}, \frac{n-1}{2})}v_1^{\frac{m-1}{2}-1}\left(n-1+(m-1)v_1\right)^{-\frac{m+n-2}{2}},$$

which implies that

$$\frac{(n-1)\sum b_j^2 v^2}{(m-1)\sum a_i^2} \sim \mathcal{F}(m-1, n-1).$$

In addition, it can be verified that when ρ is known, (53) can be transformed to

$$C_5\int_0^\infty\int_{-\infty}^\infty t_3^{n+m-1}\exp\left[-\frac{1}{2}\sum_{i=1}^n\left((\rho vu+d)\left(\sqrt{1+\rho^2v^2}\right)^{-1}+a_i\right)^2 t_3^2\right]$$
$$\times\exp\left[-\frac{1}{2}\sum_{j=1}^m\left((u-\rho vd)\left(\sqrt{1+\rho^2v^2}\right)^{-1}+b_j\right)^2 v^2t_3^2\right]dudt_3,$$

which is equivalent to

$$C_5\int_0^\infty\int_{-\infty}^\infty t_3^{n+m-1}\exp\left[-\frac{1}{2}\left(t_3^2\left(\sum a_i^2+\sum b_j^2v^2\right)\right)\right]$$
$$\times\exp\left[-\frac{1}{2}\left(t_3^2\frac{mnv^2d^2\left(1+\rho^2v^2\right)}{(n\rho^2v^2+mv^2)}+o^2\right)\right]dudt_3,$$

where

$$o^2=\frac{(u(n\rho^2v^2+mv^2)+(nd-mv^2d)\rho v)^2}{(1+\rho^2v^2)(n\rho^2v^2+mv^2)}.$$

By integrating $f(d|v,a,b)$ with respected to u, t_3, respectively, one can verify that

$$f(d|v,a,b) \;=\; C_6\left[\sum_{i=1}^n a_i^2+\sum_{j=1}^m b_j^2v^2+\frac{mnv^2d^2(1+\rho^2v^2)}{(n\rho^2v^2+mv^2)}\right]^{-\frac{1}{2}(m+n-1)}.$$

For more details, the reader is referred to the Appendix A.1.3.

Let

$$t=\left[\frac{mn(m+n-2)v^2(1+\rho^2v^2)}{(n\rho^2v^2+mv^2)(\sum a_i^2+v^2\sum b_j^2)}\right]^{\frac{1}{2}}d.$$

Then, we have

$$f(t|v,a,b) \;=\; \frac{[1+t^2/(m+n-2)]^{-\frac{1}{2}(m+n-2+1)}}{\sqrt{n+m-2}B\left(1/2,\frac{m+n-2}{2}\right)}.$$

so that

$$\left(\frac{mn(m+n-2)v^2(1+\rho^2v^2)}{(n\rho^2v^2+mv^2)(\sum a_i^2+v^2\sum b_j^2)}\right)^{\frac{1}{2}}d\sim\mathcal{T}_{m+n-2}.$$

□

By Proposition 4.1, (44) and (54) can be transformed to

$$R_{\rho 1} = \frac{\hat{\sigma}_{2obs}}{\hat{\sigma}_{1obs}} v^{-1} = \frac{\hat{\sigma}_{2obs}\sqrt{(n-1)\sum {b_{obs}}_j^2}}{\hat{\sigma}_{1obs}\sqrt{(m-1)\sum {a_{obs}}_i^2}} \left(\frac{(n-1)\sum b_j^2 v^2}{(m-1)\sum a_i^2}\right)^{-1/2},$$

hence,

$$R_{\rho 1} = \frac{\hat{\sigma}_{2obs}\sqrt{(n-1)\sum {b_{obs}}_j^2}}{\hat{\sigma}_{1obs}\sqrt{(m-1)\sum {a_{obs}}_i^2}} \left(\mathcal{F}(m-1, n-1)\right)^{-1/2}. \tag{62}$$

Furthermore,

$$\begin{aligned} R_{\delta 1} &= \hat{\mu}_{1obs} - \hat{\mu}_{2obs} - \sqrt{\hat{\sigma}_{1}{}_{obs}^2 + \hat{\sigma}_{2}{}_{obs}^2} \left(\frac{mn(m+n-2)v^2 d^2(1+\rho^2 v^2)}{(n\rho^2 v^2 + m v^2)(\sum a_i^2 + v^2 \sum b_j^2)}\right)^{\frac{1}{2}} \\ &\times \left(\frac{mn(m+n-2)v_{obs}^2(1+\rho^2 v_{obs}^2)}{(n\rho^2 v_{obs}^2 + m v_{obs}^2)(\sum {a_{obs}}_i^2 + v_{obs}^2 \sum {b_{obs}}_j^2)}\right)^{-\frac{1}{2}}, \end{aligned}$$

where $v_{obs} = \hat{\sigma}_{2}{}_{obs}^2/(\rho\hat{\sigma}_{1}{}_{obs}^2)$. Therefore,

$$\begin{aligned} R_{\delta 1} &= \hat{\mu}_{1obs} - \hat{\mu}_{2obs} - \mathcal{T}(n+m-2)\sqrt{\hat{\sigma}_{1}{}_{obs}^2 + \hat{\sigma}_{2}{}_{obs}^2} \\ &\times \left(\frac{mn(m+n-2)v_{obs}^2(1+\rho^2 v_{obs}^2)}{(n\rho^2 v_{obs}^2 + m v_{obs}^2)(\sum {a_{obs}}_i^2 + v_{obs}^2 \sum {b_{obs}}_j^2)}\right)^{-\frac{1}{2}}. \end{aligned} \tag{63}$$

As one can see, $R_{\delta 1}$ is a function of ρ and its pdf depends on ρ. Hence, if ρ is unknown, $R_{\delta 1}$ is not a GPQ. In this case, we use the second approach instead.

2. Second approach. Let

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i, \quad S_1^2 = \sum_{i=1}^{n}(a_i - \bar{a})^2, \quad \bar{b} = \frac{1}{m}\sum_{j=1}^{m} b_i, \quad \text{and} \quad S_2^2 = \sum_{j=1}^{m}(b_i - \bar{b})^2.$$

If $X_i \sim \mathcal{N}(\mu_1, \sigma_1)$ and $Y_i \sim \mathcal{N}(\mu_2, \sigma_2)$, as discussed in Example 3.4,

$$\begin{aligned} R_{31} &= \hat{\mu}_{1obs} - \hat{\sigma}_{1obs}\left(\frac{\hat{\mu}_1 - \mu_1}{\hat{\sigma}_1}\right) \\ &= \hat{\mu}_{1obs} - \hat{\sigma}_{1obs}\frac{S_{1obs}}{\sqrt{n(n-1)}}\left(\frac{\sqrt{n(n-1)}\left(\frac{\hat{\mu}_1 - \mu_1}{\hat{\sigma}_1}\right)}{S_1}\right) \end{aligned}$$

which gives,

$$R_{31} = \hat{\mu}_{1obs} - \hat{\sigma}_{1obs}\left(\frac{S_{1obs}}{\sqrt{n(n-1)}}\left(\mathcal{T}_{n-1}\right)\right).$$

Similarly,

$$R_{32} = \hat{\mu}_{2obs} - \hat{\sigma}_{2obs}\left(\frac{S_{2obs}}{\sqrt{m(m-1)}}\left(\mathcal{T}_{m-1}\right)\right).$$

Therefore,

$$R_{\delta 2} \quad = \quad R_{31} - R_{32},$$

and finally,

$$R_{\delta 2} = \hat{\mu}_{1obs} - \hat{\mu}_{2obs} - \hat{\sigma}_{1obs}\left(\frac{S_{1obs}\times\ \mathcal{T}_{n-1}}{\sqrt{n(n-1)}}\right) + \hat{\sigma}_{2obs}\left(\frac{S_{2obs}\times\ \mathcal{T}_{m-1}}{\sqrt{m(m-1)}}\right) \tag{64}$$

is the GPQ for δ obtained by the second approach.

Similarly, by applying the second approach, which is discussed in Section 4.3,

$$R_{41} \quad = \quad \hat{\sigma}_{1obs}(\hat{\sigma_1}^2/\sigma_1^2)^{-1} = S_{1obs}\hat{\sigma}_{1obs}(S_1\hat{\sigma_1}^2/\sigma_1^2)^{-1} = \frac{S_{1obs}\hat{\sigma}_{1obs}}{\sqrt{\mathcal{X}^2_{n-1}}},$$

and

$$R_{42} \quad = \quad \hat{\sigma}_{2obs}(\hat{\sigma_2}^2/\sigma_2^2)^{-1} = S_{2obs}\hat{\sigma}_{1obs}(S_2\hat{\sigma_2}^2/\sigma_2^2)^{-1} = \frac{S_{2obs}\hat{\sigma}_{2obs}}{\sqrt{\mathcal{X}^2_{m-1}}}.$$

Therefore, another GPQ of ρ, as suggested in Section 4.3, is

$$R_{\rho 2} \quad = \quad R_{42}/R_{41} = \frac{S_{2obs}\hat{\sigma}_{2obs}\sqrt{\mathcal{X}^2_{n-1}}}{S_{1obs}\hat{\sigma}_{1obs}\sqrt{\mathcal{X}^2_{m-1}}} = \frac{S_{2obs}\hat{\sigma}_{2obs}}{S_{1obs}\hat{\sigma}_{1obs}}\sqrt{\frac{\mathcal{X}^2_{n-1}}{\mathcal{X}^2_{m-1}}}.$$

It can be verified that $\bar{a} = \bar{b} = 0$, thus $S_1 = \sum_{i=1}^{n} a_i^2$ and $S_2 = \sum_{j=1}^{m} b_j^2$. Further, is well known that

$$\frac{\mathcal{X}^2_{n-1}/(n-1)}{\mathcal{X}^2_{m-1}/(m-1)} \sim \mathcal{F}(n-1, m-1). \tag{65}$$

These imply that

$$\begin{aligned} R_{\rho 2} &= \frac{S_{2obs}\hat{\sigma}_{2obs}}{S_{1obs}\hat{\sigma}_{1obs}}\sqrt{\frac{\mathcal{X}^2_{n-1}}{\mathcal{X}^2_{m-1}}} \\ &= \frac{\hat{\sigma}_{2obs}\sqrt{(n-1)\sum {b_{obs}}_j^2}}{\hat{\sigma}_{1obs}\sqrt{(m-1)\sum {a_{obs}}_i^2}} \left(\mathcal{F}(m-1,n-1)\right)^{-1/2}, \end{aligned}$$

which is same as (62).

Computation of GCI and GPV.

Based on the GPQ we constructed above, the $100\gamma\%$ GCI can be computed by applying the generalized pivotal quantity to (2). For instance, since the observed pivotal of $R_{\rho 1}$ is ρ, we can use (3), instead of (2). In this case, the 95% GCI of ρ, say CI_ρ, is given by

$$Pr(R_{\rho 1} \in CI_\rho) = 0.95.$$

In addition, the GPV methods provided in Chapter 2 can also be applied here, with the GPQ constructed above. For example, if we want to test (45), the GPV is

$$Pr(R_{\rho 1} \geqslant \rho_0).$$

In order to verify the efficiency of the generalized pivotal quantities and the related GCI and GPV methods provided above, we carry out some numerical simulations, which are shown in Chapter 5.

CHAPTER 5

Numerical results, simulation study, and applications

In this section, we carry out some simulation studies in order to evaluate the performances of the (conditional) approaches we discussed in the previous sections (i.e. Pitman estimator, generalized p-value (GPV), and generalized confidence interval (GCI), with small and moderate sample sizes. Furthermore, we also apply these approaches to some real data sets.

5.1. Point estimation approach

To study the efficiency of Pitman estimator, in the following examples we repeat the simulation 10000 times. In each replicate, we first generate the data under the given condition, then based on the data we compute the Pitman estimator and MLE (if it exists), respectively. After performing the 10000 simulated samples and obtaining the estimators, we evaluate the performance by computing the average of the 10000 estimators and the MSE.

Example 5.1. Normal distribution with known location parameter and unknown scale parameter (Example 3.2).

Here we work under the same condition as given in Example 3.2, where the exact value of σ is given by $\sigma = 2$. By applying the simulation method discussed above, we present the results with different values of size n as shown in Table 5.1.

TABLE 5.1. Numerical results of MRE and MLE in Example 3.2

Size	$\hat{\sigma}_P(X)$	MSE of $\hat{\sigma}_P(X)$	$\hat{\sigma}_M(X)$	MSE of $\hat{\sigma}_M(X)$
n=5	1.793630	0.3754643	1.884990	0.3808773
n=10	1.904615	0.1944069	1.952750	0.1970261
n=50	1.980724	0.04007357	1.990652	0.04018838
n=100	1.987849	0.02025364	1.992825	0.02025826

From Table 5.1, it is noted that the MSE of MLE is larger than MSE of Pitman estimator in scale family. But as the size increases, the MSE of MLE is getting close to the MSE of Pitman estimator.

Example 5.2. Normal distribution with unknown location and scale parameters (Example 3.3).

In Example 3.3, we verify that the Pitman estimator of μ is $\bar{X}$, which is also the MLE. However, the Pitman estimator of σ is different from the MLE. Here we carry out the simulations and study the difference between the MSE of Pitman estimator and MSE of MLE. By setting exact values $\mu = \sigma = 2$. The numerical results are presented in Table 5.2.

From Table 5.2, it is noted that the MSE of MLE is larger than MSE of Pitman estimator of scale parameter. But as the size increases, the MSE of MLE is getting

TABLE 5.2. Numerical results of MRE and MLE in Example 3.3

Size	$\hat{\sigma}_P(X)$	MSE of $\hat{\sigma}_P(X)$	$\hat{\sigma}_M(X)$	MSE of $\hat{\sigma}_M(X)$
n=5	1.756706	0.4744318	1.671564	0.4838348
n=10	1.89907	0.1991592	1.852258	0.2021038
n=50	1.967015	0.04082784	1.957205	0.04117583
n=100	1.981417	0.01999685	1.976469	0.02001215

closer to the MSE of Pitman estimator.

In the beginning of Chapter 3, we have presented 2 examples for the case where MLE does not exist. Here we apply the Pitman estimation method to these 2 examples, and evaluate the efficiency of Pitman estimator in location-scale family in which MLE does not exist.

Example 5.3.

Let $X_1, ..., X_n$ be iid with pdf $f(x|\mu, \sigma) = \sigma^{-l} g((x - \mu)/\sigma)$, where $g(x) = c(xlog^2 x)^{-1}$, $0 < x \leq k$, k is any constant that satisfies $0 < k < 1$, $c = -1/log(k)$ is a constant and $-\infty < \mu < \infty$, $0 < \sigma < \infty$ are both unknown. By choosing $\mu = \sigma = 2$ and using the simulation method with different sample size n, the results are shown in Table 5.3.

TABLE 5.3. Numerical results of Pitman estimator in Example 5.3

Size	$\hat{\mu}(x)$	MSE of $\hat{\mu}(x)$	$\hat{\sigma}(x)$	MSE of $\hat{\sigma}(x)$
n=10	2.118721	0.04340062	2.386992	0.2505716
n=100	1.989474	0.01407563	2.011956	0.01343508
n=200	1.990491	0.01373037	2.003245	0.01342210

From Table 5.3, the simulation results show that even for small sample size the Pitman point estimator is very close to the exact value. In addition, as the sample size increases, the Pitman estimators of location and scale parameters both get closer to the exact value.

Example 5.4.

Let $X_1, ..., X_n$ be iid with pdf $f(x|\mu, \sigma) = \sigma^{-l} g((x-\mu)/\sigma)$, where

$$g(x) = \frac{1}{2(1+|x|)(1+\log(1+|x|))^2}, -\infty < x < \infty,$$

with $-\infty < \mu < \infty$, $0 < \sigma < \infty$ are both unknown. In this case, if we let $F(x)$ denote the cumulative distribution function (cdf), it can be verified that

$$F(x) = \begin{cases} \dfrac{1}{2(1+\log(1-(x-\mu)/\sigma))} & \text{if } x \leqslant \mu \\ 1 - \dfrac{1}{2(1+\log(1+(x-\mu)/\sigma))} & \text{if } x > \mu \end{cases}$$

To evaluate the Pitman estimator in this location-scale family numerically, we apply the simulation method given in the previous examples. To this end, we choose $\mu = \sigma = 2$ and carry out the simulations with different sample size n, as presented in Table 5.4.

TABLE 5.4. Numerical results of Pitman estimator in Example 5.4

Size	$\hat{\mu}_P(X)$	MSE of $\hat{\mu}_P(X)$	$\hat{\sigma}_P(X)$	MSE of $\hat{\sigma}_P(X)$
n=5	1.0066701	11.247332	0.7996300	3.2886735
n=10	2.0202800	6.2806460	1.2966522	1.2754575
n=50	2.00331056	0.06228972	1.87996060	0.20682585
n=60	1.99697324	0.05619604	1.89837688	0.17093954

From Table 5.4, it is noted that when sample size is small, the Pitman estimator may produce large MSE due to the heavy tails of the distribution. But as the sample size increases, the MSE becomes small and the Pitman estimators of location and scale parameters both get closer to the exact value.

5.2. Generalized P-values and confidence intervals (univariate case)

In the previous chapters, the GCI and GPV methods for univariate case are provided and theoretically analyzed. In this subsection, we use simulation methods to illustrate the performances of these methods.

To evaluate the efficiency of GCI method, we set the confidence coefficient $\gamma = 0.95$ and study the related coverage probability, that is, the probability that the GCI will contain the exact value of the parameter of interest. Ideally, the coverage probability of a 95% GCI should be 0.95. However, due to sampling variation, the actual coverage probability of the interval may not be exactly equal to 0.95. In this case, it is necessary to simulate the coverage probabilities under different situations, which are presented in the following examples.

For GPV, in McNally, Iyer, and Mathew (2003), a simulation method is used to evaluate the power of the test based on the GPV approach. In the following examples, we use the same method and set the significance level $\alpha = 0.05$ to study the power of the suggested tests. In addition, the generalized test statistic used here is similar to that in Bebu and Mathew (2007).

Example 5.5. GPV and GCI for Example 3.4.

Under the situation of Example 3.4, in order to evaluate the generalized p-values of μ and σ under the hypothesis tests (16) and (18), respectively, we simulate the power at significance level $\alpha = .05$. By choosing $\mu_0 = \sigma_0 = 2$, the results for different exact values of μ and σ are shown in Table 5.5, with the related figures shown in Figures 5.1 and 5.2, respectively. Further, in Table 5.6, we choose $\mu = \sigma = 2$ and the simulated coverage probability is presented.

TABLE 5.5. The simulated powers in Example 5.5

(μ, σ)	Size	Power of (16)	Power of (18)
(.5, .5)	n=5	1	0.9773
	n=10	1	1
	n=100	1	1
(1, 1)	n=5	0.5781	0.4383
	n=10	0.8961	0.8512
	n=100	1	1
(1.5, 1.5)	n=5	0.1518	0.1312
	n=10	0.2576	0.2547
	n=100	0.9542	0.9733
(2, 2)	n=5	0.0463	0.0484
	n=10	0.0522	0.0465
	n=100	0.0518	0.0455
(3, 3)	n=5	0.0123	0.0134
	n=10	0.005	0.003
	n=100	0	0
(4, 4)	n=5	0.0049	0.0044
	n=10	.0017	0.0005
	n=100	0	0

From Table 5.5, one can see that under the null hypothesis, i.e. $\mu > 2$, as the exact value of the parameter of interest μ increases, the power decreases, and as μ

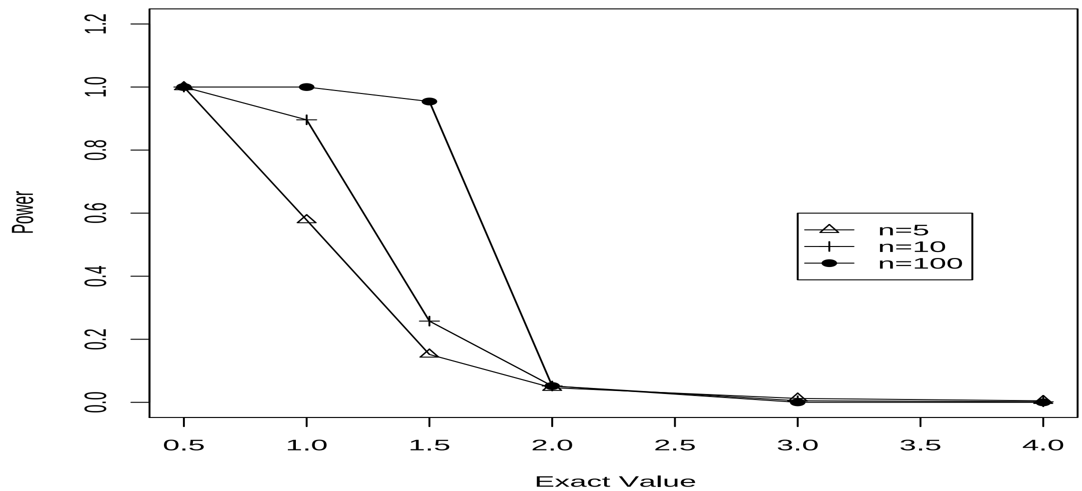

FIGURE 5.1. The simulated powers of (16) in Example 5.5

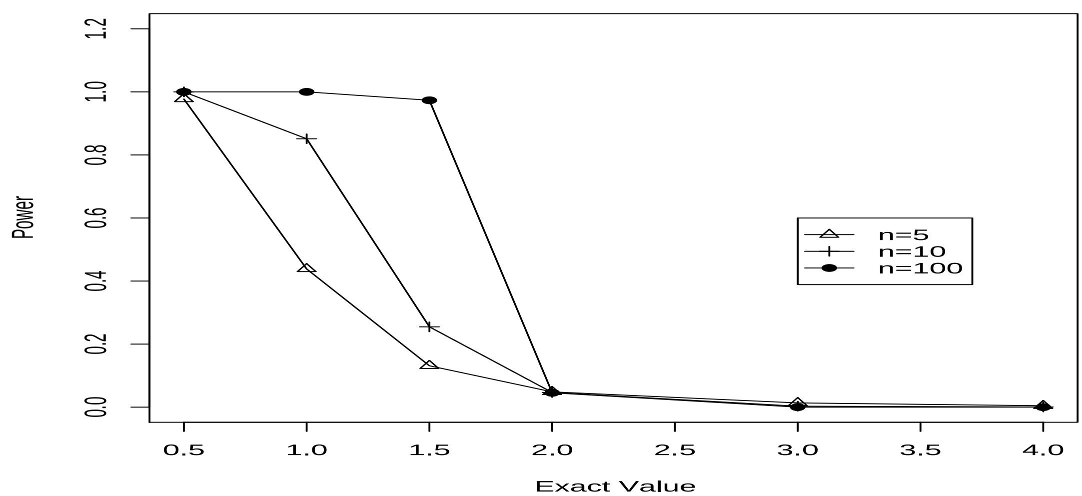

FIGURE 5.2. The simulated powers of (18) in Example 5.5

tends to infinity, the power function tends to 0. In addition, when μ tends to 2, the power is close to 0.05. This shows that the provided generalized test is consistent. The above results can also be verified in Figure 5.1 and 5.2. From the figures, it

TABLE 5.6. The simulated coverage probabilities of the 95% GCI in Example 5.5

Size	GCI of $\hat{\mu}_M(X)$	GCI of $\hat{\mu}_P(X)$	GCI of $\hat{\sigma}_M(X)$	GCI of $\hat{\sigma}_P(X)$
n=2	0.917	0.920	0.932	0.934
n=5	0.932	0.947	0.946	0.941
n=10	0.956	0.948	0.947	0.948
n=100	0.952	0.951	0.950	0.950

can be seen that when $\theta = \theta_0 = 2$ (here θ denote to the parameter of interest), the powers are all approximately equal to 0.05. But on the left hand side of 2, the power continually increases to 1 when the distance between θ and 2 increases. And in the right hand side the power decreases to 0 when the distance increases. Furthermore, in the left hand side of 2, for each exact value of θ, the power increases as the sample size increases. This implies that the hypothesis become more precise when sample size is large (note that on the right hand side of 2 implies that the alternative hypothesis is true).

In addition, it is shown in Table 5.6 that, when the sample size is small, the simulated coverage probability is close to .95, besides, as the size increases, the coverage probability gets close to .95. This implies that the GCI method performs well in this example.

In summary, the numerical results show that the provided GCI and GPV methods are optimal.

Example 5.6. GPV and GCI for Example 3.5.

In this example, consider the hypothesis tests (16) and (18), respectively. With $\mu_0 = \sigma_0 = 2$, we simulate the powers under the different values of μ and σ. The results are shown in Table 5.7, with the related figures shown in Figure 5.3 and 5.4, respectively. Further, by choosing different values of μ and σ, the simulation results of coverage probabilities are presented in Table 5.8.

TABLE 5.7. The simulated powers in Example 5.6

(μ, σ)	Size	Power of (16)	Power of (18)
(.5, .5)	n=5	0.9425	0.8506
	n=10	1	0.9969
	n=100	1	1
(1, 1)	n=5	0.4062	0.3046
	n=10	0.6107	0.5683
	n=100	1	1
(1.5, 1.5)	n=5	0.1352	0.1027
	n=10	0.1919	0.1439
	n=100	0.6587	0.7668
(2, 2)	n=5	0.0485	0.0521
	n=10	0.0567	0.0374
	n=100	0.0852	0.0243
(3, 3)	n=5	0.001	0.003
	n=10	0	0
	n=100	0	0
(4, 4)	n=5	0	0
	n=10	0	0
	n=100	0	0

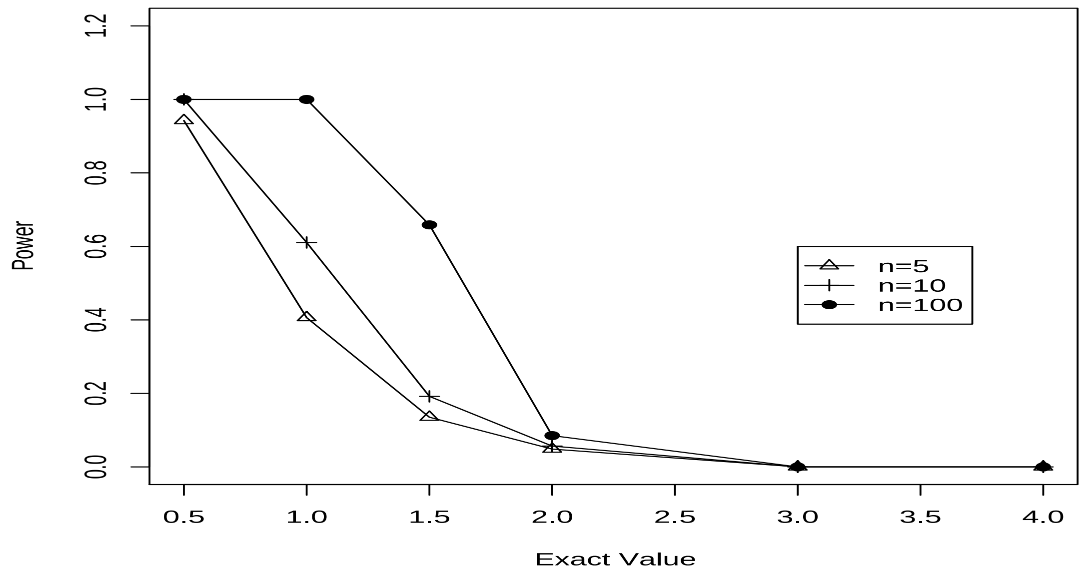

FIGURE 5.3. The simulated powers of (16) in Example 5.6

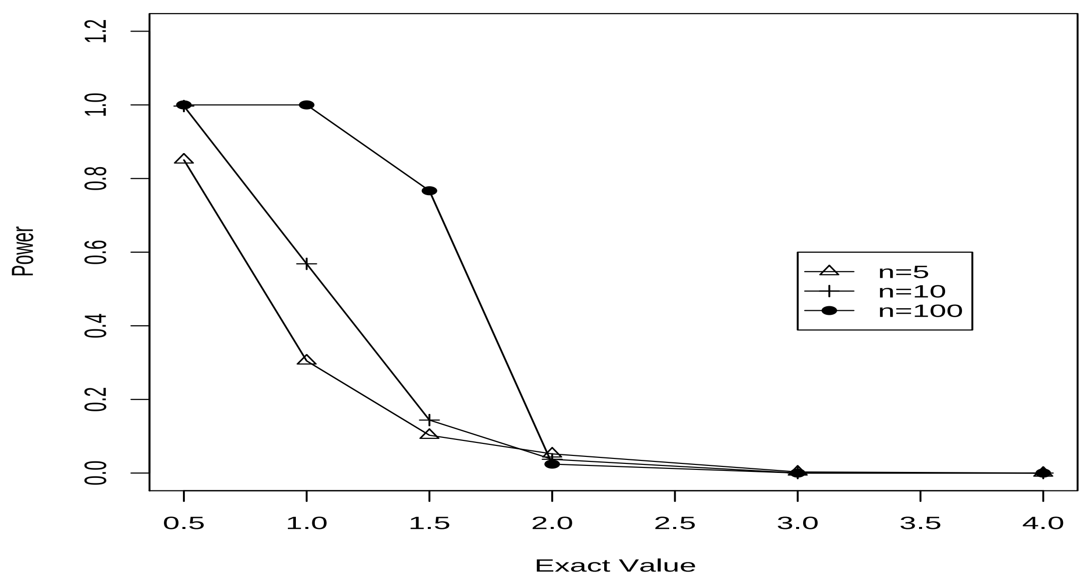

FIGURE 5.4. The simulated powers of (18) in Example 5.6

TABLE 5.8. The simulated coverage probabilities of the 95% GCI in Example 5.6

(μ, σ)	Size	GCI of μ	GCI of σ
(2, 2)	n=5	0.938	0.944
	n=10	0.942	0.957
	n=100	0.959	0.952
(-2, 0.5)	n=5	0.934	0.930
	n=10	0.941	0.943
	n=100	0.954	0.948

From Table 5.7, one can see that for the parameter of interest, the power decreases when the exact value increases. In addition, the power is close to 0.05 when the exact value is equal to 2. Also, Figure 5.3 and 5.4 show the same information. Also, from these figures, when $\theta = \theta_0 = 2$ (here θ denote to the parameter of interest), the powers are all approximately equal to 0.05. But on the left hand side of 2, the power continually increases to 1 when the distance between θ and 2 increases. And in the right hand side the power decreases to 0 when the distance increases. Furthermore, in the left hand side of 2, for each exact value of θ, the power increases as the sample size increases. This implies that the hypothesis become more precise when sample size is large (note that on the right hand side of 2 implies that the alternative hypothesis is true).

In addition, it is shown in Table 5.8 that, when the sample size is small, the simulated coverage probability is close to .95. Besides, as the size increases, the coverage probability gets close to .95. Once again, as discussed in Example 5.5, the numerical results show that the GCI and GPV methods are asymptotically optimal for the data generated by the model in Example 5.6.

Example 5.7. GPV and GCI for Example 3.6.

With the same procedure as described in the previous example, the simulation results of powers and coverage probabilities are presented in Table 10 and 11, respectively. Here the values of μ_0 and σ_0 in (16) and (18) are chosen to be $\mu_0 = \sigma_0 = 2$. The simulated powers under the different exact values of μ and σ are given in Table 5.9, with the related figures presented in Figure 5.5 and 5.6, respectively. Also, by choosing different values of μ and σ, the simulation results of coverage probabilities are given in Table 5.10.

TABLE 5.9. The simulated powers in Example 5.7

(μ, σ)	Size	Power of (16)	Power of (18)
(.5, .5)	n=5	0.9583	0.9431
	n=10	1	1
	n=100	1	1
(1, 1)	n=5	0.4062	0.4017
	n=10	0.6852	0.7541
	n=100	1	1
(1.5, 1.5)	n=5	0.1889	0.1263
	n=10	0.2773	0.2181
	n=100	0.9810	0.9533
(2, 2)	n=5	0.0851	0.0413
	n=10	0.0654	0.0439
	n=100	0.0583	0.0366
(3, 3)	n=5	0.0027	0
	n=10	0.0013	0
	n=100	0.0026	0.0012
(4, 4)	n=5	0.0003	0
	n=10	0.0004	0
	n=100	0.0018	0

From Table 5.9 and 5.10, it is noticed that the simulated results are quite similar

TABLE 5.10. The simulated coverage probabilities of the 95% GCI in Example 5.7

(μ, σ)	Size	GCI of μ	GCI of σ
(2, 2)	n=5	0.953	0.942
	n=10	0.954	0.947
	n=100	0.950	0.953
(-2, 0.5)	n=5	0.955	0.922
	n=10	0.946	0.947
	n=100	0.948	0.954

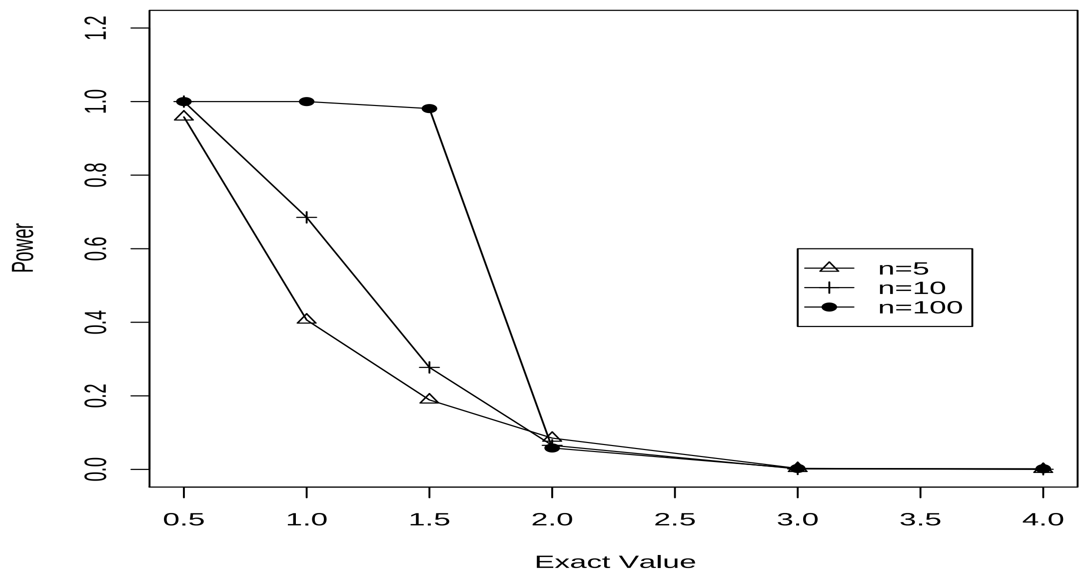

FIGURE 5.5. The simulated powers of (16) in Example 5.7

with Example 5.5 and 6.6. That is, for the parameter of interest, the power decreases when the exact value increases. In addition, the power is close to 0.05 when the exact value is equal to 2. Also, Figure 5.5 and 5.6 illustrate the same pattern. From the figures, when $\theta = \theta_0 = 2$ (here θ denote to the parameter of interest), the powers are all approximately equal to 0.05. But on the left hand side of 2, the power continually increases to 1 when the distance between θ and 2 increases. And in the right hand side

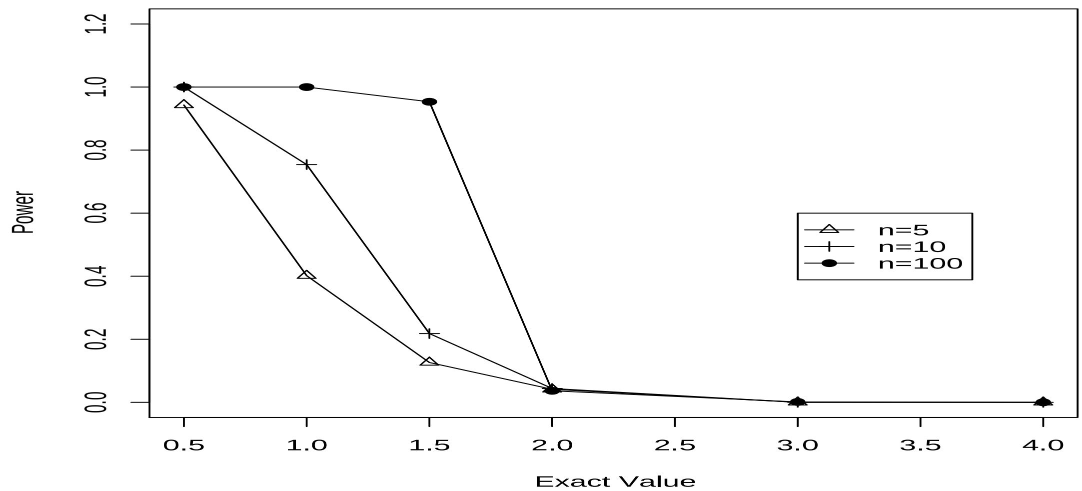

FIGURE 5.6. The simulated powers of (18) in Example 5.7

the power decreases to 0 when the distance increases. Furthermore, in the left hand side of 2, for each exact value of θ, the power increases as the sample size increases. This reflects the fact that the hypothesis test become more precise when sample size is large (note that on the right hand side of 2 implies that the alternative hypothesis is true).

In addition, for the coverage probability presented in Table 5.10, the simulated coverage probability is close to .95 when the sample size is small. Besides, as the size increases, the coverage probability get close to .95. Thus, as discussed in Example 5.5, the numerical results show that the GCI and GPV methods are also optimal for Example 5.7.

Example 5.8. GPV and GCI for Example 5.4.

In Example 5.4, we present the case where the Pitman estimators is applicable for a location-scale family, while the MLEs for location and scale parameters do not exist. In this subsection, we apply the GCI methods to the same distribution, in which the pdf is given by

$$f(x|\mu,\sigma) = \frac{1}{2\sigma(1+|\frac{x-\mu}{\sigma}|)(1+\log(1+|\frac{x-\mu}{\sigma}|))^2},$$

$-\infty < x < \infty$, and $-\infty < \mu < \infty$, $0 < \sigma < \infty$ are both unknown.

Simulation study. Under the hypothesis tests (16) and (18), let $\mu_0 = \sigma_0 = 2$. The simulated powers under the different exact values of μ and σ are given in Table 5.11, with the related figures shown in Figure 5.7 and 5.8, respectively. Also, by choosing different values of μ and σ, the simulated coverage probabilities are given in Table 5.12.

For the coverage probabilities presented in Table 5.12, the simulated coverage probability is lower than .95 when the sample size is small. These results imply that the GPQ method provided in this paper does not perform well for small sample size. This problem may be caused by the fact that the Pitman estimators for this example exist but do not perform well, which are discussed in Example 5.4. However, as the size increases, the coverage probability get close to .95. In conclusion, the GCI and GPV methods seem to be optimal for this example when the sample size is moderate. In addition, from Table 5.11 and 5.12, it can be seen that the power decreases when the exact value increases. Also, Figure 5.7 and 5.8 illustrate the same information. In fact, from these figures, when $\theta = \theta_0 = 2$ (here θ denote to the parameter of interest), the powers are higher than 0.05, when sample size is small. This result confirms that

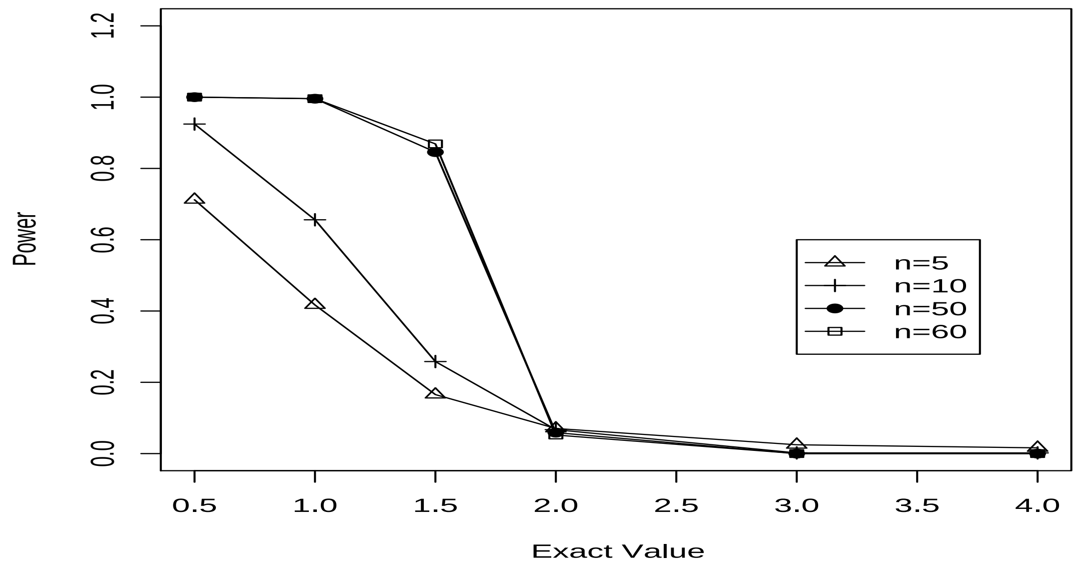

Figure 5.7. The simulated powers of (16) in Example 5.8

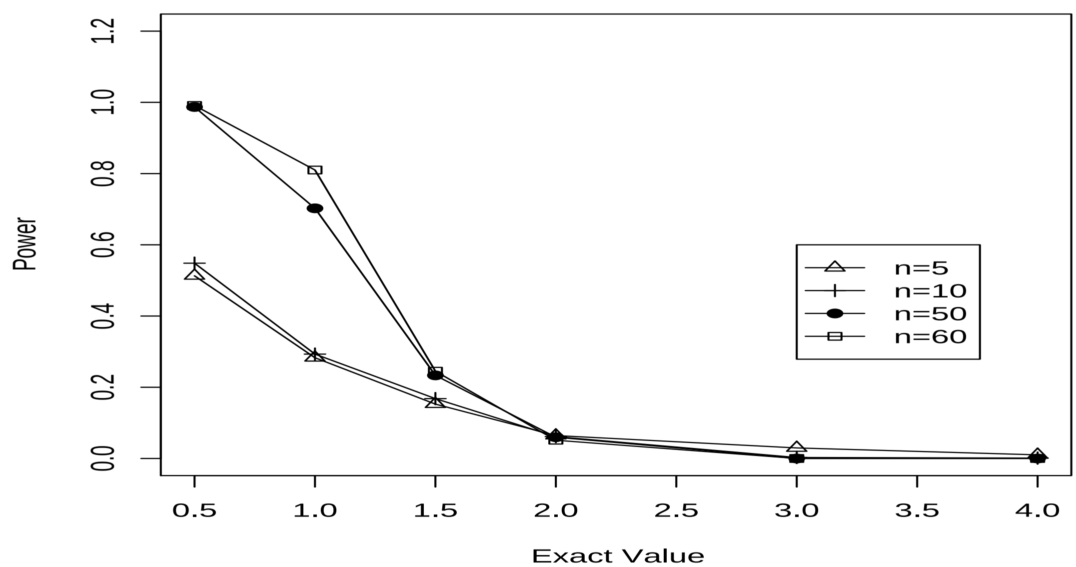

Figure 5.8. The simulated powers of (18) in Example 5.8

TABLE 5.11. The simulated powers in Example 5.8

(μ, σ)	Size	Power of (16)	Power of (18)
(.5, .5)	n=5	0.7121	0.5129
	n=10	0.9245	0.5484
	n=50	1	0.9870
	n=60	1	0.9912
(1, 1)	n=5	0.4169	0.2829
	n=10	0.6559	0.2931
	n=50	0.9956	0.7025
	n=60	0.9957	0.8100
(1.5, 1.5)	n=5	0.1655	0.1527
	n=10	0.2582	0.1681
	n=50	0.8458	0.2334
	n=60	0.8688	0.2451
(2, 2)	n=5	0.0705	0.0544
	n=10	0.0671	0.0610
	n=50	0.0588	0.0597
	n=60	0.0521	0.0513
(3, 3)	n=5	0.0247	0.0296
	n=10	0.0019	0.0028
	n=50	0	0
	n=60	0	0
(4, 4)	n=5	0.0161	0.0102
	n=10	0.0019	0.0011
	n=50	0	0
	n=60	0	0

TABLE 5.12. The simulated coverage probabilities of the 95% GCI in Example 5.8

(μ, σ)	Size	GCI of μ	GCI of σ
(2, 2)	n=5	0.804	0.793
	n=10	0.870	0.931
	n=50	0.943	0.941
	n=60	0.943	0.948
(-2, 0.5)	n=5	0.825	0.804
	n=10	0.908	0.940
	n=50	0.944	0.945
	n=60	0.946	0.949

in this example with small sample size, our methods seem to be inaccurate, which needs to be investigated in future. However, as sample size increases, the power get close to 0.05. Moveover, on the left hand side of 2, the power continually increases to 1 when the distance between θ and 2 increases. And in the right hand side the power decreases to 0 when the distance increases. Furthermore, in the left hand side of 2, for each exact value of θ, the power increases as the sample size increases. This indicates that the hypothesis test is unbiased and consistent when sample size is large.

5.3. Generalized P-value and confidence interval (bivariate case)

Example 5.9. GCI for Example 4.1.

To perform the simulation study for Example 4.1, we choose the different exact values of μ_1, σ_1, μ_2, σ_2, respectively. With the methods we provide in Example 4.1, the simulated coverage probabilities of the .95 C.I. for ρ, δ conditionally to v, are presented in Table 5.13 and Table 5.14.

TABLE 5.13. The simulated coverage probabilities of the 95% GCI in Example 5.9

$(\mu_1, \mu_2, \sigma_1, \sigma_2)$	n	m	GCI of δ (ρ known)	GCI of ρ
(2, 2, 2, 2)	5	5	0.9499	0.9521
	10	10	0.9487	0.9512
	20	20	0.9497	0.9488
	50	50	0.9496	0.9507
	100	100	0.9519	0.9525
	5	10	0.9471	0.9482
	5	100	0.9454	0.9511
	50	100	0.9527	0.9475
	10	5	0.9467	0.9485
	100	5	0.9459	0.9470
	100	50	0.9471	0.9516
(2, 1, 2, 1)	5	5	0.9508	0.9490
	10	10	0.9519	0.9498
	20	20	0.9477	0.9490
	50	50	0.9512	0.9503
	100	100	0.9516	0.9482
	5	10	0.9496	0.9466
	5	100	0.9521	0.9514
	50	100	0.9507	0.9498
	10	5	0.9491	0.9495
	100	5	0.9512	0.9492
	100	50	0.9497	0.9505

Table 5.14. Continuation of Table 5.13

$(\mu_1, \mu_2, \sigma_1, \sigma_2)$	n	m	GCI of δ (ρ known)	GCI of ρ
(1, 2, 1, 2)	5	5	0.9539	0.9511
	10	10	0.9527	0.9505
	20	20	0.9521	0.9469
	50	50	0.9516	0.9485
	100	100	0.9520	0.9545
	5	10	0.9469	0.9496
	5	100	0.9495	0.9452
	50	100	0.9516	0.9471
	10	5	0.9475	0.9511
	100	5	0.9484	0.9514
	100	50	0.9512	0.9497
(2, 1, 200, 1)	5	5	0.9463	0.9481
	10	10	0.9482	0.9476
	20	20	0.9474	0.9508
	50	50	0.9528	0.9511
	100	100	0.9473	0.9484
	5	10	0.9495	0.9470
	5	100	0.9477	0.9482
	50	100	0.9462	0.9476
	10	5	0.9492	0.9466
	100	5	0.9464	0.9509
	100	50	0.9521	0.9482
(2, 1, 2, 100)	5	5	0.9546	0.9471
	10	10	0.9458	0.9473
	20	20	0.9471	0.9534
	50	50	0.9483	0.9497
	100	100	0.9485	0.9493
	5	10	0.9519	0.9508
	5	100	0.9512	0.9464
	50	100	0.9454	0.9489
	10	5	0.9491	0.9490
	100	5	0.9483	0.9489
	100	50	0.9522	0.9473

As one can see from Table 5.13 and 5.14, the coverage probabilities for ρ are all close to 0.95 under each situation. This implies that our methods for constructing the GCI of ρ perform well. Besides, for GCI of δ conditionally to v, first we consider the case that the sample sizes of the two groups (X and Y) are equal. In this case, it is shown in Table 5.13 and 5.14 that when the sample sizes are small, i.e. $n = m = 5$, the differences between the simulated coverage probabilities .95 are around .5. Moreover, if the ratio ρ is not significant, the coverage probabilities are higher than 0.95, while it is lower than 0.95, for the case that the ratio is significant (i.e. $\rho = 100$). However, as the size increases, the coverage probabilities get close to .95.

Further, the simulated coverage probabilities for the case that the sample sizes of the two groups are unequal are also provided in the above tables. It can be seen from Table 5.13 and 5.14, unless the two samples are come from the same population, the results turn out to be inaccurate, especially for the cases that the difference of sample sizes are large.

In addition, the coverage probabilities for .95 GCI of δ with unknown ρ, as discussed in Chapter 4, only the second approach can be applied. By using the second approach, which is discussed in Example 4.1, the results are presented in Table 5.15 and 5.16.

Table 5.15. The simulated coverage probabilities of the 95% GCI for δ with unknown ρ

$(\mu_1, \mu_2, \sigma_1, \sigma_2)$	n	m	Second approach
(2, 2, 2, 2)	5	5	0.9748
	10	10	0.9606
	20	20	0.9572
	50	50	0.9526
	100	100	0.9503
	5	10	0.9657
	5	100	0.9496
	50	100	0.9528
	10	5	0.9664
	100	5	0.9507
	100	50	0.9498
(2, 1, 2, 1)	5	5	0.9698
	10	10	0.9587
	20	20	0.9545
	50	50	0.9512
	100	100	0.9501
	5	10	0.9586
	5	100	0.9507
	50	100	0.9492
	10	5	0.9648
	100	5	0.9532
	100	50	0.9509

TABLE 5.16. Continuation of Table 5.15

$(\mu_1, \mu_2, \sigma_1, \sigma_2)$	n	m	Second approach
(1, 2, 1, 2)	5	5	0.9716
	10	10	0.9586
	20	20	0.9520
	50	50	0.9495
	100	100	0.9508
	5	10	0.9709
	5	100	0.9574
	50	100	0.9508
	10	5	0.9559
	100	5	0.9493
	100	50	0.9474
(2, 1, 200, 1)	5	5	0.9513
	10	10	0.9497
	20	20	0.9485
	50	50	0.9525
	100	100	0.9505
	5	10	0.9447
	5	100	0.9496
	50	100	0.9501
	10	5	0.9540
	100	5	0.9508
	100	50	0.9493
(2, 1, 2, 100)	5	5	0.9494
	10	10	0.9482
	20	20	0.9511
	50	50	0.9509
	100	100	0.9496
	5	10	0.9479
	5	100	0.9484
	50	100	0.9495
	10	5	0.9533
	100	5	0.9508
	100	50	0.9503

It is shown in Table 5.15 and 5.16 that, if the exact value of ρ is close to 1, the coverage probabilities based on the second approached are all a little higher than 0.95, and get close to 0.95 as the sample sizes increase. However, if ρ is significant larger than 1, for instance, $\rho = 100$, in this case, as we can see from the tables, the coverage probabilities are close to 0.95, even for the small sample sizes.

Example 5.10. GPV for δ in Example 4.1.

In Example 4.1, we also provide the procedures for computing the GPV with respect to (45) and (55), respectively. Under the same condition of Example 4.1, where X and Y are normal distributed with unknown ρ, we use the same simulation method presented in Krishnamoorthy and Mathew (2003) to evaluate the performances of (59) for testing (55). For (46) and (60), the interested readers can verify in the same way. In this example, we assume the sample sizes m and n are equal, μ_2 is set to be 2, and under the significance level $\alpha = 0.05$ is, the null hypothesis is H_0: $\delta \geqslant 0$, versus the alternative hypothesis H_1: $\delta < 0$. Let SA denote the simulated powers based on (59). Then, by choosing the different values of m, n, μ_1, σ_1, σ_2, and $\delta = \mu_1 - \mu_2$, the simulated results are provided in the following tables. Firstly, we consider the case that σ_1 and σ_2 are equal. In this case, we choose $\sigma_1 = \sigma_2 = 2$ (both small) and $\sigma_1 = \sigma_2 = 200$ (both large), the results are presented in Table 5.17. Secondly, for the case that σ_1 and σ_2 are different, we choose $\sigma_1 = 1$, $\sigma_2 = 2$ (small difference) and $\sigma_1 = 1$, $\sigma_2 = 200$ (significant difference), the results are provided in Table 5.18. Besides, for the simulated powers presented in Table 5.17 and 5.18, the related figures are given in Figure 5.9∼5.12, respectively.

It is shown in Table 5.17, 5.18 and the related figures that, under the null hypothesis of (55) with $\delta_0 = 0$, when μ_2 is fixed, the power of (59) decreases as the exact value of the parameter of interest δ increases. And as μ tends to infinity, the powers both tend to 0. In addition, when μ_1 is close to μ_2 (δ tends to 0), the powers are close to 0.05. Moreover, from the related figures, it can be seen that on the left hand side of $\delta = 0$, the powers continually increase when the distance between δ and 0 increases. And in the right hand side the powers decrease to 0 when the distance increases.

Furthermore, as one can see from the tables and related figures, when the exact value of $\sigma_1^2 + \sigma_2^2$ is large, the decreasing rate of power is slower than the case that $\sigma_1^2 + \sigma_2^2$ is small. For example, in the case of $n = m = 5$, when $\sigma_1 = \sigma_2 = 200$, the power of (59) decreases from 0.1236 to 0.0332, as the exact value of δ increases from -102 to -12. However, when $\sigma_1 = 1$, and $\sigma_2 = 200$, the power decreases from 0.2482 to 0.0622, as δ increases from -102 to -12. Besides, for small sample sizes (i.e. n=m=5), when the exact value of $\sigma_1^2 + \sigma_2^2$ is small, the powers are closer to 0.05 as δ tends to 0, as compare to the case that $\sigma_1^2 + \sigma_2^2$ is large.

TABLE 5.17. The simulated powers of Example 5.10 with equal scale parameters

Sizes (m=n)	(δ, σ_1, σ_2)	SA	(δ, σ_1, σ_2)	SA
5	(-2, 2, 2)	0.3236	(-102, 200, 200)	0.1236
	(-1, 2, 2)	0.1162	(-12, 200, 200)	0.0332
	(0, 2, 2)	0.0314	(0, 200, 200)	0.0324
	(1, 2, 2)	0.0066	(8, 200, 200)	0.0276
	(2, 2, 2)	0	(98, 200, 200)	0.0056
10	(-2, 2, 2)	0.6392	(-102, 200, 200)	0.255
	(-1, 2, 2)	0.245	(-12, 200, 200)	0.053
	(0, 2, 2)	0.0412	(0, 200, 200)	0.043
	(1, 2, 2)	0.0022	(8, 200, 200)	0.0316
	(2, 2, 2)	0	(98, 200, 200)	0.003
20	(-2, 2, 2)	0.9202	(-102, 200, 200)	0.4594
	(-1, 2, 2)	0.4518	(-12, 200, 200)	0.073
	(0, 2, 2)	0.0472	(0, 200, 200)	0.0442
	(1, 2, 2)	0.0012	(8, 200, 200)	0.0322
	(2, 2, 2)	0	(98, 200, 200)	0
50	(-2, 2, 2)	0.9992	(-102, 200, 200)	0.8084
	(-1, 2, 2)	0.7898	(-12, 200, 200)	0.092
	(0, 2, 2)	0.048	(0, 200, 200)	0.0494
	(1, 2, 2)	0	(8, 200, 200)	0.0338
	(2, 2, 2)	0	(98, 200, 200)	0
100	(-2, 2, 2)	1	(-102, 200, 200)	0.9742
	(-1, 2, 2)	0.9712	(-12, 200, 200)	0.105
	(0, 2, 2)	0.0485	(0, 200, 200)	0.0506
	(1, 2, 2)	0	(8, 200, 200)	0.0274
	(2, 2, 2)	0	(98, 200, 200)	0

TABLE 5.18. The simulated powers of Example 5.10 with unequal scale parameters

Sizes (m=n)	(δ, σ_1, σ_2)	SA	(δ, σ_1, σ_2)	SA
5	(-2, 1, 2)	0.4784	(-102, 1, 200)	0.2482
	(-1, 1, 2)	0.1792	(-12, 1, 100)	0.0622
	(0, 1, 2)	0.0348	(0, 1, 200)	0.0468
	(1, 1, 2)	0.0016	(8, 1, 200)	0.0402
	(2, 1, 2)	0	(98, 1, 200)	0.0046
10	(-2, 1, 2)	0.8392	(-102, 1, 200)	0.4392
	(-1, 1, 2)	0.3624	(-12, 1, 200)	0.0634
	(0, 1, 2)	0.0445	(0, 1, 200)	0.0502
	(1, 1, 2)	0	(8, 1, 200)	0.0412
	(2, 2, 2)	0	(98, 1, 200)	0.001
20	(-2, 1, 2)	0.9866	(-102, 1, 200)	0.7073
	(-1, 1, 2)	0.6198	(-12, 1, 200)	0.087
	(0, 1, 2)	0.048	(0, 1, 200)	0.0524
	(1, 1, 2)	0	(8, 1, 200)	0.034
	(2, 2, 2)	0	(98, 1, 200)	0
50	(-2, 1, 2)	1	(-102, 1, 200)	0.9742
	(-1, 1, 2)	0.9294	(-12, 1, 200)	0.1096
	(0, 1, 2)	0.0528	(0, 1, 200)	0.0487
	(1, 1, 2)	0	(8, 1, 200)	0.0316
	(2, 2, 2)	0	(98, 1, 200)	0
100	(-2, 1, 2)	1	(-102, 1, 200)	0.9999
	(-1, 1, 2)	0.998	(-12, 1, 200)	0.1484
	(0, 1, 2)	0.046	(0, 1, 200)	0.0458
	(1, 1, 2)	0	(8, 1, 200)	0.018
	(2, 2, 2)	0	(98, 1, 200)	0

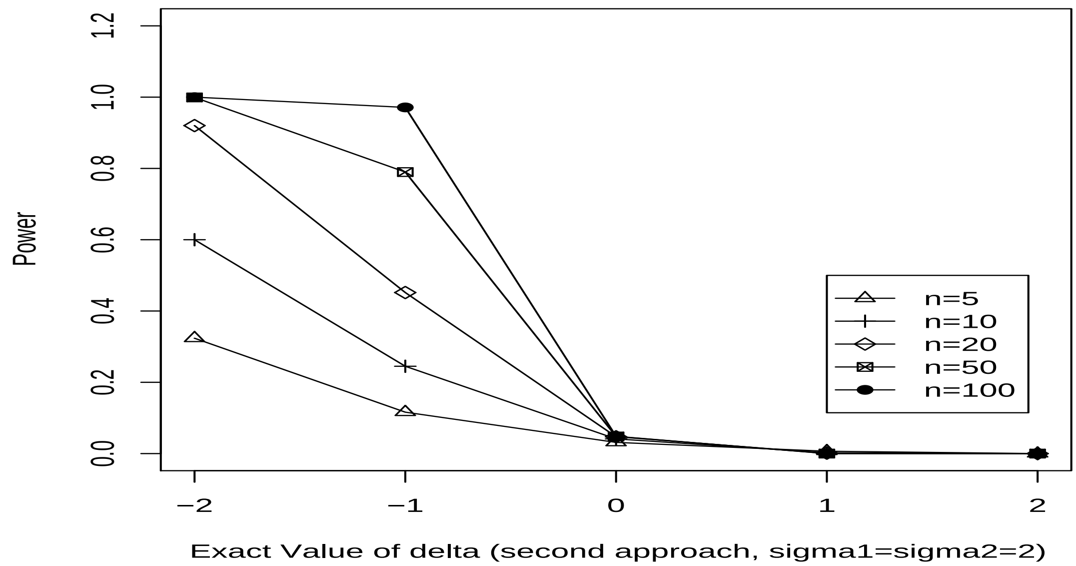

FIGURE 5.9. Example 5.10: The simulated powers with $\sigma_1 = \sigma_2 = 2$ (SA)

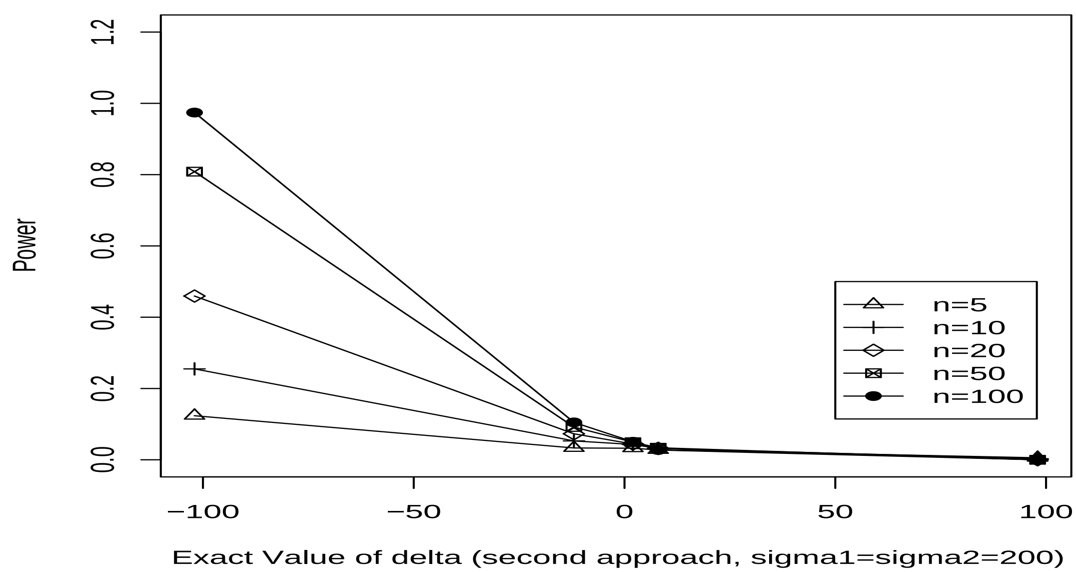

FIGURE 5.10. Example 5.10: The simulated powers with $\sigma_1 = \sigma_2 = 200$ (SA)

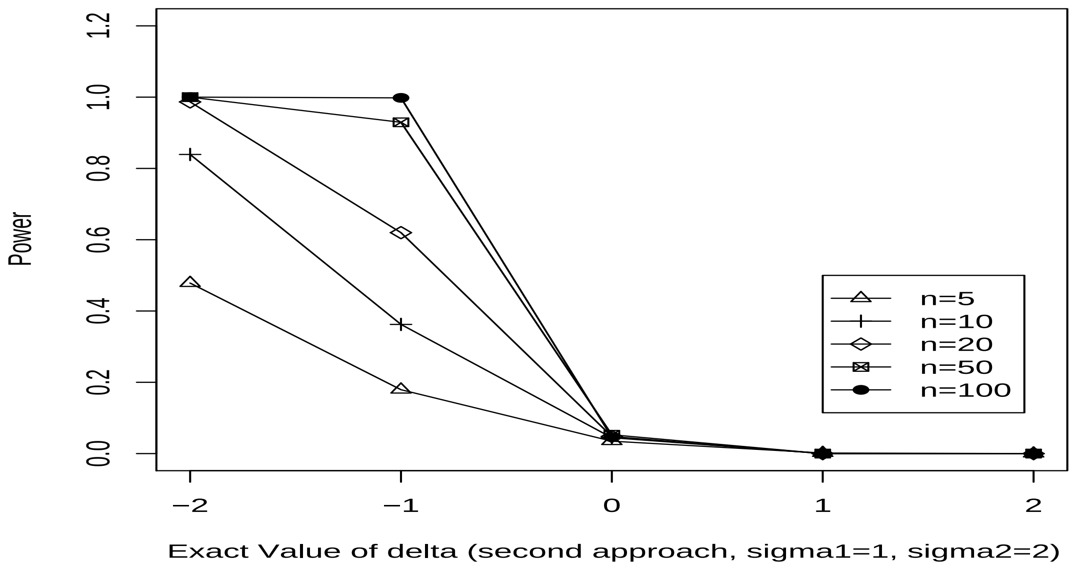

FIGURE 5.11. Example 5.10: The simulated powers with $\sigma_1 = 1$, $\sigma_2 = 2$ (SA)

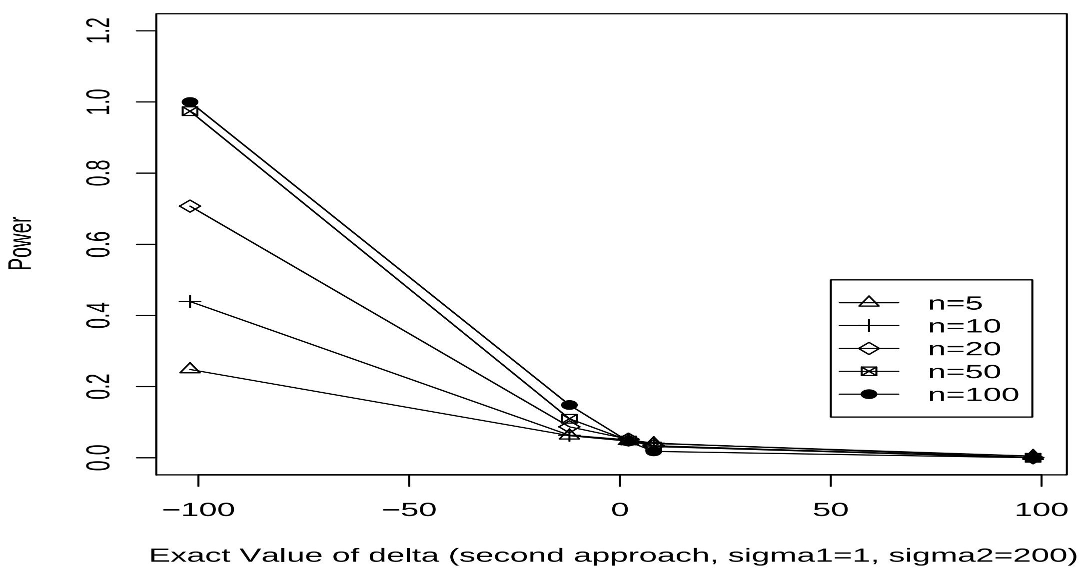

FIGURE 5.12. Example 5.10: The simulated powers with $\sigma_1 = 1$, $\sigma_2 = 200$ (SA)

5.4. Applications

In Section 5.3, we simulate the coverage probabilities for the GCI of 3 bivariate problems with respect to normal families. As one can see from the results, the methods provided in this thesis perform well, especially for the case of equal sample sizes. Therefore, in the following examples, we would like to apply the methods to some real data sets.

Example 5.11: Air lead levels (univariate case).

This data set is available in Krishnamoorthy, Mathew, and Ramachandran (2006). The data contains 15 different air lead levels, which were collected for health hazard evaluation purpose on February 23, 1989. It is already confirmed that the given sample does not follow normal distribution but lognormal distribution. In this case, we assume the air lead levels sample $X_1 \sim Lognormal(\mu, \sigma)$. Then, by taking the log-transformation to X_1, it can be verified that the logged data are normal distributed. That is, $X = \log(X_1) \sim \mathcal{N}(\mu, \sigma)$. Then, the computation results are given in the Table 5.19. From Table 5.19, the Pitman estimator for μ is 4.332862, and the .95

TABLE 5.19. Computation results of Example 5.11

Parameters of interest	Pitman estimator	0.95 GCI
μ	4.332862	(3.405445, 5.310342)
σ	1.708681	(0.9931368, 2.5532461)

GCI is (3.405445, 5.310342). Since the lower bound of 0.95 GCI is higher than 0, this indicates the fact that the exact value of μ is significant different from 0. Further, the Pitman estimator for σ is 1.708681, and the 0.95 GCI is (0.9931368, 2.5532461), which indicates that the difference between σ and 1 is not statistically significant at level 5%.

Example 5.12: Normal Body Temperature (bivariate case).

These data are derived from a dataset presented in Mackowiak, Wasserman, and Levine (1992). In this data set, a total number of 130 patients have been assigned, with 65 males and 65 females. Their body temperatures have been tested and recorded. Furthermore, it is already confirmed that the temperatures in these 2 gender groups are normal distributed, respectively. In this case, we assume male group $X \sim \mathcal{N}(\mu_1, \sigma_1)$ and female group $Y \sim \mathcal{N}(\mu_2, \sigma_2)$. Then, our interest is to use the method provided in this paper to compute the generalized confidences for ρ and δ of this data set. Then, by using the methods provided, the results are presented in Table 5.20. From Table 5.20, the 0.95 GCI for ρ is (0.6848705, 1.3106982), in which

TABLE 5.20. Computation results of Example 5.12

Parameters of interest	Pitman estimator	0.95 GCI
μ_1	98.1	(97.92873, 98.28112)
σ_1	.696	(0.5283083, 0.8186700)
μ_2	98.39	(98.20655, 98.57497)
σ_2	0.74	(0.5623908, 0.8679169)
ρ	1.064	(0.6848705, 1.3106982)
δ (Second approach)	-.2892308	(-0.54288915, -0.03725161)

the value 1 is included. This implies that we failed to reject that hypothesis that $\sigma_1 = \sigma_2$, in other words, the ratio of scale parameters is equal to 1. In addition, for testing H_0: $\rho \geqslant 1$, the GPV is 0.6912, which indicates that the null hypothesis H_0 is not rejected.

However, the 0.95 GCI for δ is $(-0.54288915, -0.03725161)$ based on the second approach. Since both intervals do not contain 0, these imply that there is a significant different between two location parameters. By using (59) for one-sided testing problem H_0: $\delta \geqslant 1$ versus H_0: $\delta < 1$, the GPV is 0.0133 from the second approach. These results indicate the null hypothesis should be rejected at 2% significant level, i.e. this confirms that $\mu_1 < \mu_2$.

Example 5.13: Cloud Seeding (bivariate case).

The data set are available in Krishnamoorthy, and Mathew (2003). The amount of rainfall (in acre-feet) from 52 clouds were recorded. In this data, 26 clouds were randomly seeded with silver nitrate, while the rest were not. The above quoted authors already confirmed that lognormal models fit the data sets very well. In this case, we assume unseeded cloud group $X_1 \sim \mathit{Lognormal}(\mu_1, \sigma_1)$ and seeded cloud group $Y_1 \sim \mathit{Lognormal}(\mu_2, \sigma_2)$. Then, by taking the log-transformation to these 2 data sets, the logged data are confirmed to be normal distributed. That is, $X = \log(X_1) \sim \mathcal{N}(\mu_1, \sigma_1)$ and $Y = \log(Y_1) \sim \mathcal{N}(\mu_2, \sigma_2)$. Then, the computation results are given in the Table 5.21.

TABLE 5.21. Computation results of Example 5.13

Parameters of interest	Pitman estimator	GCI
μ_1	3.990406	(3.325968, 4.641512)
σ_1	1.625515	(1.071410, 2.148553)
μ_2	5.134187	(4.496115, 5.775415)
σ_2	1.583602	(1.019519, 2.084743)
ρ	0.9742157	(0.4540212, 1.3706670)
δ (Second approach)	-1.143781	(-2.073196, -0.2211248)

It is shown in Table 5.21 that, the Pitman estimator for ρ is 0.9742157, and the .95 GCI is (0.4540212, 1.3706670), and accordingly, one cannot reject the null hypothesis that the ratio of scale parameters is equal to 1. Further, for testing H_0: $\rho \geqslant 1$, the GPV is 0.4548, which indicates that the null hypothesis is not rejected at 5% significance level.

However, the inference results of GCI for δ indicate that there is a significant difference between the two location parameters. In addition, for testing H_0: $\delta \geqslant 0$, the GPV is 0.007, based on (59). These results indicate that $\mu_1 < \mu_2$. This confirms the result provided in Krishnamoorthy and Mathew (2003), in which the author applied the two-sample t-test for the logged data and concluded that μ_1 is not equal to μ_2.

CHAPTER 6

Conclusion and future research

In this thesis, we are interested in developing the general procedures of constructing the GCI and GPV in location-scale family. The suggested approach is based on equivariant estimator and thus, the approach improves the existing methods given in the literature which are based on MLE.

In particular, the established GPQ and GTV are functions of the Pitman estimators which are the minimum risk equivariant estimators, and thus more general and more efficient than MLE. Indeed, as mentioned in this thesis, the suggested procedure is applicable to some location-scale families where MLE does not exist. As a preliminary step, we start by establishing the procedure for the univariate (one-sample) case. Further, extend the methods to the bivariate (two-samples) case. In this case, similar to the univariate case, we solve some inference problems concerning the location and scale parameters. Namely, we establish GCI and GPV for the ratio, ρ, between two scale parameters as well as GCI and GPV for the difference, δ, between two location parameters. For this last problem, we distinguish the case where the ratio between the scale parameters is known to the case where the ratio between the scale parameters is unknown. Thus, we suggest a solution to the classical Behrens-Fisher problem.

Methodologically, two approaches are presented in bivariate case. For the inference problem with respect to δ, the second approach is applicable to the case that ρ is unknown. This is more advanced than the first one which can be applied to the case of known ρ only.

To illustrate the performance of the provided methods, we apply our methods into some well-known location-scale families, such as normal, Cauchy, Logistic, and bi-normal. In these examples, the simulated coverage probabilities are presented. The simulation studies confirm that the suggested method performs well in most cases.

Finally, it should be noticed that some problems are still needed to be solved. For instance, the methods discussed in this thesis are only applicable to the univariate or bivariate location-scale families. Indeed, the multivariate case which contains more than 2 sample groups is beyond the scope of this thesis. In addition, further research are needed to handle the problem related to the numerical computations, especially for the case where the GPQ and GTV do not have closed form. In this case, due to the heavy trail and multiple dimensions of the integrations, the numerical computations are time consuming, particularly when the samples sizes are large. These problems will be investigated in future research.

APPENDIX A

Appendix

A.1. Some theoretical results

A.1.1. Proof of Theorem 1.1.

Since $g(x)$ is a pdf, $g(x) \geq 0$ for all value of x. So $\frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right) \geq 0$ for all value of x, μ, and $\sigma > 0$. In addition, let $y = \frac{x-\mu}{\sigma}$, we can verify that:

$$\int_{-\infty}^{\infty} \frac{1}{\sigma} g\left(\frac{x-\mu}{\sigma}\right) dx \quad = \quad \int_{-\infty}^{\infty} g(y) dy = 1,$$

which implies that $f(x|\mu, \sigma)$ is a pdf.

A.1.2. Proof of Proposition 3.8.

From (31), the conditional density of z_3 is

$$\begin{aligned} f(z_3|a) &= \frac{\displaystyle\int_0^{\infty} z_4^{n-1} \prod_{i=1}^{n} g((z_3 + a_i) z_4) dz_4}{\displaystyle\int_0^{\infty} \int_{-\infty}^{\infty} z_4^{n-1} \prod_{i=1}^{n} g((z_3 + a_i) z_4) dz_3 dz_4} \\ &= \frac{\displaystyle\int_0^{\infty} z_4^{n-1} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (z_3 + a_i)^2 z_4^2\right) dz_4}{\displaystyle\int_{-\infty}^{\infty} \int_0^{\infty} z_4^{n-1} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (z_3 + a_i)^2 z_4^2\right) dz_4 dz_3}. \end{aligned}$$

Let $\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i$ and $s^2 = \sum_{i=1}^{n}(a_i - \bar{a})^2$. It can be verified that $\bar{a} = 0$. The density function of z_3 is transformed to:

$$\begin{aligned} f(z_3|a) &= \frac{\int_0^\infty z_4^{n-1} \exp\left(-\frac{1}{2} z_4^2 s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right)\right) dz_4}{\int_{-\infty}^\infty \int_0^\infty z_4^{n-1} \exp\left(-\frac{1}{2} z_4^2 s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right)\right) dz_4 dz_3} \\ &= \frac{\int_0^\infty z_4^{n-1} \exp\left(-\frac{1}{2} z_4^2 s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right)\right) dz_4}{\int_{-\infty}^\infty \int_0^\infty z_4^{n-1} \exp\left(-\frac{1}{2} z_4^2 s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right)\right) dz_4 dz_3}. \end{aligned}$$

Let

$$I(z_3) = \int_0^\infty z_4^{n-1} \exp\left(-\frac{1}{2} z_4^2 s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right)\right) dz_4,$$

we have

$$f(z_3|a) = \frac{I(z_3)}{\int_{-\infty}^\infty I(z_3) dz_3}. \tag{66}$$

Further, let

$$r = s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right) \qquad \text{and} \qquad t = z_4^2 s^2 \left(1 + \frac{n(z_3+\bar{a})^2}{s^2}\right) = z_4^2 r.$$

One can verify that

$$z_4 = t^{1/2} r^{-1/2} \qquad \text{and} \qquad \frac{\partial z_4}{\partial t} = 1/2 t^{-1/2} r^{-1/2}.$$

Hence,

$$\begin{aligned} I(z_3) &= r^{-\frac{n-1}{2}} \int_0^\infty \frac{1}{2} t^{\frac{n-1}{2}} \exp\left(-\frac{1}{2} t\right) |1/2 t^{-1/2} r^{-1/2}| dt \\ &= r^{-\frac{n}{2}} 2^{n/2-1} \Gamma\left(\frac{n}{2}\right) \int_0^\infty \frac{\left(\frac{t}{2}\right)^{\frac{n-1}{2}}}{2\Gamma\left(\frac{n}{2}\right)} \exp\left(-\frac{1}{2} t\right) dt = r^{-\frac{n}{2}} 2^{\frac{n}{2}-1} \Gamma\left(\frac{n}{2}\right). \end{aligned}$$

Therefore, from (66), we have

$$f(z_3|a) = \frac{r^{-n/2}}{\displaystyle\int_{-\infty}^{\infty} r^{-n/2} dz_3} = \frac{\left(s^2\left(1+\frac{n(z_3+\bar{a})}{s^2}\right)\right)^{-n/2}}{\displaystyle\int_{-\infty}^{\infty}\left(s^2\left(1+\frac{n(z_3+\bar{a})}{s^2}\right)\right)^{-n/2} dz_3},$$

and then

$$f(z_3|a) \quad = \quad \frac{\left(1+\frac{n(z_3+\bar{a})^2}{s^2}\right)^{-n/2}}{\displaystyle\int_{-\infty}^{\infty}\left(1+\frac{n(z_3+\bar{a})^2}{s^2}\right)^{-n/2} dz_3}. \tag{67}$$

In addition, let $u = \dfrac{\sqrt{n(n-1)}(z_3+\bar{a})}{s}$. We have $dz_3 = \dfrac{s}{\sqrt{n(n-1)}}du$ and then,

$$\begin{aligned}\int_{-\infty}^{\infty} I(z_3)dz_3 &= \int_{-\infty}^{\infty}\left(1+\frac{u^2}{n-1}\right)^{-n/2}\frac{s}{\sqrt{n(n-1)}}du \\ &= \frac{s}{\sqrt{n(n-1)}}\sqrt{n}B(1/2,(n-1)/2).\end{aligned}$$

Therefore, from (66) and (67), we get

$$f(z_3|a) \quad = \quad \frac{\left(1+\frac{n(z_3+\bar{a})^2}{s^2}\right)^{-n/2}}{\frac{1}{\sqrt{n(n-1)}}s\sqrt{n}B(1/2,(n-1)/2)}. \tag{68}$$

Further, let $U = \dfrac{\sqrt{n(n-1)}(Z_3+\bar{a})}{S}$, where given $a_1, a_2, ..., a_n$, the conditional pdf of Z_3 is given in (68). The conditional pdf of U given $a_1, a_2, ..., a_n$ is

$$\begin{aligned}f_U(u|a) &= f\left(\frac{su}{\sqrt{n(n-1)}}-\bar{a}|a\right)\frac{s}{\sqrt{n(n-1)}} \\ &= \frac{1}{\sqrt{n}B(\frac{1}{2},\frac{n-1}{2})}\left(1+\frac{u^2}{n-1}\right)^{-n/2}, \qquad u \in R.\end{aligned}$$

Note that the last term is the pdf of t distribution with $n-1$ d.f. Since $\bar{a}=0$, this implies that

$$U = \frac{\sqrt{n(n-1)}(Z_3+\bar{a})}{S} = \frac{\sqrt{n(n-1)}(\frac{\hat{\mu}-\mu}{\hat{\sigma}})}{S}$$

follows $\mathcal{T}_{n-1}$ distribution, which complete the proof of (i).

Similarly, one can verify that

$$S^2 Z_4^2 | a = \frac{S^2 \hat{\sigma}^2}{\sigma^2} \Big| a_1, ...a_n \sim \mathcal{X}_{n-1}^2,$$

which complete the proof of (ii).

A.1.3. Simplify of $f(d|v,a,b)$ (ρ unknown) in Example 4.1.

In the proof of Proposition 4.1, conditionally to v, we verify that (53) can be transformed to

$$C_5 \int_0^\infty \int_{-\infty}^\infty t_3^{n+m-1} \exp\left[-\frac{1}{2}\sum_{i=1}^n \left((\rho v u + d)\left(\sqrt{1+\rho^2 v^2}\right)^{-1} + a_i\right)^2 t_3^2\right]$$
$$\times \exp\left[-\frac{1}{2}\sum_{j=1}^m \left((u - \rho v d)\left(\sqrt{1+\rho^2 v^2}\right)^{-1} + b_j\right)^2 v^2 t_3^2\right] du dt_3$$

and this can be rewritten as

$$C_5 \int_0^\infty \int_{-\infty}^\infty t_3^{n+m-1}$$
$$\times \exp\left[-\frac{1}{2}\left(t_3^2\left(\sum a_i^2 + \sum b_j^2 v^2 + \frac{mnv^2 d^2 (1+\rho^2 v^2)}{(n\rho^2 v^2 + mv^2)} + o^2\right)\right)\right] du dt_3,$$

where

$$o = \left(u(n\rho^2 v^2 + mv^2) + (nd - mv^2 d)\rho v\right)\left((1+\rho^2 v^2)(n\rho^2 v^2 + mv^2)\right)^{-.5}.$$

Let

$$k = \left(n\sum a_i^2 + m\sum b_j^2 v^2 + \frac{mnv^2 d^2 (1+\rho^2 v^2)}{(n\rho^2 v^2 + mv^2)}\right).$$

Then, by transforming u to o, we have

$$\begin{aligned} f(d|v,a,b) &= C_{51}\int_0^\infty t_3^{n+m-2}\exp\left[-\frac{1}{2}(t_3k)^2\right]\left[\int_{-\infty}^\infty \frac{\exp\left(-\frac{1}{2}(ot_3)^2\right)}{\sqrt{2\pi}}d(ot_3)\right]dt_3 \\ &= C_{51}\int_0^\infty t_3^{n+m-2}\exp\left[-\frac{1}{2}(t_3k)^2\right]dt_3 \\ &= C_6 k^{\frac{n+m-1}{2}}\int_0^\infty \frac{(t_3k)^{2^{\frac{n+m-1}{2}}-1}}{2^{\frac{n+m-1}{2}}\Gamma\left(\frac{n+m-1}{2}\right)}\exp\left[-\frac{1}{2}(t_3k)^2\right]d[(t_3k)^2], \end{aligned}$$

and then,

$$f(d|v,a,b) = C_6\left[\sum_{i=1}^n a_i^2 + \sum_{j=1}^m b_j^2 v^2 + \frac{mnv^2d^2(1+\rho^2v^2)}{(n\rho^2v^2+mv^2)}\right]^{-\frac{1}{2}(m+n-1)},$$

where C_{51} and C_6 are the components which do not contain t_3, u, and d.

Bibliography

[1] Bain, L. J. (1978). Statistical analysis of reliability and life-testing models. *Marcel Dekker. Inc.*.

[2] Bebu, I., and T. Mathew. (2007). Comparing the means and variances of a bivariate log-normal distribution. *Statist. Med.*, 27 (14): 2684-96.

[3] Billingsley, P. (1995). *Probability and measure*. 3rd ed. NY: John Willey and Sons.

[4] Casella, G., and R. L. Berger. (2001). *Statistical inference*. 2nd ed. CA: Duxbury Press.

[5] David, H.A. (2003). The history of statistics in the class room. *Joint Statistical Meetings 2003-2005*.

[6] Fisher, R. A. (1925). Theory of statistical estimation. *Proc. Camb. Phil. Soc.*, 22: 700-725.

[7] Gabriela V., C. F. (2007). The Pitman estimator of the Cauchy location parameter. *Journal of Statistical Planning and Inference*, 137: 1900-13.

[8] Gupta, A. K., and G. J. Székely. (1994). On location and scale maximum likelihood estimators. *Proceedings of the American Mathematical Society*, 120 (2): 585-589.

[9] Hass, G., L. Bain, and C. Antle. (1970). Inferences for the Cauchy distribution based on maximum likelihood estimators. *Biometrika*, 57: 403-8.

[10] Krishnamoorthy, K., and T. Mathew. (2003). Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. *Journal of Statistical Planning and Inference*, 115: 103-121.

[11] Krishnamoorthy, K., T. Mathew, and G. Ramachandran. (2006). Generalized p-values and confidence intervals: a novel approach for analyzing lognormally distributed exposure data. *Journal of Occupational and Environmental Hygiene*, 3: 642-650.

[12] Krishnamoorthy, K., T. Mathew, and G. Ramachandran. (2007). Upper limits for exceedance probabilities under the one-way random effects model. *Ann. Occup. Hyg.*, 51 (4): 397-406.

[13] Lawless, J.F. (1972). Conditional confidence interval procedures for the location and scale parameters of the Cauchy and logistic distributions. *Biometrika*, 59 (2): 377-386.

[14] Lehmann, E.L., and G. Casella. (1998). *Theory of point estimation.* 2nd ed. NY: Springer.

[15] Mackowiak, P. A., S. S. Wasserman, and M. M. Levine. (1992). A Critical appraisal of 98.6 degrees F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich. *Journal of the American Medical Association*, 268: 1578-1580.

[16] McNally, R.J., H. Lyer, and T. Mathew. (2003). Tests for individual and population bioequivalence based on generalized p-values. *Statist. Med.*, 22: 31-53.

[17] Pitman, E.J.G. (1979). *Some basic theory for statistical inference.* London: Chapman and Hall Ltd.

[18] Pitman, E.J.G. (1939). The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika*, 30: 391-421.

[19] Schervish, M.J. (1997). *Theory of Statistics.* NY: Springer.

[20] Sprott, D.A. (2000). *Statistical inference in science.* NY: Springer.

[21] Tsui K., and S. Weerahandi. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association*, 84 (406): 602-607.

[22] Weerahandi, S. (1993). Generalized confidence intervals. *Journal of the American Statistical Association*, 88 (423): 899-905.

Vita Auctoris

Mr. Fuqi Chen was born in 1984, Fujian, China. He got his Bachelor degree in Mathematics and Applied Mathematics from Tianjin Polytechnic University. In 2007, he began his graduate study in Statistics at University of Windsor. He is expected to graduate with a M.Sc. degree in Summer 2009.