THE UNIVERSITY OF ALBERTA

ON NONPARAMETRIC INFERENCE FOR THE COMPARISON OF SOME MEASURES OF INCOME INEQUALITY

BY

© MARILOU O. HERVAS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE

DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY

EDMONTON, ALBERTA

FALL 1994

ISBN 0-315-95046-3

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: MARILOU O. HERVAS

TITLE OF THESIS: ON NONPARAMETRIC INFERENCE FOR THE COMPARISON OF SOME MEASURES OF INCOME INEQUALITY

DEGREE: MASTER OF SCIENCE

YEAR THIS DEGREE GRANTED: 1994

(Signed) Marilou Hervas

PERMANENT ADDRESS:
38 B Pansol Street, Loyola Heights,
Quezon City, Philippines.

DATE: July 12, 1994

THE UNIVERSITY OF ALBERTA

THE FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled ON NONPARAMETRIC INFERENCE FOR THE COMPARISON OF SOME MEASURES OF INCOME INEQUALITY submitted by MARILOU HERVAS in partial fulfilment of the requirements for the degree of MASTER OF SCIENCE in STATISTICS AND APPLIED PROBABILITY.

Emad-Eldin Aly (Supervisor)

Douglas Wiens

Sudarshan Kumar Sehgal

DATE:

Dedicated to

my parents

EMMY AND RAFAEL HERVAS

Abstract

We consider the problem of estimating and comparing inequality in income distributions. In particular, we develop asymptotically distribution-free procedures for testing against intersecting Lorenz curves. We also propose and study an asymptotically distribution-free nonparametric estimator for a measure of income inequality introduced by Zenga.

Acknowledgement

I would like to express my sincere gratitude to Professor E.E. Aly for proposing the problems considered in this thesis and for his involvement in the various stages of this work. I would like to thank him for his words of wisdom and advice which have been very invaluable to me.

I wish to thank my professors for their kind advice during my graduate study at the University of Alberta.

I also wish to express my sincerest gratitude to the staff and my fellow graduate students who have made my stay in the Department meaningful and worthwhile.

I would like to thank my relatives, friends and roommates especially Ritz, Joy, Sharon, Francis, Martee, Denise, James, Angelo Jessie, Kyung, Monice, Annie, Lola Josefa and Shaun for their prayers and their patience in dealing with me during the crucial stage of my thesis. Their untiring moral support, comments and friendly chats will always be appreciated.

Most importantly, I would like to thank God for making me realize the goodness in what I was achieving.

Contents

1 Introduction **1**

1.1 Definitions and Terms 1

1.1.1 The Lorenz Curve 1

1.1.2 The Generalized Lorenz Curve 2

1.2 Economic Concepts 2

1.2.1 Social Evaluation Criteria 4

1.2.2 Reliability and Economics 7

2 On Testing for 'Intersecting' Lorenz Curves **10**

2.1 Introduction 10

2.2 One-Sample Case 11

2.2.1 Testing H_o against H_{1p} (Known Crossing Point p) 12

2.2.2 Testing H_o versus H_1 (Unknown Crossing Point p) 18

2.3 Two-Sample Case 19

2.4 Simulations 24

2.4.1 One-Sample Case 24

2.4.2 Two-Sample Case 25

3 On Testing for Generalized Lorenz Ordering **27**

3.1 Introduction 27

3.2 One Sample Case 28

3.3 Two-Sample Case 32

4 An Alternative Index of Income Inequality **36**

4.1 Introduction 36

4.2 The K-PCM 40
4.2.1 The K-PCM as derived from Z-PCM 40
4.3 Comparison between the K and Z Curves 42
4.4 On Theoretical Results 51
4.4.1 A Nonparametric Estimator of ξ 51
4.4.2 Asymptotic Distribution of $\hat{\xi}_n$ 53
4.4.3 Nonparametric Estimator for $\sigma_\xi^2(F)$ 58
4.4.4 Exact Variance Results 62
4.5 Simulations 64
4.5.1 Monte Carlo Simulation - Direct Approach 64
4.5.2 Monte Carlo Simulation - Quantile Density Approach 67
4.6 Interpretation and Analysis of Results 68

ON NONPARAMETRIC INFERENCE FOR THE COMPARISON OF SOME MEASURES OF INCOME INEQUALITY

Foreword

The theory and practice of inequality measurement is a rich source of economic literature which dates back a century ago with the pioneering contributions of Pareto [42], Lorenz [39] and Gini [27]. Each provided equally valuable insights which paved the way for a new and fascinating field of quantitative research and policy implications on income inequality. A second wave of economists, namely, Atkinson [4], Kolm [37], Sen [49] and Shorrocks [51] brought the normative concept of income inequality into perspective. They either proposed alternative measures of income inequality based on this normative empirical criteria or provided a stronger environment for the ranking of income distributions by imposing certain constraints on the social welfare function. While the statistical contributions to this field of research have somehow lagged behind the economic conceptual framework, these past two decades saw the emerging awareness among economists of the importance of statistical methodologies in this area.

The basic groundwork for this research rests on the curve introduced by M.O. Lorenz and now named after him. The Lorenz curve is frequently used to describe and compare inequality in income or wealth distribution. The Lorenz curve also underlies social welfare rankings of alternative distributions and is the basis of several summary measures of income or wealth inequality, the most popular of which is the Gini concentration coefficient (Gini [27]). The complete Lorenz Curve allows one to look at the detailed structure of inequality and to identify those regions of a distribution where significant inequality differences occur. Prior to 1981, Lorenz curves have essentially been used as descriptive devices, the reasons attributable to several factors (see Nygard and Sandstrom [41] for a complete discussion). For instance, applied researchers prefer to work and base inequality comparisons on standard summary measures for which confidence intervals have been worked out (see Gastwirth and Gale [25]) and Kakwani [32]. Partly another deficiency in the use of the Lorenz Curve for formal statistical inference is due to methodologies which require an assumption that data come from a specified distribution. A prior assumption such as this is likely not to be true in general. Sendler [50] provided the impetus towards asymptotically distribution-free procedures for Lorenz Curves. He considered the problem of estimating the theoretical Lorenz curve from data. Beach and Davidson [9], Beach and Richmond [10], Bishop et. al.[11], Gastwirth and Gale [26] and Richmond [45] advance on these results. Gastwirth [25] proposed scale-free tests for exponentiality based on the Lorenz curve and Gini statistic. Chandra and Singpurwalla [13] stated a weak convergence result for empirical Lorenz processes. It was Goldie [28] who provided a thorough convergence for empirical Lorenz and what he calls concentration processes.

Chandra and Singpurwalla [13] heightened this momentum by introducing an interesting relationship between reliability and the Lorenz curve and were the first to point out the relation between the total time on test (TTT), which is heavily used in reliability, and the theoretical Lorenz curve. Csörgő et. al. [14] built up a unified asymptotic theory for empirical TTT, Lorenz and In this thesis, we formulate statistical procedures which can be used as tools for formal statistical inference in the area of income inequality.

The thesis is divided into four chapters. Chapter 1 discusses the terms, concepts, economic rationale and motivation used throughout the paper. In Chapter 2, we formulate hypothesis testing procedures for intersecting Lorenz curves. 'Crossing-Over' Lorenz curves are common in empirical work, hence we attempt to devise asymptotically distribution-free procedures to account for such real-world situations. In Chapter 3, we will adopt Aly's [2] hypothesis-testing procedure for the Lorenz Curve to the Generalized Lorenz curve suggested by Shorrocks [51]. The use of generalized Lorenz curves in ranking income distributions generally results in more income distributions being ranked. Finally, in Chapter 4, we present asymptotically distribution-free methodologies for the concentration index Z introduced by Zenga [54].

Chapter 1

Introduction

The economic and statistical basis of the researches in this thesis are discussed in this chapter.

1.1 Definitions and Terms

1.1.1 The Lorenz Curve

Let $F(\cdot)$ be the distribution function of a nonnegative random variable with positive finite mean μ,i.e. $F(0-) = 0$ and $\int_0^\infty x\,dF(x) = \mu$, $0 \le \mu < \infty$. The Lorenz curve (LC) of F is defined for $0 < p < 1$ as

$$L(p) = \frac{1}{\mu}\int_0^p F^{-1}(t)dt, \tag{1.1}$$

where $F^{-1}(y) = \inf\{x : F(x) > y\}$ is the right continuous inverse of the right continuous $F(\cdot)$ and an integral with endpoints 'a' and 'b' means integration over the interval [a,b).

Let $x_1, x_2, \ldots, x_n$ be a random sample from F. An empirical analogue of $L(p)$ is:

$$L_n(p) = \frac{1}{\bar{x}n}\sum_{i=1}^{[np]} x_{i:n}, \tag{1.2}$$

where $x_{1:n}, x_{2:n}, \ldots, x_{n:n}$ are the order statistics of the X sample, $[t]$ denotes the integer part of t and $\bar{x}$ is the sample mean. We also employ the notation $L_x(p)$ to denote the empirical Lorenz curve for the income vector x.

When the income units are arranged in increasing order of their incomes, the Lorenz curve is then the locus of points $(p, L(p))$ where $L(p)$ is the proportion of total wealth accounted for by the $p^{th}\%$ of the poorest individuals. The curve is an increasing convex function with endpoints $L(0) = 0$ and $L(1) = 1$.

The 45^o line called the complete-equality or *egalitarian* line corresponds to $L(p) = p$ in the unit square. Typically, a Lorenz curve is bow - shaped below this line and income inequality is said to increase as the bow is bent more. Note that the Lorenz curve is scale invariant, that is, it is not susceptible to the particular monetary denomination used to measure income.

1.1.2 The Generalized Lorenz Curve

The generalized Lorenz (GL) curve of F is simply defined as the Lorenz curve scaled up by the mean of the distribution, that is, for $0 < p < 1$,

$$GL(p) = \mu L(p),$$

where μ is the mean of the distribution F and $L(\cdot)$ is the corresponding Lorenz curve. The GL has the same property as the ordinary LC except that the GL curve is not invariant to the monetary denomination used to measure income. It is continuous, convex and nondecreasing in the unit interval. It starts at the origin $(0,0)$ and ends in $(1, \mu_x)$. The slope of the diagonal is μ_x. The height reflects the levels of incomes, while the curvature indicates the degree of income inequality. The corresponding empirical analogue is:

$$GL_n(p) = \frac{1}{n}\sum_{i=1}^{[np]} x_{i:n}, \tag{1.3}$$

where $x_{i:n}$ is the ith ordered income. We also employ $GL_x(p)$ to denote the empirical Lorenz curve for the income vector x.

1.2 Economic Concepts

Basic Economic Definitions

1. Utility function is a formula showing the total satisfaction of consuming a particular commodity bundle.
2. Social welfare function $W(\cdot)$ (SWF) is a function that relates the welfare of a society to the utilities of its members. Algebraically, if there are n individuals in society, W is a function of individual utilities, $W = W(U_1, U_2, \ldots, U_n)$. It is assumed that a change that makes someone better off without making anyone worse off increases social welfare.
3. Pareto improvement is a reallocation of resources that makes at least one person better off without making anyone else worse off.
4. Income is a flow of money earned during a period while wealth is a net stock of assets owned at a point in time.
5. Principle of transfers allows income of some people to fall, provided that incomes of others, who are poorer, increase by at least the same amount.

Measures of income inequality have been addressed by economists and applied researchers to answer a wide range of questions. Is there less inequality in the past than the present year? Can we determine which of two countries have more unequal distribution? Do taxes lead to greater inequality in the distribution of income or wealth? In empirical work one would routinely apply various measures of inequality to come up with answers. For the economist however, it is more natural to begin by considering the ordinal problem of obtaining a ranking of distributions. Herein lies the concept of the social welfare function implicit in the ranking of the income distributions. Development of principles which imply empirical criteria that can be applied to evaluate and compare income distributions are addressed by a number of writers such as, Atkinson [4], Dasgupta, et al. [21], Rothschild and Stiglitz [46], Saposnik [48], Shorrocks [51], and others.

We discuss briefly these normative criteria which we refer to in this paper as the criteria for social ordering or simply social evaluation criteria.

1.2.1 Social Evaluation Criteria

Consider n households, identical in all respects except for income. Let $W(\cdot)$ be the social welfare function (SWF) defined as follows:

$$W = W(z_1, z_2, \ldots z_n) = \sum_{i=1}^{n} U(z_i), \tag{1.4}$$

where z_i is the income of individual i, $U(\cdot)$ is the individual utility function and is increasing and concave. The above definition of SWF implies that W is increasing in incomes, that is, an increase in income implies an increase in individual utilities. As a consequence, a partial ordering implied by social welfare regardless of the form of the individual utility functions is defined. This partial ordering is termed in the economic literature as the Pareto criteria. Given any two income vectors x and y, then for $i = 1, \ldots, n$ individuals,

$$\text{if } x_i \geq y_i \Rightarrow W(x) \geq W(y). \tag{1.5}$$

x is said to be Pareto superior to y if all members in x are better off and no one is worse off compared to y and there is at least one person who is better off (meaning, a Pareto improvement has taken place).

However, in real situations, some individuals will benefit and others will lose as a result of social changes and policy interventions. The effects of these changes on social welfare become ambiguous if we adhere to the Pareto criteria. To reduce the ambiguity calls for a relaxation of the the Pareto criteria. Two constraints are imposed on the SWF.

1. W is symmetric in its argument i.e., $W(z) = W(\Pi z)$, for all permutation matrices Π.

2. W is Schur - concave in individual utilities;

Symmetry (or the anonymity property), means that it is no longer necessary for every individual to be better off under x than under y: if 2 or more individuals swap incomes it makes no difference to social welfare. As a result, by ordering the elements of vectors x and y from smallest to largest, $y_{1:n}$ refers to the poorest ndividual and $y_{n:n}$ to the richest, without reference to the identity of who the poorest or the richest individual might be.

The principle of transfers is widely accepted as the definition of greater equality, i.e., the transfer of income from a richer to a poorer individual increases equality. This is equivalent to assuming that the social welfare function is Schur-concave, that is, if for any bistochastic matrix B, $W(Bz) \geq W(z)$. In the income vector Bz, each individual's income is replaced by a convex combination of all incomes in z, that is, B is a non-negative square matrix for which each row and column sums to unity. Hence, Bz gives a specific form to the general notion that transfers which are progressive reduce the dispersion of incomes by a form of averaging or convex combination of incomes. As such, Schur-concavity reflects a preference for greater equality.

The first assumption of symmetry on W enables us to define the so-called rank dominance criterion.

$$x \geq_R y \Leftrightarrow x_{i:n} \geq y_{i:n} \forall i. \tag{1.6}$$

By imposing the assumptions on W, we present the welfare criteria which allow us to draw valid welfare conclusions in empirical comparisons of income distributions from sample data.

Theorem 1.1 (which is Theorem 1 in Dasgupta [21]) allows us to characterize rank dominance in terms of the Lorenz curves. Theorem 1.2 (which is Theorem 2 in Shorrocks [51]) characterizes rank dominance in terms of the generalized Lorenz curves.

Theorem 1.1 *Let x and y be the two income vectors with corresponding means $\bar{x}$ and*

$\bar{y}$, respectively. Let $L_x(p)(L_y(p))$, $p \in [0,1]$ represent the Lorenz curve corresponding to the distribution x (resp y).

$$\text{If } \bar{x} = \bar{y}, \quad \text{then} \quad W(x) \geq W(y) \;\; \forall \text{ Schur-concave } W(\cdot),$$
$$\iff \quad L_x(p) \geq L_y(p) \;\; \forall p.$$

That is, this is the situation where one distribution x can be shown to be at least as desirable as another distribution y for any $W(\cdot)$. Such a situation arises when two distributions have identical means and non-intersecting Lorenz curves.

The above social evaluation criterion which uses the Lorenz curve as a measure breaks down in situations where $L_x(p) > L_y(p)$ but $\bar{x} < \bar{y}$. This results in inconclusive ordering among pairs of income distributions (see Shorrocks [51]). An alternative is to use the scaled Lorenz curve as a measure to arrive at a social evaluation criterion and invoking the principle of transfers. The latter is also referred to in the literature as the generalized dominance criterion. It is assumed that the monetary denomination used to measure income is uniform across distributions being compared.

Theorem 1.2 *Let x and y be the two income vectors with corresponding means $\bar{x}$ and $\bar{y}$, respectively. Let $GL_x(p)(GL_y(p))$, $p \in [0,1]$ represent the generalized Lorenz curve corresponding to the distribution x (resp y).*

$$\text{If } \bar{x} > \bar{y}, \quad \text{then} \quad W(x) \geq W(y) \;\; \forall \text{ S-concave } W(\cdot),$$
$$\iff \quad GL_x(p) \geq GL_y(p) \;\; \forall p.$$

This means that an unambiguous ranking is obtainable iff the generalized Lorenz curves do not intersect and x has both a higher mean and a higher Lorenz curve, that is, the GL curve of x must lie everywhere above the GL curve of y.

In terms of the empirical evaluation of income distributions, the welfare ordering in Theorem 1.1 is equivalent to the Lorenz ordering and Theorem 1.2 is equivalent to the generalized Lorenz ordering.

The *Lorenz order*, $\leq_L$, is defined as:

$$x \leq_L y \text{ if for all } 0 \leq p \leq 1, L_x(p) \geq L_y(p), \tag{1.7}$$

The *generalized Lorenz order*, $\leq_{GL}$ is defined by:

$$x \geq_{GL} y \text{ if } \bar{x} > \bar{y} \text{ and } GL_x(p) \geq GL_y(p), \tag{1.8}$$

If we let the income vector x (y) correspond to a distribution function F (G, resp.), then the rank dominance criterion can be written as:

$$F \geq_R G \leftrightarrow F^{-1}(p) \geq Q^{-1}(p) \forall p \in (0,1)$$

The rest of the mentioned orderings can thus be written in terms of the population measures.

In retrospect, the partial ordering of income distributions according to the Lorenz criterion is identical with the ordering implied by social welfare if the social welfare function is defined as the sum of individual utility functions and by imposing the anonymity and Schur-concavity assumptions on $W(\cdot)$. By scaling the Lorenz curve, Shorrocks [51] arrived at another criterion to reduce the ambiguity in the ordering of income distributions. In fact, Pyatt [44] commented that the ambiguity in ranking could be removed entirely if we could define a unique measure of desirability for any particular situation so that the ranking of alternative situations was always possible. However, he also pointed out that it is hardly likely that such a measure could be achieved. The reader is referred to the papers of Shorrocks [51], Pyatt [44] and Bishop, et. al. [11].

1.2.2 Reliability and Economics

Reliability theory explores classes of life distributions corresponding to various notions of aging. Each class of life distributions provides a realistic probabilistic description of a physical property occurring in the reliability context. This is important in modeling real-life problems. For example, for life distributions which are IFR (Increasing Failure Rate), failure rate increases with age (see Barlow and Proschan [7] and Hollander and Proschan [30] for thorough discussions on classes of life distributions and their properties).

Guess, Hollander and Proschan [29] proposed two new nonparametric classes of life distributions for modeling aging based on the mean residual life (MRL) - the IDMRL ('increasing initially, then decreasing MRL') and the dual DIMRL ('initially adverse, then beneficial'). MRL is defined as 'given that an item is at age t, the expected value of the random remaining life is called MRL at age t'. Tests for these nonparametric classes of life distributions are in the literature and are derived explicitly in correspondence with the TTT-transform. The total time on test (TTT) - transform has been advocated by Barlow and Campo [8] and others as a useful graphical device to analyze life data. It has likewise been utilized by Klefsjo [35] and others to derive tests of exponentiality versus nonparametric alternatives.

For a life distribution F with finite mean μ, the TTT-transform $T(p)$ of F is defined as

$$\begin{aligned} T(p) &= \int_0^{F^{-1}(p)} (1 - F(s))\, ds, \quad \text{for } 0 \le p \le 1. \\ &= \int_0^p (1 - y)\, dF^{-1}(y) \end{aligned} \tag{1.9}$$

The scaled TTT transform is

$$T_s(p) = \frac{1}{\mu} \int_0^p (1 - y)\, dF^{-1}(y). \tag{1.10}$$

Chandra and Singpurwalla [13] were the first to point out the existence of a close relationship between the TTT and the theoretical Lorenz curve, and in particular, between the various indices associated with these transforms and the Gini index. For instance, the relationship between the scaled TTT and the Lorenz curve is noted as follows.

$$L(p) = -\frac{1}{\mu}(1-p)F^{-1}(p) + T_s(p) \tag{1.11}$$

Chapter 2

On Testing for 'Intersecting' Lorenz Curves

2.1 Introduction

The ranking of income distributions has always been of great interest for researchers and economists alike. Using the Lorenz curve as a descriptive device, one is able to compare inequality in income or wealth distribution. Not only are we able to determine dominance relationships in terms of the income criteria, we are also provided with a means by which to base policy reforms (e.g., tax reforms). Indeed, the implications on the comparison of Lorenz curves are numerous and crucial to decision-making. However, because the theoretical Lorenz curves are estimated based on the sample income data, variations in measurements are likely to occur. These variations may either be inherent or due to random variation. It is therefore relevant that statistical tests be formulated to enable a policy maker to make statistically sound conclusions.

The rank dominance criterion and the generalized dominance criterion are effective welfare criteria for ranking income distributions. However, in a number of situations, it is likely that a dominance relationship cannot be concluded. This is true when the Lorenz curves intersect. For instance, in a study of inter-country comparisons, Shorrocks [51] concluded in his Table 1 that Lorenz curves intersect in at least 108 of the 190 pairwise comparisons between countries, hence barely 40 % of the inequality comparisons generate unambiguous Lorenz ranking. When the generalized Lorenz criterion was utilized, his Table 2 indicates that the generalized Lorenz curves intersect in only 31 of the 190 potential pairwise comparisons. On the basis of this GL curve, dominance is conclusive in 84% of the cases. These situations that result in ambiguous rankings are common. Although a substantial amount of economic literature on the social evaluation criteria have been proposed to obtain comparisons that are conclusive, the problem of 'intersecting' Lorenz curves remain. We have devised a statistical procedure that can be used to test equality of the two income distributions when it is suspected that their corresponding Lorenz curves intersect. Compared to a descriptive numerical comparison, this statistical procedure is hoped to assist the economist in arriving at a more reliable conclusion in situations when the sample Lorenz curves intersect.

In this chapter, the focus is to develop asymptotically distribution-free test procedures against 'crossing-over' Lorenz curves.

We first consider the situation when a sample income distribution is being compared with a known income distribution. We refer to this as the one-sample case. The case when the point of 'intersection' is known is presented in section 2.2.1. For unknown 'crossing-point', the procedure is developed in section 2.2.2. We also consider the situation when two sample income distributions are being compared. We refer to this as the two-sample case.

2.2 One-Sample Case

Let F_o be a completely specified distribution function. Let $\hat{F}(\cdot)$, $\bar{x}$ and $x_{1:n}, x_{2:n}, \ldots, x_{n:n}$ be respectively, the empirical distribution function, the mean and the order statistics of a random sample from $F(\cdot)$.

In the present section, we consider the problem of testing the null hypothesis

$$\mathbf{H_o} : F \stackrel{L}{=} F_o \tag{2.1}$$

against each of the alternatives

$\mathbf{H_{1p}}$: For a given (known) and fixed p $\in (0,1)$ where p is the 'crossing point',

$$L(t) \geq L_o(t) \quad \text{for} \quad 0 \leq t \leq p \quad \text{and} \tag{2.2}$$
$$L(t) \leq L_o(t) \quad \text{for} \quad p \leq t \leq 1$$

$\mathbf{H_1}$: (2.2) holds for p unknown, (2.3)

where $L(\cdot)$ and $L_o(\cdot)$ are the Lorenz curves of F and F_o, respectively. We notice that $F \leq_L F_o$ and $F_o \leq_L F$ if and only if $F(x) = F_o(\theta x)$ for some θ and for all real x, i.e., F is equal to F_o in the Lorenz sense. In this case, we say that $F \overset{L}{=} F_o$. Our test statistics are motivated by Aly(1990) test statistics for $IDMRL$ alternative.

In the above test procedures, we test the Lorenz equality of the two income distributions when it is suspected that their Lorenz curves intersect. This means that we cannot establish a dominance relationship. If the 'crossing-point' is known, rejection of the equality of the two distributions mean that for the poorest p^{th} per cent of the population, the income distribution of the sample exhibits less inequality than the theoretical model. When H_o is rejected for the case when the 'intersection point' is not known, then p can be estimated. In such circumstances, concluding that the sample income distribution intersects a known distribution and hence being able to estimate the 'crossing point' $\hat{p}$, a researcher is able to statistically conclude that the degree of income inequality for the holders of the p^{th} fraction of incomes is less than that for the specified distribution model (refer to Figure 2.1).

2.2.1 Testing H_o against H_{1p} (Known Crossing Point p)

Let F_1 and F_2 be two income distributions and L_1 and L_2 be their Lorenz curves, respectively. Motivated by (2.3) of Aly (1990), we consider

$$\Delta(F_1, F_2; p) \quad = \quad 2\int_0^p \{L_1(t) - L_2(t)\}\, dt - 2\int_p^1 \{L_1(t) - L_2(t)\}\, dt$$

$$= \quad 4\int_0^p \{L_1(t) - L_2(t)\}\,dt - 2\int_0^1 \{L_1(t) - L_2(t)\}\,dt \qquad (2.4)$$

and note that if $F_1 \overset{L}{=} F_2$, $\Delta(F_1, F_2; p) = 0$ and if

$$\begin{aligned} L_1(t) &\geq L_2(t) \quad \text{for} \quad 0 \leq t \leq p \quad \text{and} \\ L_1(t) &\leq L_2(t) \quad \text{for} \quad p \leq t \leq 1 \end{aligned} \qquad (2.5)$$

holds and $F_1 \overset{L}{\neq} F_2$ then $\Delta(F_1, F_2; p) > 0$.

Note that

$$\begin{aligned} \int_0^p L(t)\,dt &= \int_0^p \frac{1}{\mu}\int_0^t F^{-1}(s)\,ds dt \\ &= \frac{1}{\mu}\int_0^p (p-s)F^{-1}(s)\,ds \\ &= \frac{1}{\mu}\left\{\frac{1}{2}p^2 F^{-1}(0) I(F^{-1}(0) \neq 0) + \frac{1}{2}\int_0^p (p-s)^2 dF^{-1}(s)\right\}, \end{aligned} \qquad (2.6)$$

where $I(A)$ is the indicator function of the event A.

By (2.4) and (2.6), we have

$$\begin{aligned} \Delta(F_1, F_2; p) &= \frac{2}{\mu_1}\left\{p^2 F_1^{-1}(0) I(F_1^{-1}(0) \neq 0) + \int_0^p (p-s)^2\,dF_1^{-1}(s)\right\} \\ &\quad -\frac{2}{\mu_2}\left\{p^2 F_2^{-1}(0) I(F_2^{-1}(0) \neq 0) + \int_0^p (p-s)^2\,dF_2^{-1}(s)\right\} \\ &\quad -\frac{1}{\mu_1}\left\{F_1^{-1}(0) I(F_1^{-1}(0) \neq 0) + \int_0^1 (1-s)^2 dF_1^{-1}(s)\right\} \\ &\quad +\frac{1}{\mu_2}\left\{F_2^{-1}(0) I(F_2^{-1}(0) \neq 0) + \int_0^1 (1-s)^2 dF_2^{-1}(s)\right\}, \\ &= \frac{1}{\mu_1}\left\{(2p^2 - 1)F_1^{-1}(0) I(F_1^{-1}(0) \neq 0)\right\} \\ &\quad +\frac{1}{\mu_1}\left\{\int_0^p 2(p-s)^2\,dF_1^{-1}(s) - \int_0^1 (1-s)^2\,dF_1^{-1}(s)\right\} \\ &\quad -\frac{1}{\mu_2}\left\{(2p^2 - 1)F_2^{-1}(0) I(F_2^{-1}(0) \neq 0)\right\} \\ &\quad -\frac{1}{\mu_2}\left\{\int_0^p 2(p-s)^2\,dF_2^{-1}(s) - \int_0^1 (1-s)^2 dF_2^{-1}(s)\right\}. \end{aligned}$$

To conduct the one-sample test of (2.1) and (2.2) we propose a test statistic based on t_{1p}:

$$\triangle(\hat{F}, F_o; p) = \delta(\hat{F}; p) - \delta(F_o; p) = t_{1p}, \tag{2.7}$$

where

$$\delta(F;p) = \frac{1}{\mu}\left\{(2p^2-1)F^{-1}(0)I(F^{-1}(0) \neq 0) + g(F;p)\right\}, \tag{2.8}$$

$$g(F;p) = 2\int_0^p (p-s)^2 dF^{-1}(s) - \int_0^1 (1-s)^2 dF^{-1}(s), \tag{2.9}$$

$$\delta(\hat{F};p) = \frac{1}{\bar{x}}\left\{(2p^2-1)x_{1:n}I(\hat{F}^{-1}(0) \neq 0) + g(\hat{F};p)\right\}, \tag{2.10}$$

$$g(\hat{F};p) = 2\sum_{j=1}^{[np]-1}\left(p - \frac{j}{n}\right)^2 (x_{j+1:n} - x_{j:n}) - \sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)^2 (x_{j+1:n} - x_{j:n}) \tag{2.11}$$

and $I(A)$ is the indicator function of the event A.

In this section we derive the asymptotic distribution of $\delta(\hat{F};p)$. The following definition is needed in the statement of Theorem 2.1

Definition 2.1 *The distribution function $F(\cdot)$ is said to satisfy the Csörgó-Révész tail conditions on the support of F denoted by $(t_F, T_F) = \{x : 0 < F(x) < 1\}$, $-\infty < t_F < T_F \leq \infty$ (Csörgó-Révész [16]) if*

1. On (t_F, T_F), F is twice differentiable and $f(\cdot) = F'(\cdot) > 0$;
2. For some $\gamma_F \epsilon (0, \infty)$, we have

$$\sup_{t_F < x < T_F} F(x)(1-F(x))|f'(x)|/f^2(x) \leq \gamma_F.$$

To be able to state the next result, we define $\sigma^2(F)$ as follows.

$$\sigma^2(F) = \frac{32}{\mu^2}\int_0^p (p-y)(1-y)\int_0^y x(p-x)dF^{-1}(x)dF^{-1}(y)$$

$$
\begin{aligned}
&+2\int_0^1 h(y)(1-y)\int_0^y xh(x)\,F^{-1}(x)dF^{-1}(y) \\
&-\frac{8}{\mu}\int_0^p (p-x)(1-x)\int_0^x yh(y)dF^{-1}(y)dF^{-1}(x) \\
&-\frac{8}{\mu}\int_0^p (p-x)x\int_x^1 (1-y)h(y)dF^{-1}(y)dF^{-1}(x) \qquad (2.12)
\end{aligned}
$$

with

$$
\begin{aligned}
h(s) &= \frac{2(1-s)}{\mu} + \frac{g(F;p)}{\mu^2}; \\
g(F;p) &= 2\int_0^p (p-s)^2 dF^{-1}(s) - \int_0^1 (1-s)^2 dF^{-1}(s).
\end{aligned}
$$

Theorem 2.1 *Assume that $F(\cdot)$ has mean μ, satisfies the Csörgő- Révész tail conditions on (t_F, T_F) and if $F^{-1}(0) \neq 0$, we have $fF^{-1}(0) \neq 0$. Assume, in addition, that $0 < \sigma^2(F) < \infty$, where $\sigma^2(F)$ is as in (2.12).*

Then as $n \to \infty$,

$$
n^{\frac{1}{2}}\left\{\delta(\hat{F};p) - \delta(F;p)\right\} \xrightarrow{D} G(F;p), \qquad (2.13)
$$

where

$$
\begin{aligned}
G(F;p) &= \frac{4}{\mu}\int_0^p (p-s)B(s)dF^{-1}(s) \\
&\quad - \int_0^1 \left\{\frac{2(1-s)}{\mu} + \frac{g(F;p)}{\mu^2}\right\} B(s)dF^{-1}(s), \qquad (2.14)
\end{aligned}
$$

and $B(\cdot)$ is a Brownian Bridge[1].

Proof:

Let $\delta^*(F;p) = \mu\delta(F;p)$, $\delta^*(\hat{F};p) = \bar{x}\delta(\hat{F};p)$ and note that

[1]The process $\{B(x), 0 \leq x \leq 1\}$ is said to be a Brownian Bridge if it is Gaussian with mean zero and covariance function $E[B(s)B(t)] = s \wedge t - st$, where $\wedge$ means minimum(s,t).

$$n^{\frac{1}{2}}\left\{\delta^*(\hat{F};p)-\delta^*(F;p)\right\} = (2p^2-1)n^{\frac{1}{2}}\left\{\hat{F}^{-1}(0)-F^{-1}(0)\right\}I(F^{-1}(0)\neq 0) + n^{\frac{1}{2}}\left\{g(\hat{F};p)-g(F;p)\right\}. \quad (2.15)$$

We consider first the limiting distribution of the second term in (2.15)

$$n^{\frac{1}{2}}\left\{g(\hat{F};p)-g(F;p)\right\} = 2n^{\frac{1}{2}}\left\{\int_0^p (p-s)^2 d\hat{F}^{-1}(s) - \int_0^p (p-s)^2 dF^{-1}(s)\right\} - n^{\frac{1}{2}}\left\{\int_0^1 (1-s)^2 d\hat{F}^{-1}(s) - \int_0^1 (1-s)^2 dF^{-1}(s)\right\}$$

By Theorem 5.1 of Aly(1990),

$$n^{\frac{1}{2}}\left\{g(\hat{F};p)-g(F;p)\right\} \xrightarrow{D} 4\int_0^p (p-s)B(s)dF^{-1}(s) - 2\int_0^1 (1-s)B(s)dF^{-1}(s) \quad (2.16)$$

Next, we consider the limiting distribution of the first term in (2.15). We only need to consider the case $F^{-1}(0)\neq 0$ in which we assume that $fF^{-1}(0)\neq 0$. Write

$$n^{\frac{1}{2}}\left\{\hat{F}^{-1}(0)-F^{-1}(0)\right\} = \frac{n^{\frac{1}{2}}f(F^{-1}(0))(\hat{F}^{-1}(0)-F^{-1}(0))}{f(F^{-1}(0))} \xrightarrow{D} \frac{B(0)}{f(F^{-1}(0))} = 0. \quad (2.17)$$

Hence, as $n\longrightarrow\infty$,

$$n^{\frac{1}{2}}\left\{\delta^*(\hat{F};p)-\delta^*(F;p)\right\} \xrightarrow{D} \text{random variable of the R.H.S. of (2.16)} \quad (2.18)$$

It therefore follows that

$$n^{\frac{1}{2}}\left\{\delta(\hat{F};p)-\delta(F;p)\right\} = n^{\frac{1}{2}}\left\{\frac{g(\hat{F};p)}{x}-\frac{g(F;p)}{\mu}\right\} = \frac{n^{\frac{1}{2}}}{x}\left\{g(\hat{F};p)-g(F;p)\right\} - \frac{g(F;p)}{x\mu}n^{\frac{1}{2}}\left\{x-\mu\right\} \quad (2.19)$$

Since $\mu = \int_0^1 (1-s)dF^{-1}(s)$ and $\bar{x} = \int_0^1 (1-s)d\hat{F}^{-1}(s)$

$$\begin{aligned} n^{\frac{1}{2}}(\bar{x}-\mu) &= n^{\frac{1}{2}}\left\{\int_0^1 (1-s)d\hat{F}^{-1}(s) - \int_0^1 (1-s)dF^{-1}(s)\right\} \\ &\xrightarrow{D} \int_0^1 B(s)dF^{-1}(s) \end{aligned} \tag{2.20}$$

Hence by (2.16),(2.19), (2.20) and Slutsky's Theorem, we have as $n \to \infty$,

$$\begin{aligned} n^{\frac{1}{2}}\left\{\delta(\hat{F};p) - \delta(F;p)\right\} &\xrightarrow{D} \frac{1}{\mu}\left\{4\int_0^p (p-s)B(s)dF^{-1}(s) - 2\int_0^1 (1-s)B(s)dF^{-1}(s)\right\} \\ &\quad -\frac{g(F;p)}{\mu^2}\int_0^1 B(s)dF^{-1}(s) \\ &= \frac{4}{\mu}\int_0^p (p-s)B(s)dF^{-1}(s) \\ &\quad -\int_0^1 \left\{\frac{2(1-s)}{\mu} + \frac{g(F;p)}{\mu^2}\right\} B(s)dF^{-1}(s) \\ &= G(F;p). \end{aligned} \tag{2.21}$$

Note that for every fixed $p \in (0,1), G(F;p)$ is a $N(0,\sigma^2(F))$. This completes the proof of Theorem 2.1.

By the result that $\triangle(F,F_o) = 0$ under H_o of (2.1) and by Theorem 2.1, we have as $n \to \infty$,

$$T_{1p} = n^{\frac{1}{2}}\frac{t_{1p}}{\sigma(F_o)} \xrightarrow{D} N(0,1). \tag{2.22}$$

An asymptotic distribution-free procedure based on (2.22) for testing H_o of (2.1) against H_{1p} of (2.2) is to reject H_o at asymptotic level α if $T_{1p} > z_{1-\alpha}$, where $P(Z \leq z_{1-\alpha}) = 1-\alpha$ (ie, z_p is the p^{th} quantile of Z, $0 < p < 1$). The consistency of the proposed testing procedure follows from the result that if $\triangle(F,F_o) > 0$ then $T_{1p} = T_{1p}^{*} + R_{1p}$, where $R_{1p} = n^{\frac{1}{2}}\frac{\triangle(F,F_o)}{\sigma(F_o)} = O(n^{\frac{1}{2}})$ and $T_{1p}^{*} \stackrel{P}{=} O(1)$.

2.2.2 Testing H_o versus H_1 (Unknown Crossing Point p)

When the 'crossing-point' p is unknown, we propose the test statistic

$$t_1 = \sup_{0<p<1} n^{\frac{1}{2}} \Delta(\hat{F}, F_o; p) = \sup_{0<p<1} n^{\frac{1}{2}} \left(\delta(\hat{F}; p) - \delta(F_o; p)\right). \tag{2.23}$$

Under H_o, we have as $n \to \infty$,

$$t_1 \xrightarrow{D} \sup_{0<p<1} G(F_o; p) = T_{F_o}, \tag{2.24}$$

where $G(F_o; p)$ is as in (2.14). We reject H_o of (2.1) in favor of H_1 of (2.3) if t_1 is large.

The distribution function of T_{F_o} depends on the parent distribution $F_o(\cdot)$. For this reason, we need to obtain limiting critical values of t_1 for each F_o of interest. For any given F_o, the limiting critical values of t_1 can be approximated. $G(F_o; p)$ of can also be approximated for K large, that is,

$$\begin{aligned} G(F_o; i/K) \approx \; & \frac{4}{\mu K} \sum_{j=1}^{i} \frac{(i/K - j/K) Z_j}{f_o F_o^{-1}(j/K)} \\ & - \frac{1}{K} \sum_{j=1}^{K-1} \left\{ \frac{2(1 - j/K)}{\mu} + \frac{g(F_o; i/K)}{\mu^2} \right\} \frac{Z_j}{f_o F_o^{-1}(j/K)}, \end{aligned} \tag{2.25}$$

where

$$\{Z_1, Z_2, \ldots, Z_{K-1}\} \sim MVN(0, \Sigma),$$

$$g(F_o; i/K) = 2 \sum_{j=1}^{i} \frac{(i/K - j/K)^2}{f_o F_o^{-1}(j/K)} - \sum_{j=1}^{K-1} \frac{(1 - j/K)^2}{f_o F_o^{-1}(j/K)}, \text{ and} \tag{2.26}$$

$$\Sigma = [\Sigma_{ij}] \text{ with}$$
$$\Sigma_{ij} = i/K(1-j/K), i \leq j. \tag{2.27}$$

Hence

$$T_{F_o} \overset{D}{\simeq} \max_{1 \leq i \leq K-1} G(F_o; i/K) = \hat{T}_{F_o} \tag{2.28}$$

For any given distribution F_o, we can obtain a table of approximate limiting critical values. In particular for $F_o(x) = 1 - \exp(-x)$, $x \geq 0$, the limiting critical values of t_1 is obtained. The procedure employed is that for large K, say $K = 200$, we compute Σ of (2.27). We simulate $N = 1000$ random variables from the multivariate normal distribution with mean 0 and variance-covariance Σ of (2.27). We calculate $\hat{T}^{(i)}_{F_o}$, for $i = 1, \ldots, N$ of (2.28). The critical values are then obtained by ordering the $\hat{T}^{(i)}_{F_o}$, $i = 1, \ldots, N$. Hence, for $\alpha = 0.1, 0.05$ and 0.01, the limiting critical values are $1.56, 1.78$, and 2.13.

Remarks:

1. When H_o is rejected, we may estimate the 'crossing-point' p by $\hat{p}$ which satisfies

$$\sup_{0<p<1} \left\{ \delta(\hat{F}; p) - \delta(F_o; p) \right\} = n^{\frac{1}{2}} \left\{ \delta(\hat{F}; \hat{p}) - \delta(\hat{F}_o; \hat{p}) \right\}$$

2. Referring to Remark 5.1 of Aly [2]: It is possible to adopt the proofs of Chapter 6 of Csörgó, Csörgó & Horvath to directly prove Theorem 2.1 under the weaker conditions $EX^2 < \infty$ and $0 < \sigma^2(F) < \infty$.

2.3 Two-Sample Case

Similar to the motivation for the one-sample case, in the test procedure for intersecting Lorenz curves when two sample income distributions are being compared, we test the Lorenz equality of the two sample income vectors. Rejection of the equality hypothesis implies that before the 'point of intersection', the Lorenz curve of the first sample income distribution is above the Lorenz curve of the second sample income distribution. After the 'crossing-point', the reverse is apparent. Hence, no dominance relationship is established. If the 'crossing point' is given, then we may conclude that for the poorest p^{th} per cent of the population, there is less inequality in incomes for the first population than the second. For an economist, this type of answer is significant when he is interested in finding out if a certain tax policy has reduced income inequality for the poorest p^{th} per cent by comparing the periods before and after taxation.

Let $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$ be independent random samples from F_1 and F_2, respectively. Let $\hat{F}_1(\cdot)$, $\bar{x}_{n_1}$ and $x_{1:n_1}, \ldots, x_{n_1:n_1}$ (resp. $\hat{F}_2(\cdot)$, $\bar{y}_{n_2}$ and $y_{1:n_2}, \ldots, y_{n_2:n_2}$) be respectively, the empirical distribution function, the mean and the order statistics of the x (resp. y) sample.

We consider the problem of testing the null hypothesis

$$H_o : F_1 \stackrel{L}{=} F_2 \tag{2.29}$$

against the alternative

$$H_2 : \text{ For a given 'crossing-point' } p \in (0,1), \tag{2.30}$$

$$\begin{aligned} L_1(t) &\geq L_2(t) \quad \text{for } 0 \leq t \leq p \text{ and} \\ L_1(t) &\leq L_2(t) \quad \text{for } p \leq t \leq 1, \end{aligned} \tag{2.31}$$

where $L_1(\cdot)$ and $L_2(\cdot)$ are the Lorenz curves of F_1 and F_2, respectively. By the argument in the one-sample case, we propose to use $\triangle(F_1, F_2; p)$ of (2.7) as a measure of the deviation from H_o of (2.29) in favor of H_1 of (2.31).

To conduct the two-sample test, we propose a test statistic based on

$$\Delta(\hat{F}_1, \hat{F}_2; p) = \delta(\hat{F}_1; p) - \delta(\hat{F}_2; p) = t_{n_1,n_2} \tag{2.32}$$

where $\delta(\hat{F}_1;p)$ (resp. $\delta(\hat{F}_2;p)$) is expressed in terms of n_1 and the X's (resp. the n_2 and the Y's) similar to the expression of (2.10).

Let $B_1(\cdot)$ and $B_2(\cdot)$ be two independent Brownian bridges and define,

$$\begin{aligned} G(F_1, F_2; p) \quad = \quad & \left(\frac{n_2}{n_1+n_2}\right)^{1/2} [\frac{1}{\mu_1}\left\{4\int_0^p (p-s)B_1(s)dF_1^{-1}(s)\right\} \\ & -\int_0^1 \left\{\frac{2(1-s)}{\mu_1} + \frac{g(F_1;(p)}{\mu^2}\right\} B_1(s)dF_1^{-1}(s))] \\ & -\left(\frac{n_1}{n_1+n_2}\right)^{1/2} [(\frac{1}{\mu_2}\left\{4\int_0^p (p-s)B_2(s)dF_2^{-1}(s)\right\} \\ & -\int_0^1 \left\{\frac{2(1-s)}{\mu_2} + \frac{g(F_2;p)}{\mu^2}\right\} B_2(s)dF_2^{-1}(s))] \end{aligned} \tag{2.33}$$

Theorem 2.2 *Assume that $F_1(\cdot)$ and $F_2(\cdot)$ have means μ_1 and and μ_2 and satisfy the Csörgó- Révész tail conditions on their supports. Assume further that $0 < \sigma^2(F_1) < \infty$, $0 < \sigma^2(F_2) < \infty$, where $\sigma^2(F)$ is as defined in (2.12).*

Then,

$$\begin{aligned} &|\left(\frac{n_2}{n_1+n_2}\right)^{1/2} n_1^{1/2}\left\{\delta(\hat{F}_1;p) - \delta(F_1;p)\right\} \\ &\quad -\left(\frac{n_1}{n_1+n_2}\right)^{1/2} n_2^{1/2}\left\{\delta(\hat{F}_2;p) - \delta(F_2;p)\right\} - G(F_1,F_2;p)| \xrightarrow{P} 0 \end{aligned}$$

as $min(n_1,n_2) \to \infty$ and $(n_1,n_2) \in D_\lambda := \{(n_1,n_2) : \lambda \le \frac{n_1}{n_1+n_2} \le 1-\lambda\}$ for some $\lambda \sim (0,1/2]$, where $G(F_1,F_2;p)$ is defined in (2.33).

Proof:

Consider

$$
\begin{aligned}
&\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\left\{\Delta(\hat{F}_1,\hat{F}_2;p)-\Delta(F_1,F_2;p)\right\}\\
&= \left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\left\{\delta(\hat{F}_1;p)-\delta(F_1;p)\right\}\\
&\quad -\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\left\{\delta(\hat{F}_2;p)-\delta(F_2;p)\right\}\\
&= \left(\frac{n_2}{n_1+n_2}\right)^{1/2} n_1^{1/2}\left\{\delta(\hat{F}_1;p)-\delta(F_1;p)\right\}\\
&\quad -\left(\frac{n_1}{n_1+n_2}\right)^{1/2} n_2^{1/2}\left\{\delta(\hat{F}_2;p)-\delta(F_2;p)\right\} \qquad (2.34)
\end{aligned}
$$

where $\Delta(F_1, F_2; p)$ and $\delta(F; p)$ are defined as in (2.7) and (2.8), respectively. From (2.34), we can prove that

Hence,

$$
\left|\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\left\{\Delta(\hat{F}_1,\hat{F}_2;p)-\Delta(F_1,F_2;p)\right\}-G(F_1,F_2;p)\right| \xrightarrow{P} 0.
$$

By remark 2 in Section 2.2.2, the Csörgő-Révész tail conditions of Theorem 2.2 can be replaced by the condition that $EX^2 < \infty$ and $EY^2 < \infty$.

The basis of the two-sample test is the following theorem which follows directly from Theorem 2.2 by noting that $G(F_1, F_2; p)$ is a mean zero normal random variable with variance $\sigma^2_{n_1,n_2}$ of (2.36). The complete proof is along the lines in Aly [2].

Theorem 2.3 *Assume that* $EX^2 < \infty$, $EY^2 < \infty$, $0 < \sigma^2(F_1) < \infty$ *and* $0 < \sigma^2(F_2) < \infty$. *Then,*

$$
\left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\frac{(t_{n_1,n_2}-\Delta(F_1,F_2;p))}{\sigma_{n_1,n_2}} \xrightarrow{D} N(0,1) \qquad (2.35)
$$

as $min(n_1, n_2) \to \infty$ and $(n_1, n_2) \in D_\lambda := \{(n_1, n_2) : \lambda \le \frac{n_1}{n_1+n_2} \le 1-\lambda\}$ for some $\lambda \sim (0, 1/2]$, where $\sigma^2(\cdot)$ is as in (2.12) and

$$\sigma^2_{n_1,n_2} = \sigma^2(F_1, F_2) = \{n_1\sigma^2(F_1) + n_2\sigma^2(F_2)\}/(n_1+n_2), \tag{2.36}$$

Next, we define

$$\hat{\sigma}^2_{n_1,n_2} = \sigma^2_{n_1,n_2}(\hat{F}_1, \hat{F}_2) = \{n_1\sigma^2(\hat{F}_1) + n_2\sigma^2(\hat{F}_2)\}/(n_1+n_2), \tag{2.37}$$

where

$$\begin{aligned}
\sigma^2(\hat{F}_1) = {} & \frac{32}{\bar{x}^2}\sum_{j=1}^{[np]-1}\left(p-\frac{j}{n}\right)\left(1-\frac{j}{n}\right)\sum_{i=1}^{j}\frac{i}{n}\left(p-\frac{i}{n}\right)D_{x,i}\,D_{x,j} \\
& +2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)h\left(\frac{j}{n}\right)\sum_{i=1}^{j}\left(\frac{i}{n}\right)h\left(\frac{j}{n}\right)D_{x,i}D_{x,j} \\
& -\frac{8}{\bar{x}}\sum_{j=1}^{[np]-1}\left(1-\frac{j}{n}\right)\left(p-\frac{j}{n}\right)\sum_{i=1}^{j}\left(\frac{i}{n}\right)h\left(\frac{i}{n}\right)D_{x,i}D_{x,j} \\
& -\frac{8}{\bar{x}}\sum_{j=1}^{[np]-1}\left(1-\frac{j}{n}\right)\left(p-\frac{j}{n}\right)\sum_{i=j}^{n-1}\left(1-\frac{i}{n}\right)h\left(\frac{i}{n}\right)D_{x,i}D_{x,j},
\end{aligned}$$

$$\begin{aligned}
h\left(\frac{j}{n}\right) &= \left\{\frac{2}{\bar{x}}\left(1-\frac{j}{n}\right)+\frac{g(\hat{F};p)}{\bar{x}^2}\right\}, \\
g(F;p) &= 2\sum_{j=1}^{[np]}\left(p-\frac{j}{n}\right)^2 D_{x,j} - \sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)^2 D_{x,j},
\end{aligned}$$

where $\bar{x}$ is the mean of the sample, $D_{x,j} = x_{j+1:n} - x_{j:n}$ and $\sigma(\hat{F}_2)$ is defined similarly.

Corollary 2.1 *Assume that H_o of (2.29) holds true. Then, under the conditions of Theorem 2.3 we have*

$$T_{n_1,n_2} = \left(\frac{n_1 n_2}{n_1+n_2}\right)^{1/2}\frac{t_{n_1,n_2}}{\hat{\sigma}_{n_1,n_2}} \xrightarrow{D} N(0,1) \tag{2.38}$$

As a consequence of (2.38) we propose to reject H_o of (2.29) if $T_{n_1,n_2} > z_{1-\alpha}$,ie, z_p is the p^{th} quantile of Z, $0 < p < 1$. The consistency of the proposed testing procedure follows from the result that if $\Delta(F_1, F_2; p) > 0$ then

$$T_{n_1,n_2} = T^*_{n_1,n_2} + R_{n_1,n_2}, \tag{2.39}$$

where

$$T^*_{n_1,n_2} \xrightarrow{D} N(0,1)$$

and

$$R_{n_1,n_2} = \left(\frac{n_1 n_2}{n_1 + n_2}\right)^{1/2} \frac{\Delta(F_1, F_2; p)}{\hat{\sigma}_{n_1,n_2}} \stackrel{P}{=} O\left(\frac{n_1 n_2}{n_1 + n_2}\right)^{1/2}.$$

2.4 Simulations

2.4.1 One-Sample Case

We consider the problem of testing H_o of (2.1) against H_{1p} of (2.2). Let T_{1p} be as in (2.22) and recall that we propose to reject H_o at approximate level α if $T_{1p} > z_{1-\alpha}$.

A Monte Carlo simulation study was conducted to obtain a table of critical values of t_{1p} for each value of $p = 0.1, 0.25, 0.5, 0.75$, and 0.9. For each sample size $n = 50, 100$, the random sample $x_1, x_2, \ldots, x_n$ was generated from the exponential distribution with mean 1 and corresponding test statistic T_{1p} of (2.22) was calculated. This procedure was repeated 1000 times. The resulting percentage points are shown in Table(2.1) for $n = 50$ and in Table(2.2) for $n = 100$. The corresponding standard normal values for $\alpha = 0.10, 0.05$ and 0.1 are 1.2816, 1.6449, and 2.3263, respectively.

Table 2.1: Critical Values of T_{1p} for $n = 50$

Percentage	0.10	0.05	0.01
$p = 0.1$	1.45782	1.79441	2.32465
$p = 0.25$	1.49497	1.92712	2.56977
$p = 0.5$	1.58372	2.04085	2.83499
$p = 0.75$	1.17229	1.54033	2.19678
$p = 0.90$	1.17233	1.48873	2.29634

Table 2.2: Critical Values of T_{1p} for $n = 100$

Percentage	0.10	0.05	0.01
$p = 0.1$	1.42825	1.77570	2.34559
$p = 0.25$	1.47846	1.93438	2.47104
$p = 0.5$	1.46975	1.85528	2.64145
$p = 0.75$	1.17926	1.54951	2.02444
$p = 0.90$	1.22171	1.69560	2.17465

2.4.2 Two-Sample Case

We consider the problem of testing H_o of (2.29) against H_2 of (2.31). Let T_{n_1,n_2} be as in (2.38) and recall that we propose to reject H_o at approximate level α if $T_{n_1,n_2} > z_{1-\alpha}$.

We consider the problem of testing H_o vs. H_2.

A Monte Carlo simulation study was conducted to obtain a table of empirical critical values of $p = 0.1, 0.25, 0.5, 0.75$, and 0.9. For each sample size n_1 and $n_2 = 50, 100$, the random samples $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$ were both generated

Table 2.3: Critical Values of T_{n_1,n_2} for n_i = 50,i=1,2.

Percentage	0.10	0.05	0.01
$p = 0.1$	1.11774	1.42170	2.13231
$p = 0.25$	1.16895	1.50346	2.03769
$p = 0.5$	1.26100	1.61274	2.18157
$p = 0.75$	1.03335	1.37699	2.00665
$p = 0.90$	1.13575	1.48087	2.11772

Table 2.4: Critical Values of T_{n_1,n_2} for n_i = 100,i=1,2.

Percentage	0.10	0.05	0.01
$p = 0.1$	1.15492	1.50200	1.98969
$p = 0.25$	1.21999	1.53851	2.35534
$p = 0.5$	1.23468	1.70766	2.44096
$p = 0.75$	1.17582	1.52253	2.23334
$p = 0.90$	1.11538	1.44979	2.27025

from the exponential distribution with mean 1 and corresponding test statistic T_{n_1,n_2} was calculated. This procedure was repeated 1000 times. The resulting percentage points are shown in Table (2.3) for $n_1 = n_2 = 50$ and in Table (2.4) for $n_1 = n_2 = 100$.

Chapter 3

On Testing for Generalized Lorenz Ordering

3.1 Introduction

In Chapter 2, we developed statistical tests to determine whether the Lorenz curves of two distributions 'cross-over' or not. We have pointed out in Chapter 1 that 'crossing-over' Lorenz curves is an empirical hindrance in the exercise of ranking income distributions since it results in inconclusive ordering. Utilization of the GL curves as tools for ranking provides a clearer ordering of alternatives. Hence, rank dominance is characterized in terms of the generalized Lorenz curves (refer to Theorem 1.2). Shorrocks [51] concluded that by using the generalized Lorenz criterion, he was sucessful in ranking 85% of the possible pairwise comparisons. Bishop, et.al. [11] demonstrates that by applying statistical tests to obtain a dominance relationship between two distributions, the statistical tests result in more unambiguous ranking than mere numerical or descriptive comparison. For example, in his study to compare the income distributions of the states in the U.S. among themselves and with the income distribution of the U.S. as a whole, 32 states can be ranked according to the generalized rank dominance criterion while 17 states ranked statistically.

In this thesis, we develop a distribution-free methodology to determine ordering or dominance relationship between generalized Lorenz curves. This non-parametric procedure follows from Aly [2]. One may also develop a non-parametric procedure for 'intersecting' generalized Lorenz curves. The latter is not shown in this thesis but can be derived following the methods in Chapter 2.

3.2 One Sample Case

Let F_o be a completely specified distribution function. Let $\hat{F}(\cdot)$, $\bar{x}$ and $x_{1:n}, x_{2:n}, \ldots, x_{n:n}$ be respectively, the empirical distribution function, the mean and the order statistics of a random sample from $F(\cdot)$. It is assumed that incomes are measured in the same monetary units. Furthermore, $\mu_F > \mu_{F_o}$.

$$H_o : F \stackrel{GL}{=} F_o \tag{3.1}$$

against the alternative

$$H_1 : F \leq_{GL} F_o \text{ and } F \stackrel{GL}{\neq} F_o \tag{3.2}$$

where $\leq_{GL}$ is defined in (1.8).

Let $L(p)$ be the Lorenz curve of income distribution F and let $GL(p)$ be the Generalized Lorenz curve of F.

$$\begin{aligned} \int_0^1 GL(p) &= \mu \int_0^1 L(p)dp = \int_0^1 \int_0^p F^{-1}(s)dsdp \\ &= \int_0^1 F^{-1}(s) \int_s^1 dpds \\ &= \int_0^1 (1-s)F^{-1}(s)ds \end{aligned} \tag{3.3}$$

Then by integration by parts,

$$\int_0^1 GL(p) = \frac{1}{2}F^{-1}(0)I(F^{-1}(0) \neq 0) + \frac{1}{2}\int_0^1 (1-s)^2 dF^{-1}(s) \tag{3.4}$$

Let F_1 and F_2 be two income distribution functions.

Define

$$\begin{aligned}
\triangle(F_1,F_2) &= 2\int_0^1 \{GL_1(p)-GL_2(p)\}\,dp \\
&= 2\int_0^1 \{\mu_1 L_1(p)-\mu_2 L_2(p)\}dp \\
&= \left\{F_1^{-1}(0)I(F_1^{-1}(0)\neq 0)+\int_0^1 (1-s)^2 dF_1^{-1}(s)\right\} - \\
&\quad \left\{F_2^{-1}(0)I(F_2^{-1}(0)\neq 0)+\int_0^1 (1-s)^2 dF_2^{-1}(s)\right\} \\
&= \delta_{GL}(F_1)-\delta_{GL}(F_2), \qquad (3.5)
\end{aligned}$$

where

$$\delta_{GL}(F) = F^{-1}(0)I(F^{-1}(0)\neq 0)+\int_0^1 (1-s)^2 dF^{-1}(s). \qquad (3.6)$$

and I(A) is the indicator function of the event A. We note that

$$\begin{aligned}
\triangle(F_1,F_2) &= 0 \text{ if } F_1 \overset{GL}{=} F_2 \text{ and} \\
\triangle(F_1,F_2) &> 0 \text{ if } F_1 \overset{GL}{\neq} F_2 \qquad (3.7)
\end{aligned}$$

For this reason, we propose to use $\triangle(F_1,F_2)$ as a measure of deviation from H_o n favor of H_1.

To conduct the one-sample test, we propose the test statistic

$$t_n = \triangle(\hat{F},F_o) = \delta_{GL}(\hat{F})-\delta_{GL}(F_o) \qquad (3.8)$$

where $\delta_{GL}(\cdot)$ is as defined in (3.6).

$$\begin{aligned}
\delta_{GL}(\hat{F}) &= x_{1:n}I(F^{-1}(0)\neq 0)+\sum_{i=1}^{n-1}\left(1-\frac{i}{n}\right)^2(x_{i+1:n}-x_{i:n}) \\
&= x_{1:n}I(F^{-1}(0)\neq 0)+\frac{1}{n^2}\sum_{i=1}^{n-1}(n-i)^2(x_{i+1:n}-x_{i:n}), \qquad (3.9)
\end{aligned}$$

The basis of the one-sample test is the following theorem:

Theorem 3.1 *Assume that $F(\cdot)$ has finite variance, and that $0 < \sigma^2(F) < \infty$, where*

$$\sigma^2(F) = 8\int_0^1 (1-x)^2 \int_0^x y(1-y)\, dF^{-1}(y)\, dF^{-1}(x) \quad (3.10)$$

Then as $n \to \infty$,

$$n^{\frac{1}{2}}\left\{\delta_{GL}(\hat{F}) - \delta_{GL}(F)\right\} \xrightarrow{D} N(0, \sigma^2(F)) \quad (3.11)$$

Proof:

Let

$$n^{\frac{1}{2}}\left\{\delta_{GL}(\hat{F}) - \delta_{GL}(F)\right\} = n^{\frac{1}{2}}\left\{\hat{I} - I\right\}$$

where

$$\begin{aligned} \hat{I} &= \hat{F}^{-1}(0)I(\hat{F}^{-1}(0) \neq 0) + \int_0^1 (1-s)^2 d\hat{F}^{-1}(s) \\ I &= F^{-1}(0)I(F^{-1}(0) \neq 0) + \int_0^1 (1-s)^2 dF^{-1}(s) \end{aligned}$$

$$\begin{aligned} n^{\frac{1}{2}}\left\{\hat{I} - I\right\} &= n^{\frac{1}{2}}\left\{\hat{F}^{-1}(0) - F^{-1}(0)\right\} I(F^{-1}(0) \neq 0) && (3.12) \\ &\quad + n^{\frac{1}{2}}\left\{g(\hat{F}; p) - g(F; p)\right\}, && (3.13) \end{aligned}$$

where $g(F;p) = \int_0^1 (1-s)^2 dF^{-1}(s)$.

We consider first the limiting distribution of the second term in (3.13).

$$n^{\frac{1}{2}}\left\{g(\hat{F}; p) - g(F; p)\right\} = n^{\frac{1}{2}}\left\{\int_0^1 (1-s)^2 d\hat{F}^{-1}(s) - \int_0^1 (1-s)^2 dF^{-1}(s)\right\},$$

By Theorem 5.1 of Aly (1990),

$$\begin{aligned} n^{\frac{1}{2}}\left\{g(\hat{F};p)-g(F;p)\right\} &\xrightarrow{D} 2\int_0^1(1-s)(-1)B(s)dF^{-1}(s) \\ &= -2\int_0^1(1-s)B(s)dF^{-1}(s) = Z, \end{aligned} \tag{3.14}$$

Next, we consider the limiting distribution of the first term in (3.13). We only need to consider the case $F^{-1}(0) \neq 0$ in which we assume that $fF^{-1}(0) \neq 0$. Write

$$\begin{aligned} n^{\frac{1}{2}}\left\{\hat{F}^{-1}(0)-F^{-1}(0)\right\} &= \frac{n^{\frac{1}{2}}f(F^{-1}(0))(\hat{F}^{-1}(0)-F^{-1}(0))}{f(F^{-1}(0))} \\ &\xrightarrow{D} \frac{B(0)}{f(F^{-1}(0))} = 0. \end{aligned} \tag{3.15}$$

Hence, as $n \to \infty$,

$$n^{\frac{1}{2}}\left\{\hat{I}-I\right\} \xrightarrow{D} \text{random variable of the R.H.S. of (3.14)}$$

It therefore follows that

$$n^{\frac{1}{2}}\left\{\delta_{GL}(\hat{F};p)-\delta_{GL}(F;p)\right\} = n^{\frac{1}{2}}\left\{g(\hat{F};p)-g(F;p)\right\} \tag{3.16}$$

Hence by (3.14) and (3.16), we have as $n \to \infty$,

$$n^{\frac{1}{2}}\left\{\delta_{GL}(\hat{F};p)-\delta_{GL}(F;p)\right\} \xrightarrow{D} -2\int_0^1(1-s)B(s)dF^{-1}(s) = Z$$

where $B(\cdot)$ is a Brownian bridge and $Z \sim N(0,\sigma^2(F))$. This completes the proof of Theorem 3.1.

Theorem 3.2 *Assume that* $EX^2 < \infty$, *and* $0 < \sigma^2(F) < \infty$. *Then*

$$n^{\frac{1}{2}}\frac{(t_n - \Delta(\hat{F}, F_o)}{\sigma(F_o)} \xrightarrow{D} N(0,1),$$
$$T_n = n^{\frac{1}{2}}\frac{t_n}{\sigma(F_o)} \xrightarrow{D} Z, \text{ under } H_o \tag{3.17}$$

where Z is a $N(0,1)$ r.v. and $\sigma(F_o)$ is as in (3.10)

The proof of Theorem 3.2 is along the lines of proof in Theorem 2.1 in Aly(1991).

As a consequence of (3.17), we propose to reject H_o of (3.1), in favor of H_1 of (3.2) if $T_n > z_{1-\alpha}$, where z_p is the p^{th} quantile of Z. The consistency of the proposed testing procedure follows from the result that if $\Delta(F, F_o) > 0$ then

$$T_n = T_n^* + R_n,$$

where

$$T_n^* \stackrel{P}{=} O(1)$$

and

$$R_n = n^{\frac{1}{2}}\frac{\{\Delta(F, F_o)\}}{\sigma(F_o)} = O(n^{\frac{1}{2}})$$

We used a small scale Monte Carlo experiment to check the asymptotic normality of T_n of (3.18). We computed T_{50} and T_{100} in the case $F(x) = 1 - \exp(-x)$, $x \geq 0$ and used 1000 replications.

3.3 Two-Sample Case

In the present section, we are interested in testing the null hypothesis

Table 3.1: Critical Values for T_n

Sample Size	0.10	0.05	0.01
$n_1 = 50$	1.1432	1.4692	2.2425
$n_1 = 10($	1.1815	1.6414	2.4468

$$H_o : F_1 \overset{GL}{=} F_2 \tag{3.18}$$

against the alternative

$$H_1 : F_1 \leq_{GL} F_2 \text{ and } F_1 \overset{GL}{\neq} F_2 \tag{3.19}$$

By the argument following (3.19), we propose the use of $\Delta(F_1, F_2)$ of (3.5) as a measure of deviation from H_o of (3.18) in favor of H_1 of (3.19).

To conduct the above-mentioned two-sample test, we propose the test statistic

$$t_{2;n_1,n_2} = \Delta(\hat{F}_1, \hat{F}_2) = \delta_{GL}(\hat{F}_1) - \delta_{GL}(\hat{F}_2), \tag{3.20}$$

where $\delta_{GL}(\cdot)$ is as defined in (3.9).

The asymptotic distribution of $t_{2;n_1,n_2}$ and the basis of the two-sample test is given in the following theorem the proof of which is shown in the appendix of Aly (1991).

Theorem 3.3 *Assume that $F_1(\cdot)$ and $F_2(\cdot)$ have finite variances, $0 < \sigma^2(F_1) < \infty$ and $0 < \sigma^2(F_2) < \infty$. Then*

$$\left(\frac{n_1 n_2}{n_1 + n_2}\right)^{\frac{1}{2}} \frac{\{t_{2;n_1,n_2} - \Delta(F_1, F_2)\}}{\sigma_{n_1,n_2}} \xrightarrow{D} N(0,1)$$

as $min(n_1, n_2) \to \infty$ and $(n_1, n_2) \in D_\lambda := \{(n_1, n_2) : \lambda \leq \frac{n_1}{n_1+n_2} \leq 1 - \lambda\}$ for some $\lambda \sim (0, 1/2]$, where $\triangle(F_1, F_2)$ is as (3.5) and

$$\sigma^2_{n_1,n_2} = \frac{\{n_1\sigma^2(F_1) + n_2\sigma^2(F_2)\}}{(n_1 + n_2)}$$

Next we define,

$$\hat{\sigma}^2_{n_1,n_2} = \sigma^2_{n_1,n_2}(\hat{F}_1, \hat{F}_2) = \{n_1\sigma^2(\hat{F}_1) + n_2\sigma^2(\hat{F}_2)\}/(n_1 + n_2), \tag{3.21}$$

where

$$\sigma^2(\hat{F}_1) = 8\int_0^1 (1-x)^2 \int_0^x y(1-y)d\hat{F}_1^{-1}(y)d\hat{F}_1^{-1}(x) \tag{3.22}$$

$$= 8\sum_{j=2}^{n_1-1}\left(1 - \frac{j}{n_1}\right)^2 \sum_{i=1}^{j-1} \frac{i}{n_1}(1 - i/n_1)(x_{i+1:n_1} - x_{i:n})(x_{j+1:n_1} - x_{j:n}) \tag{3.23}$$

and $\sigma^2(\hat{F}_2)$ is defined similarly.

From the definition of $\sigma^2(\hat{F}_1)$ and $\sigma^2(\hat{F}_2)$ and by the SLLN and Theorem 11 of Sendler (1979), it follows that they are consistent estimators of $\sigma^2(F_1)$ and $\sigma^2(F_2)$ respectively This result, and the fact that under H_o, $\triangle(F_1, F_2) = 0$, imply the following corollary.

Corollary 3.1 *Assume H_o holds true. Then under the conditions of Theorem 3.3, we have*

$$T_{2;n_1,n_2} = \left(\frac{n_1 n_2}{n_1 + n_2}\right)^{\frac{1}{2}} \frac{t_{n_1,n_2}}{\hat{\sigma}_{n_1,n_2}} \xrightarrow{D} Z \tag{3.24}$$

where $Z \sim N(0,1)$ random variable.

Above corollary suggest that an asymptotic distribution free test for testing H_o of (3.18) against H_1 of (3.19) is to reject H_o if $T_{2;n_1,n_2} > z_{1-\alpha}$. The test based on $T_{2;n_1,n_2}$ is consistent.

Table 3.2: Critical Values for $T_{2;n_1,n_2}$

Sample Sizes	0.10	0.05	0.01
$n_1, n_2 = 50$	1.28649	1.54799	2.23698
$n_1, n_2 = 100$	1.33959	1.67885	2.34783

We performed a Monte Carlo simulation study to check the asymptotic normality of $T_{2;n_1,n_2}$ of (3.24) In this paper, we computed $T_{2;50,50}$ and $T_{2;100,100}$ in the case $F_1(x) = F_2(x) = 1 - \exp(-x)$, $x \geq 0$. The number of replications is 1000. The Monte Carlo critical values are shown in Table 3.2.

Chapter 4

An Alternative Index of Income Inequality

In this chapter, we consider the problem of obtaining an asymptotically distribution-free estimation procedure for the point inequality measure K and its corresponding global measure ξ^K. These measures are derived from the ratio between population and income quantiles of *Zenga* [54]. The asymptotic theory for the K empirical concentration process is also developed and consequently, a variance estimator is defined. Simulation studies are conducted under two leading income distribution models, namely the Lognormal and Pareto Distributions to check, for finite sample sizes, the asymptotic results and properties of these estimators.

4.1 Introduction

In recent years, economic and statistical research on income distribution has led to the development of several measures (curves) of income inequality or point concentration measures (PCM). Nygard and Sandstrom [41] give an excellent accounting of these concentration curves. Frosini [23] makes a thorough examination of inequality measures based on various types of comparisons between the distribution function F and the first moment distribution function Q or as a comparison between their inverses. The first incomplete moment distribution function $Q(x)$ is defined as $Q(x) = \frac{1}{\mu}\int_a^x sf(s)ds$.

As inequality increases the two curves tend to move apart although not always (Frosini [23]); hence, some *distance* measures between the two curves can be used, or at least be candidates for inequality measures. The first of such measures is the Holme's measure given by $Q^{-1}(0.5) - F^{-1}(0.5)$. Modifications to the Holme's measure were done by dividing by μ or by $Q^{-1}(0.5)$. Others have proposed similar measures of specific fractiles. Frosini [23] gives a systematic analysis of deriving point inequality measures. If the distance is computed as a function of a particular abscissa x, or of a particular ordinate p, such distances are called point concentration measures. If the distance is a function of two or more (typically all) points of the curve F and Q, this is said to be a global measure or global index. This is computed as the average of point measures and is the area under the point concentration measure.

If we let $QF^{-1}(p) = L(p)$, examples of measures derived from differences or ratios between ordinates are $A(p) = p - L(p)$, $B(p) = FQ^{-1}(p) - p$, etc. The measures $C(x) = x - F^{-1}[Q(x)]$, $a \leq x \leq b$ and $D(p) = Q^{-1}(p) - F^{-1}(p)$ are based on comparisons between the abscissas. Point measures allow comparisons between distributions, not within the same distributions. The following interpretations of the ordinates and abscissas are useful in determining the real significance that the above inequality measures possess.

1. p refers to the percentage of poorest individuals;
2. $F^{-1}(p)$ is the highest income of the share p of (poorer) individuals;
3. $Q^{-1}(p)$ is the highest income of the poorer individuals that possess together the share p of the income;
4. $L(p)$ is the percentage of total income possessed by the p^{th} % poorest of individuals;
5. $F(x)$ is the percentage of individuals with income less than or equal to x;
6. $Q(x)$ refers to the percentage of total income possessed by the same individuals;
7. $FQ^{-1}(p)$ is the percentage of (poorer) individuals that possess the fixed share of income;

8. $Q^{-1}F(x)$ refers to the highest income among the share of (poorer) individuals who possess the income proportion F(x).

In order for a PCM to be in the roster of valid inequality measures, it must satisfy the following characteristics: (i) The PCM must take values in the interval $[0, 1]$; (ii) It must not have a forced behavior; (iii) As the income distribution X tends to the null concentration (i.e., when all individuals receive the same income or $L(p) = p$), the PCM must tend to zero and as X tends to the maximum concentration (i.e., when one person receives all the income and others receive none or $L(p) = 0$ for $0 \leq p < 1$ and $L(p) = 1$ for $p = 1$), the PCM must tend to one; and (iv) The PCM allows partial ordering. While the ordinate of the Lorenz Curve, $L(\cdot)$, has been the most widely used PCM, by construction, it has a forced or predetermined behavior in the sense that $L(0) = 0$ and $L(1) = 1$ and is always an increasing convex function.

Zenga [54], introduced the Z-curve which is derived based on the analysis outlined above but normalized in the sense that the difference $D(p)$ is divided by $Q^{-1}(p)$. It is a comparison of the *difference* of the highest income of the share p of (poorer) individual and the highest income among the poorer individuals that possess together the share p of the total income, *relative* to the highest income of the poorer individuals that possess the fixed share of total income. It satisfies the properties of a good PCM particularly that of a free behavior. Not having a forced behavior property means that when p varies from 0 to 1 or when x varies from a to b, the PCM does not obey a pre-assigned pattern (say, a curve being convex). Its global index ξ^Z, which is the average of point measures $Z(p)$, can be visualized as the area under the Z-curve. Moreover, when x tends to the null concentration, Z(p) tends to one and when x tends to the maximum concentration, $Z(p)$ tends to one. This implies that the $Z(p) = 0$ is the complete-equality line and the deviation from the Z-curve to this line is the point concentration. The larger the deviation, the higher the point concentration. Dancelli [20], Latorre [41], Salvaterra [47], Pollastri [43], and Dagum [17] have critically surveyed the advantages and disadvantages of the Z measure.

These theoretical and empirical studies have shown the superiority of the Z-curve relative to other measures. However, since this is a relatively new area of research and mainly due to the complexity of the expression for Z, little has been done to develop the corresponding tools for formal statistical inference. In particular, only the *parametric* approach of estimation has been extensively studied (see Latorre [38]).

In this paper, our focus is on developing nonparametric asymptotically distribution-free methods for the K-PCM. This inequality measure can be derived from the comparison of the abscissa (particularly $C(x)$) or from the Z-measure. The (1-$K(p)$) curve is interpreted as the measure of highest income among the share p of (poorer) individuals who possess a percentage of total income em relative to the average income of the same individuals. The global index of the K-curve is also the average of the point measures and is the area under the K-curve.

In *Section 2*, expressions for the K-PCM and the corresponding global index ξ are shown. A comparison is performed against the *Lorenz* and *Z-PCM* in terms of their properties. Applications of these measures to classical distribution models and to empirical data of USA and ITALY are explored as well. The development of the statistical theories are contained in *Section 3*, particularly the nonparametric approach for estimating ξ. Further, asymptotic theory of our proposed estimator of ξ is developed leading to inferences regarding the sampling variability of the income distribution. The variance is estimated nonparametrically. Unbiasedness, consistency and coverage probability are verified through a simulation study in *Section 4*.

4.2 The K-PCM

4.2.1 The K-PCM as derived from Z-PCM

Let the income X of a unit be a nonnegative continuous random variable with density function f(x), cumulative distribution function (cdf) $F(x)$ and finite mean $E[X] = \mu$. Let $Q(x)$ be the cdf of the first incomplete moment, that is, $Q(x) = \frac{1}{\mu}\int_a^x sf(s)ds$. Note that both $F(\cdot)$ and $Q(\cdot)$ are invertible, that is, both are stricly increasing and positive on their supports. Zenga [54] defines the Z-PCM or Z-Curve as:

$$
\begin{aligned}
Z(p) &= \frac{Q^{-1}(p) - F^{-1}(p)}{Q^{-1}(p)} \\
&= 1 - \frac{F^{-1}(p)}{Q^{-1}(p)}, \qquad 0 < p < 1, \qquad (4.1)
\end{aligned}
$$

where

$F^{-1}(p) = inf\{x : F(x) >\}$ is the population quantile,

$Q^{-1}(p) = inf\{x : Q(x) >\}$ is the income quantile.

The corresponding global measure of Z-PCM, denoted by ξ^Z, is the area under the concentration curve Z, that is,

$$
\begin{aligned}
\xi^Z &= \int_0^1 Z(p)dp \\
&= 1 - \int_0^1 \frac{F^{-1}(p)}{Q^{-1}(p)}\,dp. \qquad (4.2)
\end{aligned}
$$

The Lorenz Curve of F is defined as

$$
L(p) = \frac{1}{\mu}\int_0^p F^{-1}(t)\,dt. \qquad (4.3)
$$

In empirical work, the Lorenz curve maps the cumulative distribution of households arranged by increasing the size of their incomes (ie p=F(x)) into the cumulative distribution of their corresponding aggregate income.

We note that the Lorenz Curve $L(\cdot)$ can be written as $L(y) = Q \circ F^{-1}(y)$ and hence $L^{-1}(p) = FQ^{-1}(p)$ and $Q^{-1}(p) = F^{-1}L^{-1}(p)$. This leads to an alternative expression for ξ^Z as explained below.

By (4.2), we have

$$\begin{aligned} \xi^Z &= \int_0^1 Z(p)dp \\ &= 1 - \int_0^1 \frac{F^{-1}(p)}{Q^{-1}(p)} dp \\ &= 1 - \int_0^1 \frac{F^{-1}(p)}{F^{-1}L^{-1}(p)} dp \end{aligned} \tag{4.4}$$

Let $L^{-1}(p) = y$ and note that $dp = L'(y)dy = \frac{1}{\mu}F^{-1}(y)dy$. Substituting into (4.4), we get

$$\xi^Z = 1 - \int_0^1 \frac{F^{-1}L(y)\frac{1}{\mu}F^{-1}(y)}{F^{-1}(y)} dy = 1 - \frac{1}{\mu}\int_0^1 F^{-1}L(y)dy, \tag{4.5}$$

Note that by (4.5), $\xi^Z = \int_0^1 K(p)\,dp$, where

$$K(p) = 1 - \frac{1}{\mu}F^{-1}L(p), \qquad 0 < p < 1 \tag{4.6}$$

This suggests that the new measure[1](curve) $K(\cdot)$ defined by (4.6) and denoted by K-PCM is a point concentration measure. Note that ξ^Z is the area under the K-Curve (and also under the Z-Curve). We will use the notation $\xi(= \xi^Z)$ when we use the representation in terms of the K-Curve.

[1]Frossini has examined several point measures based on the comparison between the distribution function F and the 1st moment incomplete distribution function Q. The K PCM can also be derived using such method of comparison.

Remarks:

From (4.2), we can express ξ as

$$\begin{aligned}\xi &= 1 - \frac{1}{\mu}\int_0^1 F^{-1}L(p)\,dp \\ &= \frac{1}{\mu}\left\{\mu - \int_0^1 F^{-1}L(p)\,dp\right\} \\ &= \frac{1}{\mu}\left\{\int_0^1 (F^{-1}(y) - F^{-1}L(y))\,dy\right\} \\ &= \int_0^1 K^{*}(\cdot)\,dy,\end{aligned}$$

where

$$K^{*}(p) = \frac{1}{\mu}\left\{F^{-1}(p) - F^{-1}L(p)\right\}, \qquad \forall p \in (0,1)$$

which suggests the additional measure $K^{*}(\cdot)$ as a point concentration measure.

4.3 Comparison between the K and Z Curves

In this section, we establish the properties of the K-Curve and compare it with Zenga's Z-Curve.

Properties of the K-PCM

1. $K(p)$ takes values in

$$\left[1 - \frac{t_F}{\mu}, 1 - \frac{T_F}{\mu}\right],$$

   where $t_F = F^{-1}(0)$ and $T_F = F^{-1}(1-)$.

2. $K(p)$ has a pre-established behavior since the curve is a decreasing function of p($\frac{dK(p)}{dp} = \frac{-F^{-1}(p)}{\mu}\frac{1}{\mu f(F^{-1}L(p))} \leq 0$). The curve can be either convex or concave.

3. The null and maximum concentration results apply for $p \in (0,1)$.

Functional Forms to Fit Income Distributions

Behavior of L and Z under the Uniform, Lognormal and Pareto Models have been analyzed in Zenga [54]. Here we presen' 'he graphs of these curves together with those from the K-PCM.

Behavior for Uniform Distribution

$X \rightsquigarrow Uniform(\mu(1-\alpha), \mu(1+\alpha))$; $E(X) = \mu$;

$$f(x) = \begin{cases} \frac{1}{2\mu\alpha} & \mu(1-\alpha) \leq x \leq \mu(1+\alpha), \mu > 0; 0 < \alpha \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < \mu(1-\alpha) \\ \frac{x-\mu(1-\alpha)}{2\mu\alpha} & \mu(1-\alpha) \leq x \leq \mu(1+\alpha) \\ 1 & x > \mu(1+\alpha) \end{cases}$$

$$\begin{aligned} F^{-1}(p) &= \mu\{2\alpha p + (1-\alpha)\} \\ \frac{dF^{-1}(p)}{dp} &= 2\mu\alpha \\ Q^{-1}(p) &= \sqrt{4\alpha p + (1-\alpha)^2} \end{aligned}$$

1. $L(p) = \frac{1}{\mu}\int_0^p F^{-1}(t)dt = \alpha p^2 + (1-\alpha)p$
2. $Z(p) = 1 - \frac{F^{-1}(p)}{Q^{-1}(p)} = 1 - \frac{2\alpha p+(1-\alpha)}{\sqrt{4\alpha p+(1-\alpha)^2}}$
3. $K(p) = 1 - \frac{F^{-1}L(p)}{\mu} = 1 - \{2\alpha^2p^2 + 2\alpha(1-\alpha)p + (1-\alpha)\}$
4. $K^*(p) = \frac{1}{\mu}\{F^{-1}(p) - F^{-1}L(p)\} = \{2\alpha p + (1-\alpha)\} - \{2\alpha^2p^2 + 2\alpha(1-\alpha)p + (1-\alpha)\}$

Figure 4.1 (a),(b),(c) and (d) are the $L, Z, K,$ *and* K^* curves of the uniform distribution with $\alpha = 0.5$ and $\alpha = 2.5$.

Behavior for Lognormal

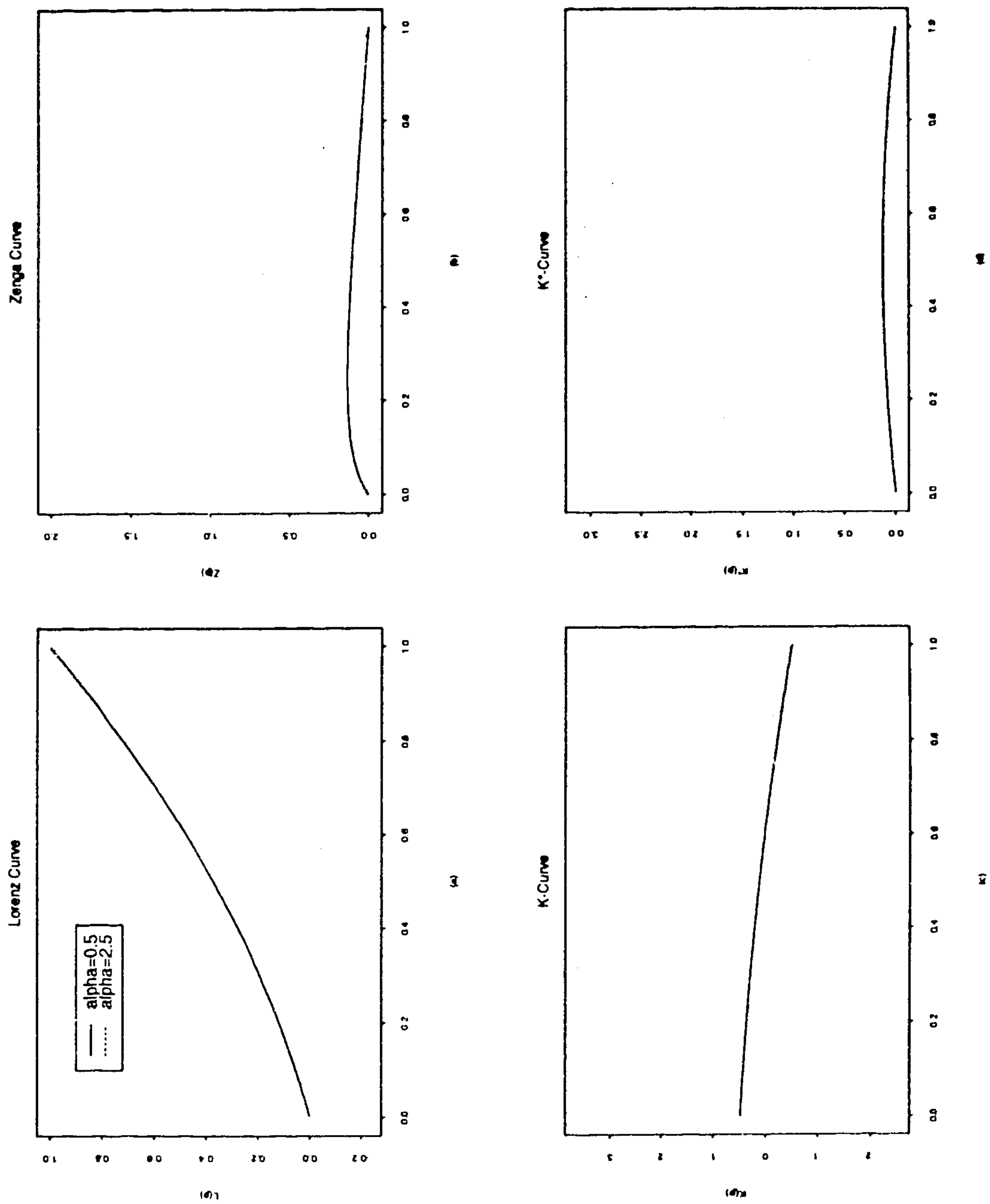

Figure 4.1: Comparison of Curves under the Uniform

$X \rightsquigarrow$ Lognormal with $E(X) = \exp(\gamma + \frac{1}{2}\delta^2)$,

$Var(X) = \exp(2\gamma + 2\delta^2) - \exp(2\gamma + \delta^2),\ \delta > 0.$

$$
\begin{aligned}
f(x) &= \begin{cases} \frac{1}{\delta\sqrt{2\pi}}\frac{1}{2}\exp\left(-\frac{1}{x}\left[\frac{\log x-\gamma}{\delta}\right]^2\right) & x > 0; |\gamma| < +\infty, 0 < \delta < \infty \\ 0 & otherwise \end{cases} \\
F(x) &= \Phi\left(\frac{\log x - \gamma}{\delta}\right) \\
Q(x) &= \Phi\left(\frac{logx - \gamma - \delta^2}{\delta}\right) \\
F^{-1}(p) &= \exp\left(\gamma + \delta\,\Phi^{-1}(p)\right) \\
\frac{dF^{-1}(p)}{dp} &= \delta\exp\left(\gamma + \delta\,\Phi^{-1}(p)\right) d\Phi^{-1}(p) \\
Q^{-1}(p) &= \exp\left(\gamma + \delta^2 + \delta\,\Phi^{-1}(p)\right)
\end{aligned}
$$

where $\Phi^{-1}(p)$ is the inverse of the standard normal distribution function.

1. $L(p) = \Phi\left[\Phi^{-1}(p) - \delta\right]$
2. $Z(p) = 1 - \exp(-\delta^2)$
3. $K(p) = 1 - \exp\left(\delta\,\Phi^{-1}(p) - \frac{\delta^2}{2}\right)$
4. $K^*(p) = \frac{\exp\left(\delta\,\Phi^{-1}(p)\right) - \exp\left(\delta\,\Phi^{-1}(p) - \delta^2\right)}{\exp\left(\frac{1}{2}\delta^2\right)}$

Note: Figures 4.2 (a), (b), (c) and (d) are the L, Z, K and K* - Curves of the Lognormal distribution with $\gamma = 2.8$ and $\delta = 0.35, 1.5$.

Behavior for Pareto Type I

$X \rightsquigarrow$ Pareto with $E(X) = \frac{x_o\theta}{\theta-1}$, $\theta > 1$ and $E(X^2) = \frac{x_o^2\theta}{\theta-2}$, $\theta > 2$

$$
\begin{aligned}
f(x) &= \theta\, x_o^\theta\, x^{-\theta-1},\ x \geq x_o > 0, \\
F(x) &= 1 - x_o^\theta x^{-\theta} = 1 - \left(\frac{x_o}{x}\right)^\theta
\end{aligned}
$$

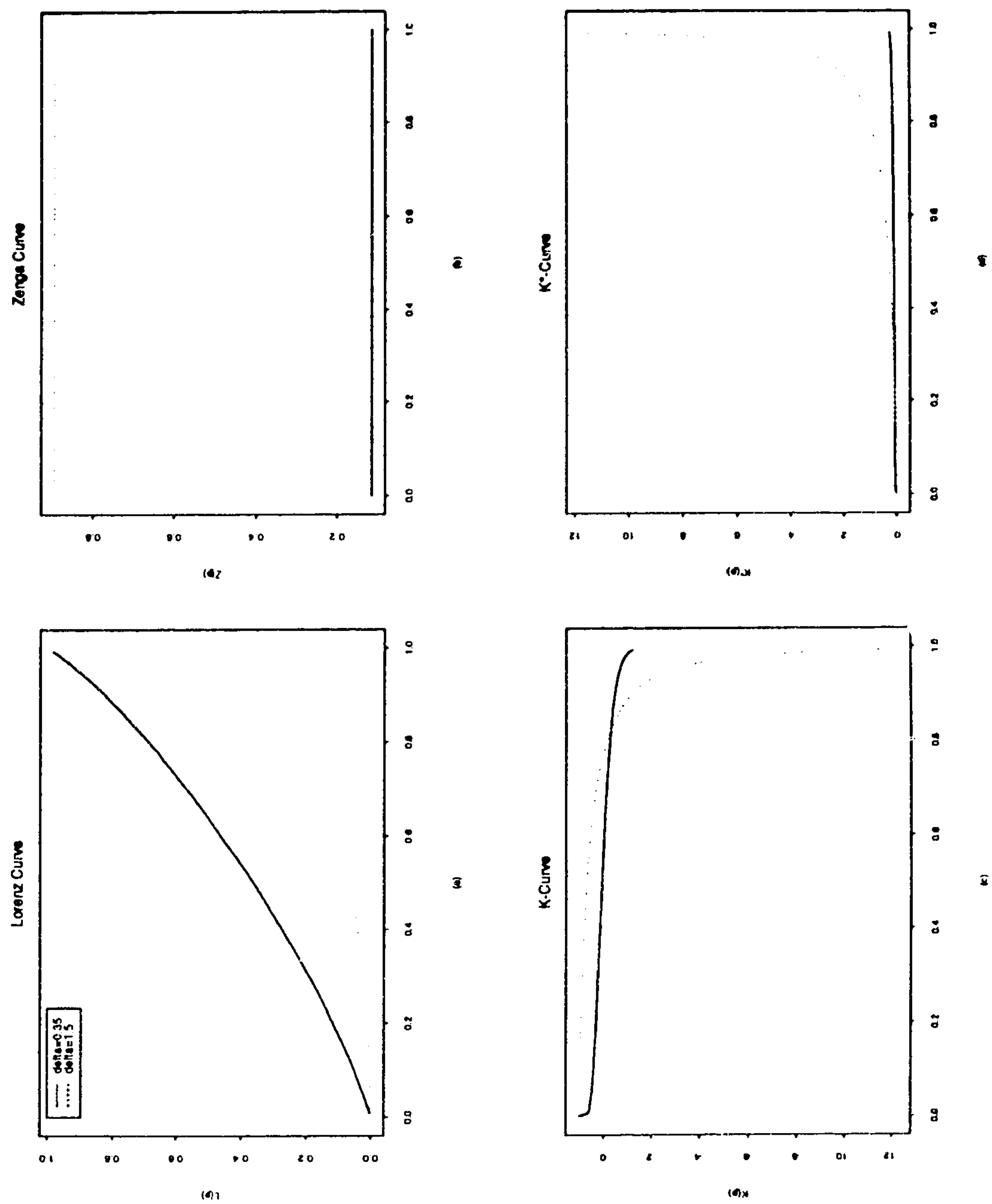

Figure 4.2: Comparison of Curves under the Lognormal

$$
\begin{aligned}
Q(x) &= 1-\left(\frac{x}{x_o}\right)^{1-\theta} \\
F^{-1}(p) &= x_o(1-p)^{-\frac{1}{\theta}} \\
\frac{dF^{-1}(p)}{dp} &= \frac{x_o}{\theta}(1-p)^{-(1/\theta+1)} \\
Q^{-1}(p) &= x_o(1-p)^{-\frac{1}{(\theta-1)}}
\end{aligned}
$$

1. $L(p) = 1-(1-p)^{\frac{(\theta-1)}{\theta}}$

2. $Z(p) = 1-(1-p)^{\frac{1}{\theta(\theta-1)}}$

3. $K(p) = 1-\frac{\left[(1-p)^{((1-\theta)/\theta^2)}\right]}{\theta(\theta-1)^{-1}}$

4. $K^*(p) = \frac{\theta-1}{x_o\theta}\left[(1-p)^{-\frac{1}{\theta}}-(1-p)^{\frac{1-\theta}{\theta^2}}\right]$

Note: Figures 4.3 (a), (b), (c), and (d) are the L, Z, K and K* - Curves of the Pareto distribution with $\theta = 1.2$ and 2.9, 4.5.

Application to Empirical Income Distribution

Salvaterra [47] has studied the behavior of the L and Z concentration curves in grouped data taken from the 1986 Survey of the Bank of Italy on income distribution of Italian families. Pollastri [43] has likewise determined the behavior for both L and Z for the data taken from the 1935-36 and 1979-80 Survey on the U.S.A. Personal Income Distribution. In Figures 4.4 (a), (b) and (c) and Figure 4.5 (a)-(f), we reproduce their results together with the behavior for the K Curve.

Inspection of the curves indicates that the curve $Z(p)$ seems more flexible in describing inequality and is more sensitive to variations in the empirical distributions. In particular, when the Z-curves in Figures 4.3.5 (c) and (d) are superimposed, we notice that in the USA during the two periods studied, the point concentration has increased for the richest and has decreased for the poorest. We are unable to arrive at a similar conclusion when the Lorenz curves are being compared.

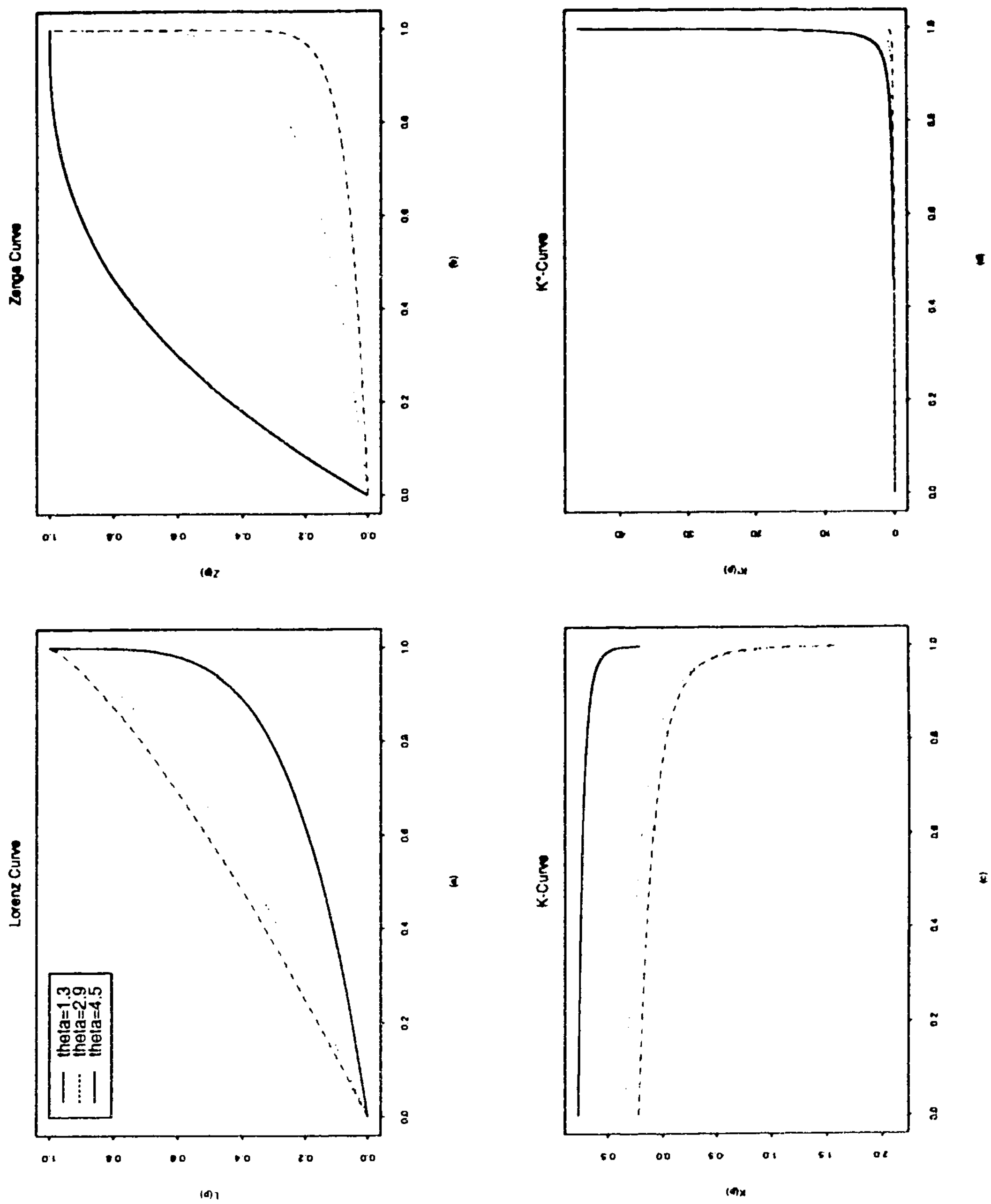

Figure 4.3: Comparison of Curves under the Pareto

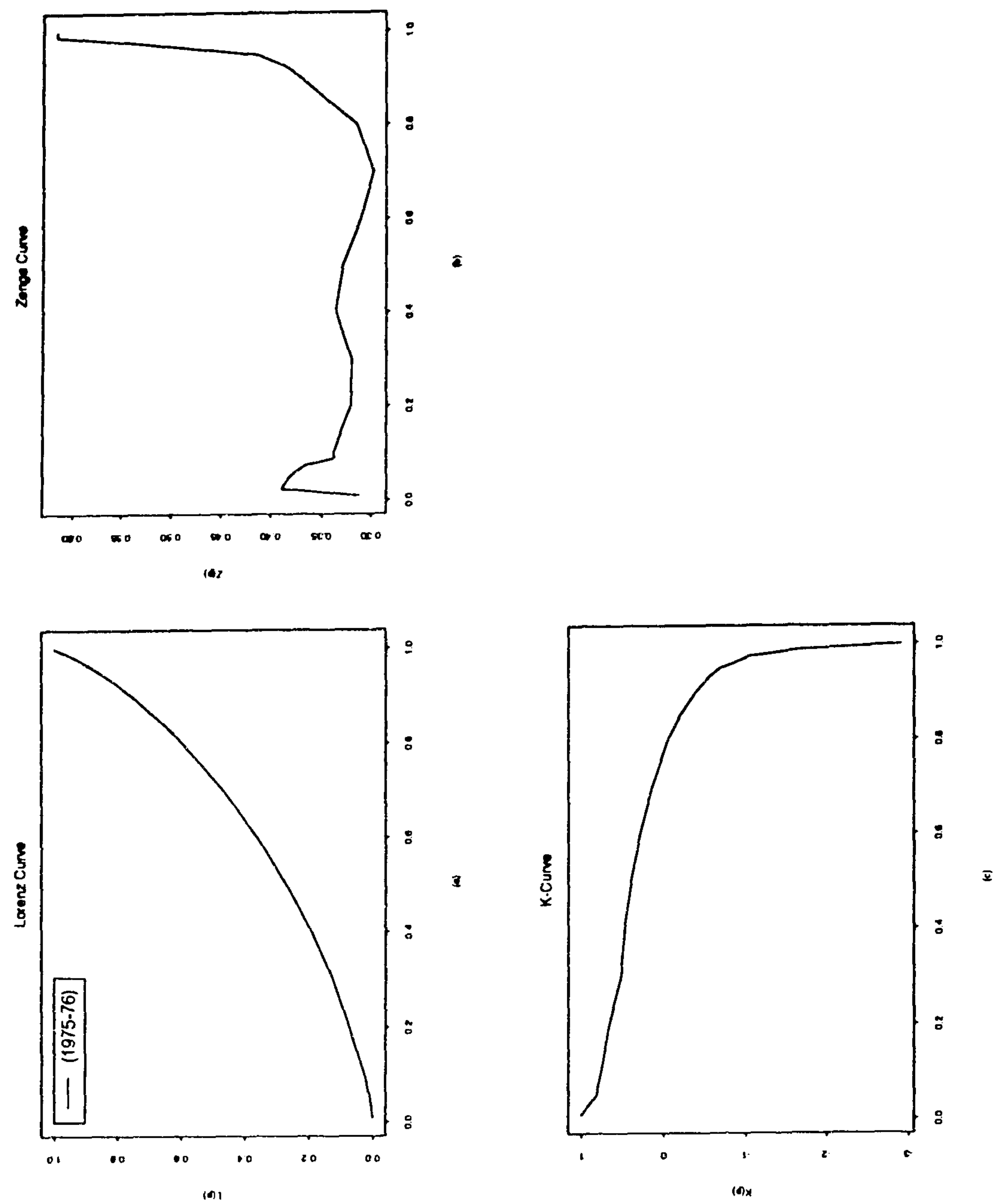

Figure 4.4: Comparison of Curves in Grouped Data (Italy)

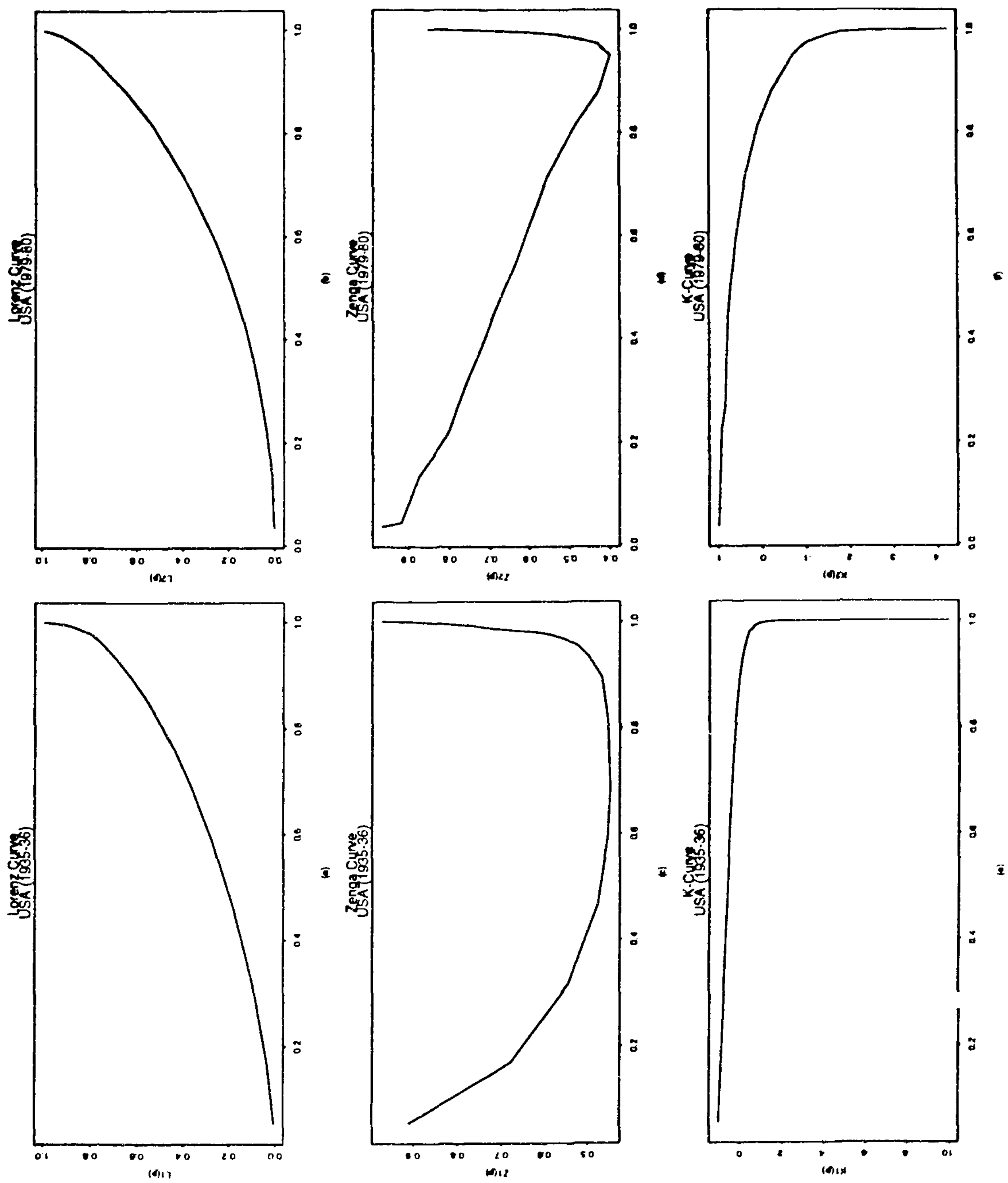

Figure 4.5: Comparison of Curves in Grouped Data (USA)

4.4 On Theoretical Results

In a descriptive sense, the $Z(p)$ Curve is shown to be more superior than the $L(p)$, $K(p)$ or $K^{*}(p)$ Curves. This is evident in its properties, through its behavior for various classical income distribution models and because of its flexibility and sensitivity to variations of income inequality in the empirical distributions. Statistical inference based on the Z Curve is hindered by its expression as a ratio. However, for the global index ξ, we are able to obtain a nonparametric estimate of ξ and develop its properties.

The first subsection gives a derivation of the nonparametric estimator $\hat{\xi}_n$ for the global index ξ. Next in section 4.4.2, a Central Limit Theorem for $\hat{\xi}_n$ is established. A nonparametric estimate of the variance of $\hat{\xi}_n$ is given in section 4.4.3 Furthermore, we present another method of estimating $\sigma^2(\hat{\xi}_n)$ through the quantile density estimation approach. The form $q(\cdot) = \frac{1}{fF^{-1}(\cdot)}$ appears in the expression of $\sigma^2(\hat{\xi}_n)$, hence the problem of estimating q in terms of either a histogram-type or kernel-type estimator is investigated.

The last subsection of this section provides expressions for the exact asymptotic variance under the Pareto and Lognormal distributions.

4.4.1 A Nonparametric Estimator of ξ

Let F_n be the empirical distribution function (EDF) based on a random sample $x_1, x_2, \ldots, x_n$ from F, that is,

$$F_n(x) = \begin{cases} 0 & x_{1:n} > x \\ \frac{k}{n} & x_{k:n} \leq x < x_{k+1:n},\ k = 1 \ldots n \\ 1 & x_{n:n} \leq x, \end{cases} \tag{4.7}$$

where $x_{1:n}, x_{2:n}, \ldots, x_{n:n}$ are the order statistics of the X sample.
Let $F_n^{-1}(y) = inf\{x : F_n(x) > y\}$ be the corresponding empirical quantile function. Note that we will use the following (asymptotically equivalent) definition of $F_n^{-1}(\cdot)$,

$$F_n^{-1}(y) = \begin{cases} x_{1:n} & y = 0 \\ x_{<ny>:n} & 0 < y \leq 1, \end{cases} \quad (4.8)$$

where $< x >$ is the smallest integer $\geq$ x.

Let $\tilde{L}_n(\cdot)$ be the empirical counterpart of $L(\cdot)$. It can be shown that $\tilde{L}_n(0) = 0$, $\tilde{L}_n(1) = 1$ and

$$\tilde{L}_n(y) = \frac{1}{n\bar{x}} \sum_{i=1}^{[ny]} x_{i:n} + \frac{1}{n\bar{x}}(ny - [ny])\, x_{[ny]+1:n} \quad 0 < y < 1, \quad (4.9)$$

where $\sum_{i=1}^{0} = 0$ and $[t]$ is the greatest integer $\leq t$.

Note that

$$\frac{1}{n\bar{x}} \sum_{i=1}^{[ny]} x_{i:n} \leq \tilde{L}_n(y) < \frac{1}{n\bar{x}} \sum_{i=1}^{[ny]+1} x_{i:n} \quad (4.10)$$

For this reason, the empirical Lorenz curve has been defined in the literature by any of the three (asymptotically) equivalent expressions appearing in (4.10). In this paper, we define the empirical Lorenz curve, $L_n(\cdot)$, as

$$L_n(y) = \frac{1}{n\bar{x}} \sum_{i=1}^{[ny]} x_{i:n} \qquad 0 \leq y \leq 1. \quad (4.11)$$

The left continuous inverse, $L_n^{-1}(\cdot)$, of $L_n(\cdot)$, is known in the literature as the *empirical concentration curve.* It can be shown that $L_n^{-1}(0) = 0$ and

$$\begin{aligned} L_n^{-1}(p) &= inf\{y : L_n(y) \geq p\} \\ &= \begin{cases} 0 & p = 0 \\ \frac{k}{n} & \frac{1}{n\bar{x}} \sum_{i=1}^{k-1} x_{i:n} < p \leq \frac{1}{n\bar{x}} \sum_{i=1}^{k} x_{i:n} \end{cases} \end{aligned} \quad (4.12)$$

This implies that

$$L_n^{-1}(\frac{j}{n}) = \begin{cases} 0 & j=0 \\ \frac{k}{n} & \frac{1}{\bar{x}}\sum_{i=1}^{k-1} x_{i:n} < j \le \frac{1}{\bar{x}}\sum_{i=1}^{k} x_{i:n} \end{cases} \tag{4.13}$$

Now, recall that the concentration index ξ is defined in (4.5), as

$$\xi = 1 - \frac{1}{\mu}\int_0^1 F^{-1}L(y)\,dy$$

Our proposed nonparametric estimate of ξ is given by

$$\hat{\xi}_n = 1 - \frac{1}{\bar{x}}\int_0^1 F_n^{-1}L_n(y)dy \tag{4.14}$$

By (4.8), (4.11), (4.14) we obtain

$$\begin{aligned} \hat{\xi}_n &= 1 - \frac{1}{\bar{x}}\sum_{j=0}^{n-1}\int_{[\frac{j}{n},\frac{j+1}{n})} F_n^{-1}\left(\frac{1}{n\bar{x}}\sum_{i=1}^{j} x_{i:n}\right) \\ &= 1 - \frac{1}{n\bar{x}}\sum_{j=1}^{n-1} F_n^{-1}\left(\frac{1}{n\bar{x}}\sum_{i=1}^{j} x_{i:n}\right) - \frac{1}{n\bar{x}}F_n^{-1}(0) \\ &= 1 - \frac{1}{n\bar{x}}\left\{x_{1:n} + \sum_{j=1}^{n-1} x_{\langle\sum_{i=1}^{j} x_{i:n}/\bar{x}\rangle:n}\right\}. \end{aligned} \tag{4.15}$$

Remark:

If we estimate $L_n(\cdot)$ in (4.14) by the third expression of (4.10), we arrive at

$$\hat{\xi}_n^* = 1 - \frac{1}{n\bar{x}}\left\{\sum_{j=1}^{n-1} x_{\left[\frac{\sum_{i=1}^{j} x_{i:n}}{\bar{x}}\right]+1:n} + x_{n:n}\right\} \tag{4.16}$$

as an alternative estimator of ξ. We have noticed that $\hat{\xi}_n^*$ overestimates ξ.

4.4.2 Asymptotic Distribution of $\hat{\xi}_n$

In this section we derive the asymptotic distribution of $\hat{\xi}_n$.

Theorem 4.1 *Assume that $F(\cdot)$ satisfies the Csörgő- Révész tail conditions on (t_F, T_F) (cf. Definition 2.1) and that $0 < \sigma_\xi^2(F) < \infty$, where*

$$\begin{aligned}
\sigma_\xi^2(F) &= \sigma_\xi^2(\hat{\xi}_n) = 2\int_0^1 \left\{(1-y)h(y)\int_0^y xh(x)dF^{-1}(x)\right\} dF^{-1}(y); && (4.17)\\
h(t) &= \frac{1}{F^{-1}L^{-1}(t)} + \frac{g(t)}{\mu^2} - \frac{cl}{\mu^2} - \frac{(1-\xi)}{\mu}; && (4.18)\\
g(t) &= \int_{L(y)}^1 \frac{\mu}{F^{-1}L^{-1}(t)} dF^{-1}(t) \text{ and} && (4.19)\\
cl &= \int_0^1 \frac{t\mu}{F^{-1}L^{-1}(t)} dF^{-1}(t). && (4.20)
\end{aligned}$$

Then, as $n \longrightarrow \infty$,

$$n^{\frac{1}{2}}\left\{\hat{\xi}_n - \xi\right\} \xrightarrow{D} N\left(0, \sigma_\xi^2(F)\right).$$

Proof:

Recall that

$$\hat{\xi}_n = 1 - \frac{I_n}{\bar{x}} = 1 - \frac{1}{\bar{x}}\int_0^1 F_n^{-1}L_n(y)\,dy$$

and

$$\xi = 1 - \frac{I}{\mu} = 1 - \frac{1}{\mu}\int_0^1 F^{-1}L(y)\,dy.$$

Hence,

$$\begin{aligned}
\hat{\xi}_n - \xi &= \frac{I}{\mu} - \frac{I_n}{\bar{x}}\\
&= \frac{1}{\bar{x}}(I_n - I) - \frac{I}{\mu\bar{x}}(\bar{x} - \mu)\\
&= \frac{1}{\bar{x}}(I_n - I) - \frac{(1-\xi)}{\bar{x}}(\bar{x} - \mu). && (4.21)
\end{aligned}$$

Define $U_n(\cdot) = FF_n^{-1}(\cdot)$ and $u_n(t) = n^{\frac{1}{2}}(U_n(t) - t), 0 \le t \le 1$.

Note that

$$\begin{aligned} n^{\frac{1}{2}}(I_n - I) &= n^{\frac{1}{2}} \int_0^1 \left\{F_n^{-1}L_n(y) - F^{-1}L(y)\right\} dy \\ &= n^{\frac{1}{2}} \int_0^1 \left\{[F_n^{-1}L_n(y) - F^{-1}L_n(y)] + [F^{-1}L_n(y) - F^{-1}L(y)]\right\} dy \\ &= J_1(y) + J_2(y), \end{aligned}$$

where

$$J_1(y) = n^{\frac{1}{2}} \int_0^1 \left\{F_n^{-1}L_n(y) - F^{-1}L_n(y)\right\} dy$$

and

$$J_2(y) = n^{\frac{1}{2}} \int_0^1 \left\{F^{-1}L_n(y) - F^{-1}L(y)\right\} dy$$

By a two term Taylor series expansion,

$$\begin{aligned} J_2(y) &= n^{\frac{1}{2}} \int_0^1 \{L_n(y) - L(y)\} \frac{1}{fF^{-1}L(y)} dy - \frac{1}{2}n^{\frac{1}{2}} \int_0^1 \{L_n(y) - L(y)\}^2 \frac{f'F^{-1}(\epsilon_y)}{f^2F^{-1}L(\epsilon_y)} dy \\ &= n^{\frac{1}{2}} \int_0^1 \{L_n(y) - L(y)\} \frac{1}{fF^{-1}L(y)} dy - R_2, \qquad (4.22) \end{aligned}$$

where

$$R_2 = \frac{1}{2}n^{\frac{1}{2}} \int_0^1 \{L_n(y) - L(y)\}^2 \frac{f'F^{-1}(\epsilon_y)}{f^2F^{-1}L(\epsilon_y)}$$

and

$$min(L(y), L_n(y)) \le \epsilon_y \le max(L(y), L_n(y)).$$

Also,

$$\begin{aligned} J_1(y) &= n^{\frac{1}{2}} \int_0^1 \left\{ F_n^{-1} L_n(y) - F^{-1} L_n(y) \right\} dy \\ &= n^{\frac{1}{2}} \int_0^1 \left\{ F^{-1}(F F_n^{-1} L_n(y)) - F^{-1} L_n(y) \right\} dy \end{aligned} \tag{4.23}$$

By a two term Taylor series expansion,

$$\begin{aligned} J_1(y) &= n^{\frac{1}{2}} \int_0^1 \left\{ F F_n^{-1} L_n(y) - L_n(y) \right\} \frac{1}{f F^{-1} L_n(y)} dy \\ &\quad - \frac{1}{2} n^{\frac{1}{2}} \int_0^1 \{U_n L_n(y) - L_n(y)\}^2 \frac{f' F^{-1} L_n(\delta_y)}{f^2 F^{-1} L_n(\delta_y)} dy \\ &= \int_0^1 \frac{u_n(L_n(y))}{f F^{-1} L_n(y)} dy - R_1, \end{aligned} \tag{4.24}$$

where

$$R_1 = \frac{1}{2} n^{\frac{1}{2}} \int_0^1 \{U_n L_n(y) - L_n(y)\}^2 \frac{f' F^{-1} L_n(\delta_y)}{f^2 F^{-1} L_n(\delta_y)} dy$$

and

$$min(U_n L_n(y), L_n(y)) < \delta_y < max(U_n L_n(y), L_n(y)).$$

By Definition 2.1 ii, a result of Goldie [28] and arguments similar to those used in Csörgó-Révész [16], we can prove that as $n \to \infty$, $R_1 \xrightarrow{P} 0$ and $R_2 \xrightarrow{P} 0$.

Hence in terms of $l_n(y) = n^{\frac{1}{2}} \{L_n(y) - L(y)\}$ we have

$$\begin{aligned} n^{\frac{1}{2}}(I_n - I) &\stackrel{P}{=} \int_0^1 \frac{l_n(y)}{f F^{-1} L(y)} dy + \int_0^1 \frac{u_n(L_n(y))}{f F^{-1} L_n(y)} dy + o(1) \\ &\stackrel{P}{=} \lambda_{1,n} + \lambda_{2,n} + o(1). \end{aligned} \tag{4.25}$$

Csörgó et. al. [15] proved that

$$l_n(y) \xrightarrow{D} \Lambda(y)$$

and

$$u_n(y) \xrightarrow{D} B(y),$$

where $B(\cdot)$ is a Brownian bridge [2] and

$$\Lambda(y) = \frac{1}{\mu}\left\{\int_0^y B(t)\,dF^{-1}(t) - L(y)\int_0^1 B(t)dF^{-1}(t)\right\}$$

Consequently, as $n \to \infty$,

$$\lambda_{1,n} \xrightarrow{D} \int_0^1 \frac{\Lambda(y)}{fF^{-1}L(y)}\,dy \tag{4.26}$$

and

$$\lambda_{2,n} \xrightarrow{D} \int_0^1 \frac{B(L(y))}{fF^{-1}L(y)}\,dy. \tag{4.27}$$

By (4.25), (4.26) and (4.27), we have

$$n^{\frac{1}{2}}(I_n - I) \xrightarrow{D} \int_0^1 \frac{\Lambda(y)}{fF^{-1}L(y)}\,dy + \int_0^1 \frac{B(L(y))}{fF^{-1}L(y)}\,dy \tag{4.28}$$

as $n \to \infty$.

By (4.21), (4.28) and Slutsky's Theorem, we have as $n \to \infty$

$$\begin{aligned} n^{\frac{1}{2}}\{\hat{\xi}_n - \xi\} \quad \xrightarrow{D} \quad & \frac{1}{\mu}\int_0^1 \frac{B(L(y))}{fF^{-1}L(y)}\,dy + \frac{1}{\mu^2}\int_0^1 \frac{1}{fF^{-1}L(y)}\int_0^y B(t)\,dF^{-1}(t)\,dy \\ & - \frac{1}{\mu^2}\int_0^1 \frac{L(y)}{fF^{-1}L(y)}\,dy \int_0^1 B(t)\,dF^{-1}(t) \\ & - \frac{(1-\xi)}{\mu}\int_0^1 B(t)\,dF^{-1}(t) = S. \end{aligned} \tag{4.29}$$

[2]The process $\{B(x), 0 \le x \le 1\}$ is said to be a Brownian Bridge if it is Gaussian with mean zero and covariance function $E[B(s)B(t)] = s \wedge t - st$, where $\wedge$ means minimum(s,t).

Note that in (4.29), we use the result that

$$n^{\frac{1}{2}}(\bar{x}-\mu) \xrightarrow{D} \int_0^1 B(t)\, dF^{-1}(t), \text{ as } n \to \infty.$$

By letting $t = L(y)$ and use the fact the $L'(y) = \frac{1}{\mu}F^{-1}(y)$ and $dF^{-1}(t) = \frac{dt}{fF^{-1}(t)}$, we can show that

$$S = \int_0^1 \left\{ \frac{1}{F^{-1}L^{-1}(t)} + \frac{g(t)}{\mu^2} - \frac{c1}{\mu^2} - \frac{(1-\xi)}{\mu} \right\} B(t)\, dF^{-1}(t) \tag{4.30}$$

and $h(t)$, $g(t)$ *and* $c1$ are as in (4.18),(4.19), (4.20), respectively.

Note that S is a mean zero normal random variable with variance $\sigma_\xi^2(F)$ of (4.17). This completes the proof of Theorem 4.1.

4.4.3 Nonparametric Estimator for $\sigma_\xi^2(F)$

The Direct Nonparametric Variance Estimator $\hat{\sigma}^2(\hat{\xi}_n)$

Let $F_n(\cdot$, and $L_n(\cdot)$ be the empirical distribution function and the empirical Lorenz function of the given sample and $\bar{x}$ be the sample mean. The direct nonparametric estimator of $\sigma_\xi^2(F)$ is obtained by replacing $F^{-1}(\cdot)$ by $F_n^{-1}(\cdot)$ in (4.17). The resulting estimator is given by

$$\begin{aligned} \hat{\sigma}_\xi^2(F_n) &= \hat{\sigma}^2(\hat{\xi}_n) \qquad (4.31) \\ &= 2\sum_{j=1}^{n-1} \left\{ \left(1-\frac{j}{n}\right) h_n\left(\frac{j}{n}\right) \sum_{i=1}^{j} \frac{i}{n} h_n\left(\frac{i}{n}\right)(x_{i+1:n}-x_{i:n}) \right\} (x_{j+1:n}-x_{j:n}), \end{aligned}$$

where

$$\begin{aligned} h_n\left(\frac{k}{n}\right) &= \frac{1}{F_n^{-1}L_n^{-1}(\frac{k}{n})} + \frac{g_n(\frac{k}{n})}{\bar{x}^2} - \frac{c1_n}{\bar{x}^2} - \frac{(1-\xi_n)}{\bar{x}} \\ c1_n &= \bar{x}\sum_{j=1}^{n-1} \frac{\frac{j}{n}}{F_n^{-1}L_n^{-1}(\frac{j}{n})}(x_{j+1:n}-x_{j:n}) \\ g_n\left(\frac{k}{n}\right) &= \bar{x}\sum_{j=[n\,L_n(\frac{k}{n})]}^{n-1} \frac{1}{F_n^{-1}L_n^{-1}(\frac{j}{n})}(x_{j+1:n}-x_{j:n}) \end{aligned}$$

and

$$L_n^{-1}\left(\frac{k}{n}\right), F_n^{-1}L_n^{-1}\left(\frac{k}{n}\right) \text{ and } L_n\left(\frac{k}{n}\right)$$

are as defined in (4.13), (4.8) and (4.11), respectively.

Nonparametric Variance Estimator via Quantile Density Estimation Approach

In this section, we derive another expression for the variance in terms of the quantile density function where $q(\cdot) = \frac{1}{fF^{-1}(\cdot)}$. From here on, we will refer to this alternative expression as the quantile density approach.

Let S be as in (4.29), it can be shown that

$$S = S_1 + S_2 - S_3,$$

$$\begin{aligned} S_1 &= \frac{1}{\mu}\int_0^1 \frac{B(y)}{fF^{-1}(y)}\, dL^{-1}(y), \\ S_2 &= \frac{1}{\mu^2}\int_0^1 B(y)g(y)\, dF^{-1}(y), \\ S_3 &= \frac{1}{\mu^2}\{\lambda + \mu(1-\xi)\}\int_0^1 B(x)\, dF^{-1}(x), \end{aligned}$$

$$g(x) = \int_x^1 \frac{dt}{fF^{-1}L(t)} = \int_x^1 \frac{1}{fF^{-1}(z)}\, dL^{-1}(z),$$

$$\lambda = \int_0^1 \frac{L(y)}{fF^{-1}L(y)}\, dy = \int_0^1 \frac{t}{fF^{-1}(t)}\, dL^{-1}(t).$$

Hence,

$\sigma_\xi^2(F) = Var(S) = E[S]^2 = \sigma_1 + \sigma_2 + \sigma_3 + 2\sigma_{12} - 2\sigma_{13} - 2\sigma_{23}$, where $\sigma_i = E[S_i]^2$ and $\sigma_{ij} = E[S_i S_j]$ are given by

$$\begin{aligned}
\sigma_1 &= \frac{2}{\mu^2}\int_0^1 \frac{(1-y)}{fF^{-1}(y)}\int_0^y \frac{x}{fF^{-1}(x)}\,dL^{-1}(x)\,dL^{-1}(y),\\
\sigma_2 &= \frac{2}{\mu^4}\int_0^1 (1-y)\,g(y)\int_0^y x\,g(x)\,dF^{-1}(x)\,dF^{-1}(y),\\
\sigma_3 &= \frac{2}{\mu^4}\{\lambda+\mu(1-\xi)\}^2\int_0^1 y\int_y^1 (1-x)\,dF^{-1}(x)\,dF^{-1}(y),
\end{aligned}$$

$$\begin{aligned}
\sigma_{12} &= \frac{1}{\mu^3}\int_0^1\int_0^1 \left\{\frac{x\wedge y - xy}{fF^{-1}(x)}\right\} g(y)\,dL^{-1}(x)\,dF^{-1}(y),\\
&= \frac{1}{\mu^3}\int_0^1 g(y)\,(1-y)\int_0^y \frac{x}{fF^{-1}(x)}\,dL^{-1}(x)dF^{-1}(y)\\
&\quad + \frac{1}{\mu^3}\int_0^1 y\,g(y)\int_y^1 \frac{(1-x)}{fF^{-1}(x)}\,dL^{-1}(x)dF^{-1}(y),\\
&= \frac{1}{\mu^3}\int_0^1 \frac{(1-x)}{fF^{-1}(x)}\int_0^x y\,g(y)\,dF^{-1}(y)dL^{-1}(x)\\
&\quad + \frac{1}{\mu^3}\int_0^1 \frac{x}{fF^{-1}(x)}\int_x^1 (1-y)\,g(y)\,dF^{-1}(y)dL^{-1}(x),
\end{aligned}$$

$$\begin{aligned}
\sigma_{13} &= \frac{1}{\mu^3}\{\lambda+\mu(1-\xi)\}\int_0^1\int_0^1 \left\{\frac{x\wedge y - xy}{fF^{-1}(x)}\right\} dF^{-1}(y)dL^{-1}(x),\\
&= \frac{1}{\mu^3}\{\lambda+\mu(1-\xi)\}\int_0^1 \frac{(1-x)}{fF^{-1}(x)}\int_0^x y\,dF^{-1}(y)dL^{-1}(x)\\
&\quad + \frac{1}{\mu^3}\{\lambda+\mu(1-\xi)\}\int_0^1 \frac{x}{fF^{-1}(x)}\int_x^1 (1-y)\,dF^{-1}(y)dL^{-1}(x),
\end{aligned}$$

and

$$\begin{aligned}
\sigma_{23} &= \frac{1}{\mu^4}\{\lambda+\mu(1-\xi)\}\int_0^1\int_0^1 \{x\wedge y - xy\}\,g(x)dF^{-1}(y)dF^{-1}(x),\\
&= \frac{1}{\mu^4}\{\lambda+\mu(1-\xi)\}\int_0^1 (1-x)\,g(x)\int_0^x y\,dF^{-1}(y)\,dF^{-1}(x)\\
&\quad + \frac{1}{\mu^4}\{\lambda+\mu(1-\xi)\}\int_0^1 x\,g(x)\int_x^1 (1-y)\,dF^{-1}(y)\,dF^{-1}(x).
\end{aligned}$$

Let $\hat{q}(\cdot)$ be an appropriate estimator for $q(\cdot)$. We propose to estimate σ_i and σ_{ij} by $\hat{\sigma}_{i,n}$ and $\hat{\sigma}_{ij,n}$ which are given below.

$$
\begin{aligned}
\hat{\sigma}_{1n} &= \frac{2}{(n\bar{x})^2}\sum_{i=2}^{n-1}(1-Y_i)\,\hat{q}(Y_i)\sum_{j=1}^{i-1}Y_j\,\hat{q}(Y_j),\\
\hat{\sigma}_{2n} &= \frac{2}{\bar{x}^4}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)\hat{g}_n\left(\frac{j}{n}\right)\hat{\psi}_n\left(\frac{j}{n}\right)(x_{j+1:n}-x_{j:n}),\\
\hat{\sigma}_{3n} &= \frac{2}{\bar{x}^4}\left\{\hat{\lambda}_n+\bar{x}(1-\hat{\xi}_n)\right\}^2\sum_{i=1}^{n-1}(i/n)\left(\sum_{j=i}^{n-1}(1-j/n)\,(x_{j+1:n}-x_{j:n})\right)(x_{i+1:n}-x_{i:n}),\\
\hat{\sigma}_{12n} &= \frac{1}{n\bar{x}^3}\sum_{i=1}^{n-1}(1-Y_i)\hat{q}(Y_i)\sum_{j=1}^{[nY_i]}(j/n)\,\hat{g}(j/n)\,(x_{j+1:n}-x_{j:n})\\
&\quad+\frac{1}{n\bar{x}^3}\sum_{i=1}^{n-1}Y_i\,\hat{q}(Y_i)\sum_{j=[nY_i]+1}^{n-1}(1-j/n)\,\hat{g}(j/n)\,(x_{j+1:n}-x_{j:n}),\\
\hat{\sigma}_{13n} &= \frac{1}{n\bar{x}^3}\left\{\hat{\lambda}_n+\bar{x}(1-\hat{\xi}_n)\right\}\sum_{i=1}^{n-1}(1-Y_i)\,\hat{q}(Y_i)\sum_{j=1}^{[nY_i]}(j/n)\,(x_{j+1:n}-x_{j:n})\\
&\quad+\frac{1}{n\bar{x}^3}\left\{\hat{\lambda}_n+\bar{x}(1-\hat{\xi}_n)\right\}\sum_{i=1}^{n-1}Y_i\,\hat{q}(Y_i)\sum_{j=[nY_i]+1}^{n-1}(1-j/n)(x_{j+1:n}-x_{j:n}),\\
\hat{\sigma}_{23n} &= \frac{1}{\bar{x}^4}\left\{\hat{\lambda}_n+\bar{x}(1-\hat{\xi}_n)\right\}\sum_{i=2}^{n-1}(1-i/n)\,\hat{g}(i/n)\,(x_{i+1:n}-x_{i:n})\sum_{j=1}^{i-1}(j/n)\,(x_{j+1:n}-x_{j:n})\\
&\quad\frac{1}{\bar{x}^4}\left\{\hat{\lambda}_n+\bar{x}(1-\hat{\xi}_n)\right\}\sum_{i=2}^{n-1}(i/n)\,\hat{g}(i/n)\,(x_{i+1:n}-x_{i:n})\sum_{j=i}^{n-1}(1-j/n)\,(x_{j+1:n}-x_{j:n}),
\end{aligned}
$$

where

$$
\begin{aligned}
Y_i &= L_n\left(\frac{i}{n}\right)=\frac{1}{n\bar{x}}\sum_{j=1}^{i}x_{j:n},\\
\hat{g}_n(j/n) &= \int_{\frac{1}{n}}^{1}\hat{q}(z)\,dL_n^{-1}(z)=\frac{1}{n}\sum_{i=nL_n^{-1}(j/n)}^{n}\hat{q}(Y_i),\\
\hat{\psi}_n(j/n) &= \sum_{k=1}^{j-1}(k/n)\,\hat{g}_n(k/n)\,(x_{k+1:n}-x_{k:n}),\\
\hat{\lambda}_n &= \frac{1}{n}\sum_{i=1}^{n-1}\hat{q}(Y_i)Y_i.
\end{aligned}
$$

The problem of estimating the quantile density $q(\cdot) = \frac{1}{fF^{-1}(\cdot)}$ is explored through a 'histogram-type' estimator or through a 'kernel-type' estimator.

The *histogram-type* estimator suggested by Siddiqui [53] and investigated by Block and Gastwirth [12] and consequently used and modified by Falk [22] is of the form:

$$\begin{aligned} \hat{q}_1(p) &= \frac{\{F_n^{-1}(p+b) - F_n^{-1}(p-b)\}}{2b}, \\ &= \frac{\{x_{<np+nb>:n} - x_{<np-nb>:n}\}}{2b}, \end{aligned} \tag{4.32}$$

where b is the bin width, bandwidth or smoothing parameter and satisfies the conditions that $b = b(n) > 0$, $b \to 0$ and $nb \to \infty$ as $n \to \infty$. The optimal value of b is $b = cn^{\frac{4}{5}}$ where $c = \left[\frac{9q^2(p)}{2(q''(p))^2}\right]^{\frac{1}{2}}$, $q''(p) = .\left[\frac{d^2q(p)}{dp^2}\right]$.

Kernel-type estimators are proposed by various authors (see Jones [31], Babu and Rao [6], Babu [5], Falk [22], and Muller [40]). We draw on the work of Jones such that the Kernel estimator is of the form:

$$\begin{aligned} \hat{q}_2(u) &= \sum_{i=2}^{n} (x_{i:n} - x_{i-1:n}) \frac{1}{b} k\left(\frac{u - \frac{(i-1)}{n}}{b}\right) \\ &\quad - x_{n:n}\frac{1}{b}k\left(\frac{u-1}{b}\right) + x_{1:n}\frac{1}{b}k\left(\frac{u}{b}\right), \end{aligned}$$

where $k(\cdot)$ is a symmetric probability density function say,

$$k(u) = \begin{cases} 0 & u < -1 \\ 1 - |u| & -1 < u < 1 \\ 0 & u > 1 \end{cases} \tag{4.33}$$

We take $b = b(n) = \frac{1}{(n\log(n))^{\frac{1}{3}}}$ which satisfies the condition that $b(n) \to 0$ and $nb(n) \to \infty \, as \, n \to \infty$.

4.4.4 Exact Variance Results

Exact Variance Under the Lognormal Distribution

Recall (4.17), (4.18), (4.19) and (4.20) for $\sigma_\xi^2(F)$, $h(t)$, $g(t)$, and cl, respectively. Under the Lognormal, the following expressions are derived:

$$\begin{aligned}
\mu &= \exp(\gamma + \frac{1}{2}\delta^2) \\
L(t) &= \Phi\left(\Phi^{-1}(t) - \delta\right) \\
F^{-1}L^{-1}(t) &= \exp\left(\gamma + \delta^2 + \delta\Phi^{-1}(t)\right) \\
F^{-1}(t) &= \exp\left(\gamma + \delta\Phi^{-1}(t)\right) \\
dF^{-1}(t) &= \delta\exp\left(\gamma + \delta\Phi^{-1}(t)\right) d\Phi^{-1}(t).
\end{aligned}$$

It can be shown that:

$$h(y) = \frac{1}{\exp(\gamma + \delta^2 + \delta\Phi^{-1}(y))} - \frac{\delta}{\exp(\delta^2)} - \frac{\Phi^{-1}L(t)}{\exp(\gamma + \frac{1}{2}\delta^2)} - \frac{\exp(-\delta^2)}{\exp(\gamma + \frac{1}{2}\delta^2)}$$

Exact Variance under the Pareto

Recall the formulas under (4.17), (4.18), (4.19) and (4.20) for $\sigma_\xi^2(F)$, $h(t)$, $g(t)$, and cl, respectively.

Under the Pareto distribution on $(1, \infty)$, the following expressions are derived:

$$\begin{aligned}
\mu &= \frac{\theta}{\theta - 1} \\
L(t) &= 1 - (1 - t)^{\frac{\theta - 1}{\theta}} \\
F^{-1}L^{-1}(t) &= (1 - t)^{-\frac{1}{\theta - 1}} \\
F^{-1}(t) &= (1 - t)^{-\frac{1}{\theta}} \\
dF^{-1}(t) &= \frac{1}{\theta}(1 - t)^{-(1/\theta + 1)} \\
\xi &= \frac{1}{1 + \theta(\theta - 1)}
\end{aligned}$$

$$\begin{aligned}
g(t) &= \frac{\mu}{\theta}\frac{(1 - t)^{a3}}{a0}; \quad cl = \frac{\mu}{\theta}B(2, a0); \text{ and} \\
h(t) &= \frac{1}{(1 - t)^{a4}} + \frac{(1 - t)^{a3}}{\mu\theta a0} - \frac{B(2, a0)}{\mu\theta} - \frac{(1 - \xi)}{\mu}.
\end{aligned}$$

where $a0 = \frac{1}{\theta-1} - \frac{1}{\theta}; a1 = \frac{1}{\theta} - \frac{1}{\theta^2}; a2 = \frac{1}{\theta}; a3 = \frac{1}{\theta^2}; a4 = \frac{1}{\theta-1}; B(2, a0) \approx Beta(2, a0)$

Plug in $h(t)$ and $dF^{-1}(t)$ into (4.17) and evaluate the integral in terms of the parameter θ. The derived variance is called the *exact variance* under the Pareto Model in this paper.

4.5 Simulations

Monte Carlo simulation studies are conducted in order to establish the unbiasedness and consistency properties of $\hat{\xi}_n$ and $\hat{\sigma}^2(\hat{\xi}_n)$ as estimators of ξ and of $\sigma^2(\hat{\xi}_n)$, respectively. Two leading models of income distributions are assumed under study, namely, the Lognormal and Pareto.

Latorre [38] proposed parametric estimates of ξ under several models of income distributions. Part of his results are reported in Table (4.1) and Table (4.2) together with the results of this research.

The simulation study are shown in Section 4.5.1 for the direct approach while results for the quantile density estimation approach in Section 4.4. The MTS FORTRAN language together with the subroutines in IMSL(1992) are used in the computer simulation programs (refer to the appendix for the computer programs). For sample size less than or equal to 100, 1000 generations are taken. For sample size 1000, only 300 repetitions are done. Analysis and interpretation of results are covered in Section 4.3.

4.5.1 Monte Carlo Simulation - Direct Approach

Notations

1. $E_{MC}(\hat{\xi}_n)$ = Monte Carlo estimate of $\hat{\xi}_n$;

2. $Var_{MC}(\hat{\xi}_n)$ = Monte Carlo variance estimate of $\hat{\xi}_n$;

3. $E_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$ = Monte Carlo estimate of $\sigma^2_{MC}(\hat{\xi}_n)$;

4. $Var_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$ = Monte Carlo variance estimate of $\sigma^2(\hat{\xi}_n)$;

5. $\sigma^2_{exact}(\hat{\xi}_n)$ = exact variance under the population by evaluating (4.17);

6. $\sigma^2_{par}(\xi^Z)$ = 'parametric' variance , ie, variance obtained by the linearization method;

7. $\hat{\sigma}^2(\hat{\xi}_n)$ = nonparametric asymptotic variance estimator;

Under the Lognormal Model

X $\rightsquigarrow$ lognormal with $\gamma = 2.8$ and $\delta = 0.35$

$$\begin{aligned}
f(x) &= \begin{cases} \frac{1}{\delta\sqrt{2\pi}}\frac{1}{2}\exp\left(-\frac{1}{2}\left[\frac{\log x-\gamma}{\delta}\right]^2\right) & x > 0 \\ 0 & otherwise \end{cases} \\
F(x) &= \Phi\left(\frac{\log x - \gamma}{\delta}\right) \\
\xi &= 1 - \exp(-\delta^2) = 0.1153; \\
nVar_{MC}(\hat{\xi}_n) &\approx 0.03, \text{ n is the number of observations.}
\end{aligned} \tag{4.34}$$

The value for the exact variance $\sigma^2_{exact}(\hat{\xi}_n)$ is not evaluated. We rely on (4.34) as gauge of the 'exact' variance under the Lognormal.

Under Pareto Model

$X \rightsquigarrow$ Pareto on $(1, \infty)$ with $\theta = 2.9$.

The asymptotic variance under the Pareto does not exist for $\theta \leq 2$.

$$\begin{aligned}
f(x) &= \theta\, x^{-\theta-1},\ x \geq 1 > 0, \\
F(x) &= 1 - x^{-\theta} = 1 - \left(\frac{1}{x}\right)^\theta
\end{aligned}$$

Table 4.1: Simulation Results under Lognormal Model

Properties	Sample Size 100	Sample Size 1000
$E_{MC}(\hat{\xi}_n)$	0.12190	0.11571
$Var_{MC}(\hat{\xi}_n)$	0.00026	0.00003
$E_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.02783	0.02636
$Var_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.00026	0.00003
coverage probability (95 %)		
Using NONPAR Variance	96 %	93 %
Using PARAMETRIC Variance	91 %	91 %

$$\begin{aligned} \xi &= \frac{1}{1+\theta(\theta-1)} = 0.1536 \\ \sigma^2_{exact}(\hat{\xi}_n) &= 0.4932 \end{aligned}$$

Note that $\sigma^2_{exact}(\hat{\xi}_n)$ under this model is evaluated (see Section 4.4.4). This number has been verified using two approaches, namely by evaluating equation (4.17) and by the expression of the variance in section 3.3.3 (via quantile density).

Table 4.2: Simulation Results under Pareto Model

Properties	Sample Size 100	Sample Size 1000
$E_{MC}(\hat{\xi}_n)$	0.15486	0.15408
$Var_{MC}(\hat{\xi}_n)$	0.00293	0.00041
$E_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.22466	0.44118
$Var_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.16901	0.54548
coverage probabilities (95 %)		
Using NONPAR Variance	80 %	89 %
Using PARAMETRIC Variance	95 %	95 %

4.5.2 Monte Carlo Simulation - Quantile Density Approach

In this section, we determine whether the quantile density $\hat{q}(\cdot)$ which appears in the expression of $\sigma^2(\hat{\xi}_n)$, improves the estimation of this asymptotic variance of $\hat{\xi}_n$. Nonparametric estimation results were good in the lognormal case, for this reason, we will only consider the *Pareto distribution* in this subsection.

There is a tendency for spurious noise to appear in the tails of the estimates of $\hat{q}(\cdot)$. Hence, we need to adjust $\hat{q}_1(\cdot)$ appropriately to obtain accurate estimates aside from employing variable smoothing parameters. Similarly, the symmetric density function K is variable at the tails of the distribution.

Using the Histogram-Type Estimator

For $0.26 < p < 0.85$,

$$\hat{q}_1(p) = \frac{\{X_{<np+nb>:n} - X_{<np-nb>:n}\}}{2b}, \quad b > 0.$$

For $p \geq 0.85$,

$$\hat{q}_1(p) = \frac{\{X_{<np>:n} - X_{<np-nb>:n}\}}{b}, \quad b > 0.$$

For $p \leq 0.26$,

$\hat{q}_1(p) = \frac{\{X_{<np+nb>:n} - X_{<np>:n}\}}{b}, \; b > 0.$
where $b = \frac{0.85336(1-p)^{\frac{4}{5}}}{n^{\frac{1}{5}}}$.

Using Kernel-Type Estimator

Recall (4.33) the Jones Kernel estimator:

$$
\begin{aligned}
\hat{q}_2(u) &= \sum_{i=2}^{n} (X_{i:n} - X_{i-1:n}) \frac{1}{b} k\left(\frac{u - \frac{(i-1)}{n}}{b}\right) \\
&\quad - X_{n:n} \frac{1}{b} k\left(\frac{u-1}{b}\right) + X_{1:n} \frac{1}{b} k\left(\frac{u}{b}\right)
\end{aligned}
$$

For $0.10 < p < 0.85$,

$$K(u) = \begin{cases} 1 - |u| & -1 < u < 1 \\ 0 & else \end{cases}$$

For $p \leq 0.10$,

$$K(u) = \begin{cases} 2(1-u) & 0 \leq u \leq 1 \\ 0 & else \end{cases}$$

For $p \geq 0.85$,

$$K(u) = \begin{cases} 2(1+u) & -1 \leq u \leq 0 \\ 0 & else \end{cases}$$

where

$b = b(n) = \frac{1}{(n log(n))^{1/3}}$.

4.6 Interpretation and Analysis of Results

By Theorem 3.1, we have shown that for large n, $\hat{\xi}_n$ is asymptotically normal, unbiased and efficient with asymptotic variance given by (4.17). Under the population Lognormal and Pareto models, these results are checked based on the knowledge of a random sample from these populations. The estimators $\hat{\xi}_n$ is still unbiased for very large n. In order to check empirically whether our proposed procedure is effective in providing correct inferences about the population, we judge by the coverage probabilities for ξ. The coverage probability under the Lognormal is very close to the

Table 4.3: Simulation Results under the Pareto (Quantile Density)

Properties	Sample Size 40	Sample Size 100	Sample Size 1000
HISTOGRAM-TYPE			
$E_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.3192	0.3145	0.4681
$Var_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.1817	0.2631	0.5576
coverage probability (95%)	86%	86%	91%
KERNEL-TYPE			
$E_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.2868	0.3038	0.4816
$Var_{MC}\left(\hat{\sigma}^2(\hat{\xi}_n)\right)$	0.1442	0.2496	0.6038
coverage probability (95%)	87%	86 %	91%

expected true coverage of 95% for both $n = 100$ and $n = 1000$. However, under the Pareto Model this is not the case.

We then verify the unbiasedness and consistency properties of $\hat{\sigma}^2(\hat{\xi}_n)$. Based on the sample, under both models, the asymptotic variance estimator is unbiased for $n = 1000$. We are not able to conclude fully the stability of our nonparametric variance estimator, in the sense of having a small variance, based on the Monte Carlo variance estimator of $\sigma^2(\hat{\xi}_n)$. We try another approach at estimating this variance. Since the terms in $\sigma^2(\hat{\xi}_n)$ depends on the quantity $q(\cdot) = \frac{1}{fF^{-1}(\cdot)}$, we investigate quantile density estimators for $q(\cdot)$. Upon using this approach to get an alternative estimator for $\sigma^2(\hat{\xi}_n)$, we arrive at slightly improved results. By employing an estimator of $q(\cdot)$ which is more efficient than the ones conducted, we would greatly improve our estimator of $\sigma^2(\xi)$. This would result in a coverage probability much closer to 95%.

Bibliography

[1] Aly, E.E.A.A.(1990). Tests for Monotinicity Properties of the mean residual life function. *Scand J. Statistics.* 17:189-200.

[2] Aly, E.E.A.A.(1991). On Testing for Lorenz Ordering. *Metrika.* Vol. 38, pp. 117-124.

[3] Aly, E.E.A.A.(1993). On Testing for Decreasing Mean Residual Life Ordering. *Naval Research Logistics*, Vol.40, pp. 633-642.

[4] Atkinson, A.B. (1970). On the Measurement of Inequality. *Journal of Economic Theory.* Vol. 12, pp. 244-263.

[5] Babu, G.J. (1986). Efficient Estimation of the Reciprocal of the Density Quantile Function at a Point. *Statistics and Probability Letters.* Vol. 4, pp.133-139.

[6] Babu, G.J. and Rao, C. (1990). Estimation of the Reciprocal of the Density Quantile Function at a Point. *Journal of Multivariate Analysis.* Vol. 33, pp. 106-124.

[7] Barlow, R.E. and F. Proschan (1981). Statistical Theory in Reliability and Life Testing. To Begin With, Silver Springs, MD.

[8] Barlow, R.E. and R.Campo (1975). Total time on test processes and applications to failure data analysis. In Reliability and fault tree analysis, pp. 451-481, SIAM, Philadelphia.

[9] Beach, C.M. and R. Davidson (1983). Distribution-Free Statistical Inference with Lorenz Curves and Income Shares. *Review of Economic Studies.* Vol. 50, pp. 723-735.

[10] Beach C.M. and J. Richmond. (1985). Joint Confidence Intervals for Income Shares and Lorenz Curves. *International Economic Review.* Vol. 26, pp. 439-450.

[11] Bishop, J.A., J. Formby and P. Thistle (1988). Statistical Inference, Income Distributions and Social Welfare. *Research on Income Inequality, Vol.1, pp. 49-82.*

[12] Block, D. and Gastwirth, J.(1968). On a Simple Estimate of the Reciprocal of the Density Function. *The Annals of Mathematical Statistics.* Vol. 39, No. 9, pp. 1083-1085.

[13] Chandra, M. and N.D. Singpurwalla (1981). Relationships Between Some Notions which are Common to Reliability Theory and Economics. *Mathematics of Operations Research.* Vol. 6, No.1.

[14] Csorgo, et.al.(1986). An asymptotic Theory for Empirical Reliability and Concentration Processes. *Lecture Notes in Statistics.* Vol.33.

[15] Csorgo, M.(1983). Quantile Processes with Statistical Applications (CBMS-NSF Regional Conference Series in Applied Mathematics 4.2). SIAM, Philadelphia.

[16] Csorgo, M. and P. Revesz (1978). Strong Approximation of the Quantile Process. *Annals of Statistics.* Vol. 6, pp. 882-894.

[17] Dagum, C.(1980). The Generation and Distribution of Income, the Lorenz Curve and the Gini Ratio. *Economie Appliquee.* Vol. 33, n.2, pp.327-367.

[18] Dagum, C. - Zenga, M.(eds)(1990). *Income and Wealth Distributions, Income Inequality and Poverty.* Berlin, Springer-Verlag.

[19] Dagum,C. Generation and Properties of Income Distribution Functions in C. Dagum-M. Zenga.

[20] Dancelli, L.(1990), On the Behavior of the $Z(,)$ Concentration Curve, in C.Dagum-M.Zenga(eds)(1990)., pp.111-127.

[21] Dasgupta, P., A.K. Sen and D. Starret (1973). Notes on the Measurement of Inequality. *Journal of Economic Theory.* Vol. 6, pp. 180-187.

[22] Falk, M. (1986). On the Estimation of the Quantile Density Function. *Statistics and Probability Letters.* Vol. 4, pp. 69-73.

[23] Frosini, B.V. Point Measures and Global Measures of Inequality. *Giornale degli Economisti e Annali di Economia.* 69: 83-105.

[24] Gastwirth, J. (1974). Large Sample Theory of Some Measures of Income Inequality. *Econometrica.* Vol. 42, pp. 191-196.

[25] Gastwirth, J.L. and M.H. Gale (1978). A Scale-Free Goodness of Fit Test for the Exponential Distribution Based on the Lorenz Curve. *Journal of the American Statistical Association.* Vol. 73, pp. 787-793.

[26] Gastwirth, J.L. and M.H. Gale (1985). Simple Asymptotically Distribution-Free Methods for Comparing Lorenz Curves and Gini Indices Obtained from Complete Data. In R.L. Basmann and G.F. Rhodes, Jr. (eds.). Advances in Econometrics, Vol4.

[27] Gini, C. (1912). Variabilita e mutabilita contribut allo studio della distribuzioni e relazioni statistiche. Studi Economic-Giuridici della facol'a di Guirisprodenza dell' Universita di Cagliari, III parts II.

[28] Goldie,M.(1977). Convergence Theorems for Empirical Lorenz Curves and Their Inverses. *Advances in Appl. Probability.* Vol. 9, pp.765-791.

[29] Guess, F.,M. Hollander and F. Proschan. (1986) Testing exponentiality versus a trend change in mean residual life. *Ann. Statist*, Vol.14, pp.1388-1398.

[30] Hollander, M. and F. Proschan (1984). Nonparametric Concepts and Methods in Reliability. In P.R. Krishnaiah and P.K. Sen (eds), Handbook of Statistics, Vol. 4, pp. 613-655.

[31] Jones, M.C. (1992). Estimating Densities, Quantiles, Quantile Densities and Density Quantiles. *Ann. Inst. Statist. Math.* Vol. 44, No. 4, pp. 721-727.

[32] Kakwani, N.C. (1974). A Note on the Efficient Estimation of the New Measures of Income Inequality. *Econometrica.* Vol. 42, pp. 597-600.

[33] Kakwani, N.C. (1977). Applications of Lorenz Curves in Economic Analysis. Vol. 45, No.3, pp.719-727.

[34] Kakwani, N.C. and N. Podder (1973). On the Estimation of Lorenz Curves from Grouped Observations. *International Economic Review.* Vol. 14, pp. 278-292.

[35] Klefsjo, B. (1983) Testing exponentiality against HNBUE. *Scandinavian Journal of Statistics.* Vol. 10, pp. 65-75.

[36] Klefsjo, B. (1984). Reliability interpretations of Some Concepts from Economics. *Naval Research Logistic Quarterly.* Vol. 31, pp.301-308.

[37] Kolm, S.-C. (1969). Unequal Inequalities:I. *Journal of Economic Theory.* Vol. 12, pp. 416-442.

[38] Latorre, G.(1990). Asymptotic Distributions of Indices of Concentration: Empirical Verification and Applications, in C.Dagum-M.Zenga(eds)(1990), pp.149-169.

[39] Loren, M.O. (1893). Methods of Measuring the Concentration of Wealth. *Journal of American Statistical Association* Vol. 9, pp. 209-219.

[40] Muller, H-G. (1991). Smooth Optimum Kernel Estimators Near Endpoints. *Biometrika.* Vol. 78, No. 3, pp. 521-530.

[41] Nygard, F.-Sandstrom, A.(1981). *Measuring Income Inequality.* Stockholm, Almquist & Wiksell.

[42] Pareto, V.(1895). La legge della domanda. *Giornale degli Economisti Gennaio.* pp. 59-68.

[43] Pollastri, A. (1987). Le curve di concentrazione Lp e Zp nella distribuzione log-normale generalizzata. *Giornale degli Economisti ed Annali di Economia.* pp. 632-663.

[44] Pyatt, G. (1989). Social Evaluation Criteria. In C. Dagum-M. Zenga (eds) (1990), pp. 243-253.

[45] Richmond, J. (1982). A General Method for Constructing Simultaneous Confidence Intervals. *Journal of the American Statistical Association.* Vol. 77, pp. 455-460.

[46] Rothschild, M. and J.E. Stiglitz (1973). Some Further Results on the Measurement of Inequality. *Journal of Economic Theory.* Vol. 6, pp. 188-204.

[47] Salvaterra, T. Comparisons Among Concentration Curves and Indices in Some Empirical Distributions, in C. Dagum-M. Zenga(eds)(1990), pp.194-213.

[48] Saposnik, R. (1981). Rank Dominance in Income Distribution. *Public Choice.* Vol. 36, pp. 147-151.

[49] Sen, A.K. (1973). On Economic Inequality. Oxford:University Press.

[50] Sendler, W. (1979) On Statistical Inference in Concentration Measurement. *Metrika.* Vol. 26, pp. 109-122.

[51] Shorrocks, (1983). Ranking Income Distributions. *Econometrica.* Vol. 50, pp. 3-17.

[52] Stigler, S.M. (1974). Linear functions of order statistics with smooth weight functions. *Ann. Statist., Vol2,pp.676-693.*

[53] Siddiqui, M.M. (1960). Distribution of Quantiles in Samples from a Bivariate Population. *J. Res. Nat. Bur. Standards.* 64B, pp. 145-150.

[54] Zenga, M.(1984). Proposta per un indice di concentrazione basato sui rapporti fra quantili di popolazione e quantili di reddito. Giornale degli economisti e annali di economia, Vol. 43, n.5-6, pp.301-326.

[55] Zenga, M.(1990) Concentration Curves and Concentration Indices Derived from Them, in C. Dagum-M. Zenga (eds)(1990), pp.94-110.