Aaberge, Rolf

Working Paper
Asymptotic Distribution Theory of Empirical Rank-dependent Measures of Inequality

Discussion Papers, No. 402

Provided in Cooperation with:
Research Department, Statistics Norway, Oslo

Suggested Citation: Aaberge, Rolf (2005) : Asymptotic Distribution Theory of Empirical Rank-dependent Measures of Inequality, Discussion Papers, No. 402, Statistics Norway, Research Department, Oslo

This Version is available at:
http://hdl.handle.net/10419/192384

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

www.ECONSTOR.eu
Abstract:
A major aim of most income distribution studies is to make comparisons of income inequality across time for a given country and/or compare and rank different countries according to the level of income inequality. However, most of these studies lack information on sampling errors, which makes it difficult to judge the significance of the attained rankings.

The purpose of this paper is to derive the asymptotic properties of the empirical rank-dependent family of inequality measures. A favourable feature of this family of inequality measures is that it includes the Gini coefficients, and that any member of this family can be given an explicit and simple expression in terms of the Lorenz curve. By relying on a result of Doksum (1974) it is easily demonstrated that the empirical Lorenz curve, regarded as a stochastic process, converges to a Gaussian process. Moreover, this result forms the basis of the derivation of the asymptotic properties of the empirical rank-dependent measures of inequality.

Keywords: The Lorenz curve, the Gini coefficient, rank-dependent measures of inequality, nonparametric estimation methods, asymptotic distribution theory.

JEL classification: C14, D 63

Acknowledgement: I would like to thank Anne Skoglund for technical assistance and word processing.

Address: Rolf Aaberge, Statistics Norway, Research Department. E-mail: rolf.aaberge@ssb.no
Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:
Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no
1. Introduction

The standard practice in empirical analyses of income distributions is to make separate comparisons of the overall level of income (the size of the cake) and the distribution of income shares (division of the cake), and to use the Lorenz curve as a basis for analysing the distribution of income shares. By displaying the deviation of each individual income share from the income share that corresponds to perfect equality, the Lorenz curve captures the essential descriptive features of the concept of inequality.

When Lorenz curves do not intersect it is universally acknowledged that the higher Lorenz curve displays less inequality than the lower Lorenz curve. This is due to the fact that the higher of two non-intersecting Lorenz curves can be obtained from the lower Lorenz curve by means of rank-preserving income transfers from richer to poorer individuals. However, since observed Lorenz curves normally intersect weaker ranking criteria than the dominance criterion of non-intersecting Lorenz curves are required. In this case one may either search for weaker dominance criteria, see e.g. Shorrocks and Foster (1987), Dardanoni and Lambert (1988), Lambert (1993) and Aaberge (2000b), or one may apply summary measures of inequality. The latter approach also offers a method for quantifying the extent of inequality in income distributions, which may explain why numerous alternative measures of inequality are introduced in the literature. The most well-known and widely used measure of inequality is the Gini coefficient, which is equal to twice the area between the Lorenz curve and its equality reference. However, to get a broader picture of inequality than what is captured by the Gini coefficient the use of alternative measures of inequality is required.

By making explicit use of the Lorenz curve Mehran (1976), Donaldson and Weymark (1980, 1983), Weymark (1981), Yitzhaki (1983) and Aaberge (2000a, 2001) introduce various “generalized” Gini families of inequality measures. Moreover, Aaberge (2000a) demonstrates that one of these families, called the Lorenz family of inequality measures, can be considered as the moments of the Lorenz curve and thus provides a complete characterization of the Lorenz curve. This means that the Lorenz curve can be uniquely recovered from the knowledge of the corresponding Lorenz measures of inequality, i.e. without loss of information examination of inequality in an income distribution can be restricted to application of the Lorenz measures of inequality. Note that a subclass of the extended Gini family introduced by Donaldson and Weymark (1980, 1983) is uniquely determined by the Lorenz family of inequality measures.

---


2 See Aaberge (2000a) for a proof.
Since the different alternative “generalized” families of inequality measures can be considered as subfamilies of Mehran’s (1976) general family of rank-dependent measures of inequality it appears useful to consider the asymptotic properties of the empirical version of the general family of rank-dependent measures of inequality rather than to restrict to the empirical version of the Lorenz family of inequality measures.

The plan of the paper is as follows. Section 2 provides formal definitions of the Lorenz curve and the family of rank-dependent measures of inequality and the corresponding non-parametric estimators. By relying on a result of Doksum (1974) it is demonstrated in Section 3.1 that the empirical Lorenz curve (regarded as a stochastic process) converges to a Gaussian process. This result forms the basis of the derivation of the asymptotic properties of the empirical rank-dependent measures of inequality that are presented in Section 3.2.

2. Definition and estimation of the Lorenz curve and rank-dependent measures of inequality

Let $X$ be an income variable with cumulative distribution function $F$ and mean $\mu$. Let $\left[0, \infty\right)$ be the domain of $F$ where $F^{-1}$ is the left inverse of $F$ and $F^{-1}(0) \equiv 0$. The Lorenz curve $L$ for $F$ is defined by

$$L(u) = \frac{1}{\mu} \int_0^u F^{-1}(t) \, dt, \quad 0 \leq u \leq 1.$$  

Thus, the Lorenz curve $L(u)$ shows the share of total income received by the 100u per poorest of the population. By introducing the conditional mean function $H(\cdot)$ defined by

$$H(u) = E\left(X \mid X \leq F^{-1}(u)\right) = \frac{1}{F(u)} \int_0^u F^{-1}(t) \, dt, \quad 0 \leq u \leq 1,$$

Aaberge (1982) found that the Lorenz curve can be written on the following form

$$L(u) = \frac{H(u)}{H(1)} \quad 0 \leq u \leq 1.$$  

Let $X_1, X_2, \ldots, X_n$ be independent random variables with common distribution function $F$ and let $F_n$ be the corresponding empirical distribution function. Since the parametric form of $F$ is not known, it is natural to use the empirical distribution function $F_n$ to estimate $F$ and to use
(2.4) \[ H_u(u) = \frac{1}{u} \int_0^u F_n^{-1}(t) \, dt, \quad 0 \leq u \leq 1 \]

to estimate \( H(u) \), where \( F_n^{-1} \) is the left inverse of \( F_n \). Now replacing \( H(u) \) by \( H_u(u) \) in the expression (2.3) for \( L(u) \), we get the empirical Lorenz curve

(2.5) \[ L_u(u) = u \frac{H_u(u)}{H_u(1)}, \quad 0 \leq u \leq 1. \]

To obtain an explicit expression for \( H_u(u) \) and the empirical Lorenz curve, let \( X_{(i)} \leq X_{(2)} \leq \ldots \leq X_{(n)} \) denote the ordered \( X_1, X_2, \ldots, X_n \). For \( u = i/n \) we have

(2.6) \[ H_u\left( \frac{i}{n} \right) = \frac{1}{n} \sum_{j=1}^{i} X_{(j)}, \quad i = 1, 2, \ldots, n \]

and

(2.7) \[ L_u\left( \frac{i}{n} \right) = \frac{\sum_{j=i}^{n} X_{(j)}}{\sum_{j=1}^{n} X_j}, \quad i = 1, 2, \ldots, n \]

which is the familiar estimate formula of the empirical Lorenz curve.

As mentioned in Section 1 the ranking of Lorenz curves becomes problematic when the Lorenz curves in question intersect. For this reason and to be able to quantify the inequality in distributions of income it is common to apply summary measures of inequality. As justified in Section 1 it appears attractive to consider the family of rank-dependent measures of inequality introduced by Mehran (1976) and defined by

(2.8) \[ J_R(L) = 1 - \int_0^1 R(u) L(u) \, du \]

where \( R \) is a non-negative weight-function\(^3\).

By inserting for the following two alternative subclasses \( R_1 \) and \( R_2 \) of \( R \),

\(^3\) A slightly different version of \( J_R \) was introduced by Piesch (1975), whereas Giaccardi (1950) considered a discrete version of \( J_R \). For alternative normative motivations of the \( J_R \)-family and various subfamilies of the \( J_R \)-family we refer to Donaldson and Weymark (1981), Yaari (1987, 1988), Ben Porath and Gilboa (1984) and Aaberge (2001).
\[ R_{1k}(u) = k(k+1)(1-u)^{k-1}, \quad k > 0 \]

and

\[ R_{2k}(u) = (k+1)u^{k-1}, \quad k > 0 \]

we get the following subfamilies of the general rank-dependent family of inequality measures \( J_R \).

\[ G_k \equiv J_{\mathcal{G}_k}(L) = 1 - k(k+1) \int_0^1 (1-u)^{k-1} L(u) du, \quad k > 0 \]

and

\[ D_k \equiv J_{\mathcal{D}_k}(L) = 1 - (k+1) \int_0^1 u^{k-1} L(u) du, \quad k > 0. \]

Note that \( \{G_k : k > 0\} \) was denoted the extended Gini family and \( \{D_k : k > 0\} \) the “illfare-ranked single series Ginis” by Donaldson and Weymark (1980). However, as mentioned in Section 1 Aaberge (2000a) proved that each of the subfamilies \( \{D_k : k = 1, 2, \ldots\} \) (denoted the Lorenz family of inequality measures) and \( \{G_k : k = 1, 2, \ldots\} \) provides a complete characterization of the Lorenz curve, independent of whether the distribution function \( F \) is defined on a bounded interval or not. Thus, any distribution function \( F \) defined on \( \mathbb{R}^+ \) can be specified by its mean and Lorenz measures of inequality even if some of the conventional moments do not exist.

It follows directly from expressions (2.11) and (2.12) that the Gini coefficient defined by

\[ G = 1 - 2 \int_0^1 L(u) du \]

is included in the extended Gini family as well as in the Lorenz family of inequality measures.

By replacing \( L \) by \( L_n \) in the expression (2.8) for \( J_R \), we get the following estimator of \( J_R \),

\[ \hat{J}_R = J_R(L_n) = 1 - \int_0^1 R(u)L_n(u) du. \]

For \( R(u) = 2, \) (2.14) gives the estimator of \( G, \)
Asymptotic distribution theory of the empirical Lorenz curve and empirical rank-dependent measures of inequality

As demonstrated by expressions (2.8) and (2.14), the rank-dependent measures of inequality and their empirical counterparts are explicitly defined in terms of the Lorenz curve and its empirical counterpart, respectively. Thus, in order to derive the asymptotic distribution of the empirical rank-dependent measures of inequality it is convenient to firstly derive the asymptotic properties of the empirical Lorenz curve. To this end we utilize the close formal connection between the shift function of Doksum (1974) and the Lorenz curve.

3.1. Asymptotic properties of the empirical Lorenz curve

Since \( F_n \) is a consistent estimate of \( F \), \( H_n(u) \) and \( L_n(u) \) are consistent estimates of \( H(u) \) and \( L(u) \), respectively.

Approximations to the variance of \( L_n \) and the asymptotic properties of \( L_n \) can be obtained by considering the limiting distribution of the process \( Z_n(u) \) defined by

\[
Z_n(u) = n^2 \left[ L_n(u) - L(u) \right].
\]

In order to study the asymptotic behavior of \( Z_n(u) \) we find it useful to start with the process \( Y_n(u) \) defined by

\[
Y_n(u) = n^2 \left[ H_n(u) - H(u) \right] = \int_0^u n^2 \left( F_n^{-1}(t) - F^{-1}(t) \right) dt.
\]

Assume that the support of \( F \) is a non-empty finite interval \([a,b]\). (When \( F \) is an income distribution, \( a \) is commonly equal to zero.) Then \( Y_n(u) \) and \( Z_n(u) \) are members of the space \( D \) of functions on \([0,1]\) which are right continuous and have left hand limits. On this space we use the

\[
(2.15) \quad \hat{G} = 1 - 2 \int_0^1 L_n(u) du = 1 - \frac{\sum_{j=1}^n \sum_{j'=1}^i X_{(j)}}{(n+1) \sum_{j=1}^n X_j}.
\]
Skorokhod topology and the associated σ-field (e.g. Billingsley (1968), page 111). We let $W_0(t)$ denote a Brownian Bridge on $[0,1]$, that is, a Gaussian process with mean zero and covariance function $s(1-t)$, $0 \leq s \leq t \leq 1$.

**THEOREM 1:** Suppose that $F$ has a continuous nonzero derivate $f$ on $[a,b]$. Then $Y_n(u)$ converges in distribution to the process

$$Y(u) = \int_0^u \frac{W_0(t)}{f(F^{-1}(t))} \, dt.$$  \hspace{1cm} (3.3)

**PROOF:** It follows directly from Theorem 4.1 of Doksum (1974) that

$$n^2 \left( F_n^{-1}(t) - F^{-1}(t) \right)$$

converges in distribution to the Gaussian process $W_0(t)/f(F^{-1}(t))$.

Using the arguments of Durbin (1973, section 4.4), we find that $Y(u)$ as a function of $\left( W_0(t)/f(F^{-1}(t)) \right)$ is continuous in the Skorokhod topology. The result then follows from Billingsley (1968, Theorem 5.1).

Q.E.D.

The following result states that $Y(u)$ is a Gaussian process and thus that $Y_n(u)$ is asymptotically normally distributed, both when considered as a process, and for fixed $u$.

**THEOREM 2:** Suppose the conditions of Theorem 1 are satisfied. Then the process $uY(u)$ has the same probability distribution as the Gaussian process

$$\sum_{j=1}^n q_j(u) Z_j$$

where $q_j(u)$ is given by
(3.4) \[ q_j(u) = \frac{2^{j-1}}{j\pi} \int_0^u \frac{\sin(j\pi t)}{f(F^{-1}(t))} dt \]

and \( Z_1, Z_2, \ldots \) are independent \( N(0,1) \) variables.

PROOF: Put

\[ V_N(t) = \frac{2^{1/2}}{f(F^{-1}(t))} \sum_{j=1}^N \frac{\sin(j\pi t)}{j\pi} Z_j \]

and note that

\[ 2 \sum_{j=1}^N \frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2} = s(1-t), \quad 0 \leq s \leq t \leq 1. \]  

(3.5)

Thus, the process \( V_N(t) \) is Gaussian with mean zero and covariance function

\[ \text{cov}(V_N(s), V_N(t)) = \frac{2}{f(F^{-1}(s))f(F^{-1}(t))} \sum_{j=1}^N \frac{\sin(j\pi s)\sin(j\pi t)}{(j\pi)^2} \rightarrow \text{cov}(V(s), V(t)), \]

where

\[ V(t) = \frac{W(t)}{f(F^{-1}(t))}. \]

In order to prove that \( V_N(t) \) converges in distribution to the Gaussian process \( V(t) \), it is, according to Hajek and Sidak (1967, Theorem 3.1.a, Theorem 3.1.b, Theorem 3.2) enough to show that

\[ E[V_N(t) - V_N(s)]^4 \leq M(t-s)^2, \quad 0 \leq s, t \leq 1, \]

where \( M \) is independent of \( N \).

Since for normally distributed random variables with mean 0,

\[ EX^4 = 3[EX^2]^2, \]

we have
\[
E[V_N(t) - V_N(s)]^4 = 3\left[\text{var}(V_N(t) - V_N(s))\right]^2
\]

\[
= 3\left\{2\text{var}\left[\sum_{j=1}^{N} \frac{1}{j\pi} \left(\frac{\sin(j\pi t)}{f(F^{-1}(t))} - \frac{\sin(j\pi s)}{f(F^{-1}(s))}\right)Z_j\right]\right\}^2
\]

\[
= 3\left\{2\sum_{j=1}^{N} \left[\frac{1}{j\pi} \left(\frac{\sin(j\pi t)}{f(F^{-1}(t))} - \frac{\sin(j\pi s)}{f(F^{-1}(s))}\right)\right]^2\right\}^2 \leq 3\left\{2\sum_{j=1}^{N} \left[\frac{1}{j\pi} \left(\frac{\sin(j\pi t)}{f(F^{-1}(t))} - \frac{\sin(j\pi s)}{f(F^{-1}(s))}\right)\right]^2\right\}^2
\]

\[
= 3\left[\frac{t(1-t)}{f^2(F^{-1}(t))} + \frac{s(1-s)}{f^2(F^{-1}(s))} - 2\text{cov}(W_N(s),W_N(t))\right]^2.
\]

Since \(0 < f(x) < \infty\) on \([a,b]\), there exists a constant \(M\) such that

\[
f(F^{-1}(t)) \geq M\frac{1}{t^4} \text{ for all } t \in [0,1].
\]

Then

\[
E[V_N(t) - V_N(s)]^2 \leq 3M(t - s)^2(1 - |t - s|)^2 \leq 3M(t - s)^2.
\]

Hence \(V_N(t)\) converges in distribution to the process \(V(t)\). Thus, according to Billingsley (1968, Theorem 5.1)

\[
\int_0^u V_N(t)dt = \sum_{j=1}^{N} q_j(u)Z_j
\]

converges in distribution to the process

\[
\int_0^u V(t)dt = \int_0^u \frac{W_N(t)}{f(F^{-1}(t))}dt = uY(u).
\]

Q.E.D.

Now, let \(h_j\) be a function defined by

\[
h_j(u) = \frac{1}{\mu} \left[ q_j(u) - q_j(1)L(u) \right]
\]

where \(q_j(u)\) is given by (3.4).
THEOREM 3: Suppose the conditions of Theorem 1 are satisfied. Then $Z_n(u)$ given by (3.1) converges in distribution to the Gaussian process

\[(3.7) \quad Z(u) = \sum_{j=1}^{\infty} h_j(u)Z_j \]

where $Z_1, Z_2, \ldots$ are independent $N(0,1)$ variables and $h_j(u)$ is given by (3.6).

PROOF: By combining (2.5), (3.1) and (3.2) we see that

\[
Z_n(u) = \frac{1}{H_n(1)} \left[ uY_n(u) - L(u)Y_n(1) \right]
\]

where $Y_n(u)$ is given by (3.2).

Now, Theorem 1 implies that the process

\[uY_n(u) - L(u)Y_n(1)\]

converges in distribution to the process

\[uY(u) - L(u)Y(1)\]

where $Y(u)$ is given by (3.3). Then, since $H_n(1)$ converges in probability to $\mu$, Cramer-Slutsky’s theorem gives that $Z_n(u)$ converges in distribution to the process

\[
\frac{1}{\mu} \left[ uY(u) - L(u)Y(1) \right].
\]

Thus, by applying Theorem 2 the proof is completed.

Q.E.D.

In order to derive the asymptotic covariance functions of the processes $Y_n(u)$ and $Z_n(u)$, the following lemma is needed.

LEMMA 1: Suppose the conditions of Theorem 1 are satisfied. Then

\[(3.8) \quad \sum_{\nu=1}^{\infty} q_{\nu}(u)q_{\nu}(v) = \tau^2(u) + \lambda(u,v), \quad 0 \leq u, v \leq 1, \]
where $q_i(u)$ is defined by (3.4) and $\tau^2(u)$ and $\lambda(u,v)$ are given by

\begin{equation}
\tau^2(u) = 2 \int_a^{F^{-1}(u)} \int_a^{y} F(x)\{1-F(y)\}dx\,dy, \quad 0 \leq u \leq 1
\end{equation}

and

\begin{equation}
\lambda(u,v) = \int_{F^{-1}(u)}^{F^{-1}(v)} \int_a^{y} F(x)\{1-F(y)\}dx\,dy, \quad 0 \leq u \leq v \leq 1.
\end{equation}

**PROOF:** Assume that $0 \leq u \leq v \leq 1$. From the definition of $q_i(u)$ we have that

\[ \sum_{i=1}^{w} q_i(u)q_i(v) = \sum_{i=1}^{w} \int_0^{u} \int_0^{v} \left[ \frac{2}{f(F^{-1}(t))f(F^{-1}(s))} \frac{\sin(i\pi t)\sin(i\pi s)}{(i\pi)^2} \right] dt\,ds. \]

By applying Fubini’s theorem (e.g. Royden (1963)) and the identity (3.5) we get

\[ \sum_{i=1}^{w} q_i(u)q_i(v) = \int_0^{u} \int_0^{v} \left[ \frac{2}{f(F^{-1}(t))f(F^{-1}(s))} \sum_{i=1}^{w} \sin(i\pi t)\sin(i\pi s) \right] dt\,ds = 2 \int_a^{F^{-1}(u)} \int_a^{y} (F(x)\{1-F(y)\}dx\,dy + \int_{F^{-1}(u)}^{F^{-1}(v)} \int_a^{y} F(x)\{1-F(y)\}dx\,dy = \tau^2(u) + \lambda(u,v). \]

Q.E.D.

As an immediate consequence of Theorem 1, Theorem 2 and Lemma 1 we have the following corollary.

**COROLLARY 1:** Under the conditions of Theorem 1, $Y_i(u)$ has asymptotic covariance function $\theta^2(u,v)$ given by

\begin{equation}
\theta^2(u,v) = \frac{1}{uv}[\tau^2(u) + \lambda(u,v)], \quad 0 < u \leq v \leq 1.
\end{equation}
From Theorem 3 and Lemma 1 we get the next corollary.

**COROLLARY 2**: Under the conditions of Theorem 1, $Z_n(u)$ has asymptotic covariance function $\nu^2(u,v)$ given by

$$
\nu^2(u,v) = \frac{1}{\mu^2} \left[ \tau^2(u) + \lambda(u,v) - L(u)\left(\tau^2(v) + \lambda(v,1)\right) - L(v)\left(\tau^2(u) + \lambda(u,1)\right) + L(u)L(v)\tau^2(1) \right], \quad 0 < u \leq v \leq 1.
$$

(3.12)

In order to construct confidence intervals for the Lorenz curve at fixed points, we apply the results of Theorem 3 and Corollary 2 which imply that the distribution of

$$\frac{\frac{1}{n^2} L_n(u) - L(u)}{\nu(u,u)}$$

tends to the $N(0,1)$ distribution for fixed $u$, where $\nu^2(u,u)$ is given by

$$
\nu^2(u,u) = \frac{1}{\mu^2} \left[ \tau^2(u) - 2L(u)\left(\tau^2(u) + \lambda(u,1)\right) + L(u)L(1)\tau^2(1) \right], \quad 0 < u \leq 1.
$$

(3.13)

Before this result can be applied, we must estimate the asymptotic variance $\nu^2(u,u)$, i.e., we must estimate $\mu$, $L$, $\tau^2$ and $\lambda$. The estimates of $\mu$ and $L$ are given by $\bar{X}$ and (2.7), respectively. Now, by introducing the statistics $a_k$ and $b_k$ defined by

$$
a_k = \left(1 - \frac{k}{n}\right)(X_{(\hat{k})} - X_{(\bar{k})})
$$

(3.14)

and

$$
b_k = \frac{k}{n}(X_{(\hat{k})} - X_{(\bar{k})}),
$$

(3.15)

we obtain the following consistent estimates of $\tau^2$ and $\lambda$,

$$
\hat{\tau}^2\left(\frac{i}{n}\right) = 2\sum_{i=1}^{[\frac{i}{n}]} a_i \sum_{i=1}^{\frac{i}{n}} b_i, \quad i = 2,3,...,n
$$

(3.16)

and
Thus, replacing $\mu$, $L$, $\hat{\lambda}$ and $\lambda$ by their respective estimates in the expression (3.13) for $\nu^2$ we obtain a consistent estimate of $\nu^2$.

To get an idea of how reliable $L_n(u)$ is as an estimate for $L(u)$, we have to construct a confidence band based on $L_n(u)$ and $L(u)$. Such a confidence band can be obtained from statistics of the type

$$K_n = n^2 \sup_{0 < u < 1} \frac{|L_n(u) - L(u)|}{\psi(L_n(u))}$$

where $\psi$ is a continuous nonnegative weight function. By applying Theorem 3 and Billingsley (1968, Theorem 5.1), we find that $K_n$ converges in distribution to

$$K = \sup_{0 \leq u < 1} \left| \sum_{j=1}^{u} \frac{h_j(u)}{\psi(L(u))} Z_j \right|.$$

Let

$$T_n(u) = \sum_{j=1}^{u} \frac{h_j(u)}{\psi(L(u))} Z_j,$$

$$T(u) = \sum_{j=1}^{u} \frac{h_j(u)}{\psi(L(u))} Z_j$$

and

$$K_n' = \sup_{0 \leq u < 1} |T_n(u)|.$$

Since $T_n$ converges in distribution to $T$, we find by applying Billingsley (1968, Theorem 5.1) that $K_n'$ converges in distribution to $K$. Hence, for a suitable choice of $m$ and $\psi$, for instance $\psi = 1$, simulation methods may be used to obtain the distribution of $K_n'$ and thus an approximation for the distribution of $K$. 
3.2. Asymptotic properties of the empirical rank-dependent family of inequality measures

We shall now study the asymptotic distribution of the statistics \( \hat{J}_R \) given by (2.14). Mehran (1976) states without proof that \( n^{\frac{1}{2}} \left( \hat{J}_R - J_R \right) \) is asymptotically normally distributed with mean zero. The asymptotic variance, however, cannot be derived, as maintained by Mehran (1976), from Stigler (1974, Theorem 1). However, as will be demonstrated below Theorem 3 forms a helpful basis for deriving the asymptotic variance of \( \hat{J}_R \).

Let \( \omega^2 \) be a parameter defined by

\[
\omega^2 = \frac{1}{\mu^2} \left\{ 2 \int_0^1 \int_0^1 \left[ \tau^2(v) + \lambda(u,v) \right] R(u)R(v)\,du\,dv \right. \\
\left. - 2 \left[ \int_0^1 uR(u)\,du - J_R \right] \left[ \int_0^1 \left( \tau^2(u) + \lambda(u,1) \right) R(u)\,du \right] + \tau^2(1) \left[ \int_0^1 uR(u)\,du - J_R \right] \right\}.
\]

(3.23)

**THEOREM 4:** Suppose the conditions of Theorem 1 are satisfied and \( \omega^2 < \infty \). Then the distribution of \( n^{\frac{1}{2}} \left( \hat{J}_R - J_R \right) \) tends to the normal distribution with zero mean and variance \( \omega^2 \).

**PROOF:** From (2.8), (2.14) and (3.1) we see that

\[
\int_0^1 \int_0^1 \left( \tau^2(u) + \lambda(u,v) \right) R(u)R(v)\,du\,dv.
\]

By Theorem 3 we have that \( Z(u) \) converges in distribution to the Gaussian process \( Z(u) \) defined by (3.7). By applying Billingsley (1968, Theorem 5.1) and Fubini’s theorem we get that \( n^{\frac{1}{2}} \left( \hat{J}_R - J_R \right) \) converges in distribution to

\[
- \int_0^1 R(u)Z(u)\,du = - \int_0^1 R(u) \left( \sum_{j=1}^J h_j(u)Z_j \right) \,du = - \sum_{j=1}^J \left[ \int_0^1 R(u)h_j(u)\,du \right] Z_j.
\]
where $Z_1, Z_2, \ldots$ are independent $N(0,1)$ variables and $h_j(u)$ is given by (3.6), i.e., the asymptotic distribution of $n^{\frac{1}{2}} \left( J_n - J_h \right)$ is normal with mean zero and variance

$$
(3.24) \quad \sum_{j=1}^{\infty} \left[ \int_0^1 R(u)h_j(u)du \right]^2.
$$

Then it remains to show that the asymptotic variance is equal to $\omega^2$.

Inserting (3.6) in (3.24), we get

$$
\sum_{j=1}^{\infty} \left[ \int_0^1 R(u)h_j(u)du \right]^2 = \frac{1}{\mu^2} \sum_{j=1}^{\infty} \left[ \int_0^1 R(u) \left( q_j(u) - q_j(1)L(u) \right) du \right]^2
$$

$$
= \frac{1}{\mu^2} \left[ \sum_{j=1}^{\infty} \left[ \int_0^1 R(u)q_j(u)du \right]^2 - 2 \left[ \int_0^1 R(u)L(u)du \right] \left[ \sum_{j=1}^{\infty} q_j(1) \int_0^1 R(u)q_j(u)du \right] \right]
$$

$$
+ \left[ \sum_{j=1}^{\infty} q_j^2(1) \int_0^1 R(u)L(u)du \right]^2.
$$

In the following derivation we apply Fubini’s theorem and the identity (3.5),

$$
\sum_{j=1}^{\infty} \left[ \int_0^1 R(u)q_j(u)du \right]^2 = \sum_{j=1}^{\infty} \int_0^1 \int_0^1 R(u)q_j(u)R(v)q_j(v)du dv
$$

$$
= \int_0^1 \int_0^1 \int_0^1 \int_0^1 \frac{2}{f(F^{-1}(t))f(F^{-1}(s))} \left( \sum_{j=1}^{\infty} \sin \left( j\pi t \right) \sin \left( j\pi s \right) \right) dt ds R(u)R(v)du dv
$$

$$
= 2 \int_0^1 \int_0^1 \int_0^1 \int_0^1 \frac{t(1-s)}{f(F^{-1}(t))f(F^{-1}(s))} dt ds R(u)R(v)du dv + \int_0^1 \int_0^1 \frac{t(1-s)}{f(F^{-1}(t))f(F^{-1}(s))} dt ds R(u)R(v)du dv
$$

$$
= 2 \int_0^1 \int_0^1 \left[ F^{-1}(u) \int_a^u F(x)(1-F(y))dx dy + F^{-1}(v) \int_a^v F(x)(1-F(y))dx dy \right] R(u)R(v)du dv
$$

$$
= 2 \int_0^1 \left[ \tau^2(u) + \lambda(u,v) \right] R(u)R(v)du dv
$$

where $\tau^2(u)$ and $\lambda(u,v)$ are given by (3.9) and (3.10), respectively. Similarly, we find that

$$
\sum_{j=1}^{\infty} q_j(1) \int_0^1 R(u)q_j(u)du = \int_0^1 \left[ \tau^2(u) + \lambda(u,1) \right] R(u)du.
$$
From Lemma 1 it follows that

\[ \sum_{j=1}^{\infty} q_j(l) = \tau^2(l). \]

Finally, by noting that

\[ \int_{0}^{1} R(u)L(u)du = \int_{0}^{1} uR(u)du - I, \]

the proof is completed. Q.E.D.

For \( R(u) = 2 \), Theorem 4 states that \( \omega^2 = \gamma^2 \), where \( \gamma^2 \) is defined by

\[
\gamma^2 = \frac{4}{\mu^2} \left[ 2 \int_{0}^{1} \left[ \tau^2(u) + \lambda(u,v) \right] du dv - (1-G) \int_{0}^{1} \left[ \tau^2(u) + \lambda(u,1) \right] du + \frac{1}{4}(1-G)\tau^2(l) \right],
\]

is the asymptotic variance of \( n^2 \hat{G} \).

An alternative version of (3.25) is given by Hoeffding (1948).

The estimation of \( \gamma^2 \) is straightforward. As in Section 2 we assume that the parametric form of \( F \) is not known. Thus, replacing \( F \) by the empirical distribution function \( F_n \) in expression (4.1) for \( \gamma^2 \), we obtain a consistent nonparametric estimator for \( \gamma^2 \). The current estimator is given by

\[
\hat{\gamma}^2 = \frac{4}{X^2} \left[ \frac{2}{n^2} \sum_{i=1}^{n} \sum_{j=2}^{n} \tilde{\tau}^2 \left( \frac{i}{n}, \frac{j}{n} \right) + \frac{2}{n^2} \sum_{i=1}^{n} \sum_{j=2}^{n} \hat{\lambda} \left( \frac{i}{n}, \frac{j}{n} \right) \right]
\]

\[
- \frac{1}{n^2} (1-\hat{G}) \left[ \sum_{i=2}^{n} \tilde{\tau}^2 \left( \frac{i}{n}, \frac{1}{n} \right) + \sum_{i=2}^{n} \hat{\lambda} \left( \frac{i}{n}, \frac{1}{n} \right) \right] + \frac{1}{4}(1-\hat{G})\hat{\tau}^2(l),
\]

where \( \tilde{\tau}^2 \), \( \hat{\lambda} \) and \( \hat{G} \) are given by (3.16), (3.17) and (2.15), respectively.

Similarly, a consistent estimator for \( \omega^2 \) is obtained by replacing \( \hat{\tau}, \hat{\lambda}, \mu \) and \( I \) by their respective estimates in the expression (3.23) for \( \omega^2 \).
References


Recent publications in the series Discussion Papers

313  G.H. Bjertnuer (2001): Optimal Combinations of Income Tax and Subsidies for Education
316  A. Bruvoll and K. Nyborg (2002): On the value of households' recycling efforts
317  E. Biorn and T. Skjerpen (2002): Aggregation and Aggregation Biases in Production Functions: A Panel Data Analysis of Translog Models
323  E. Roed Larsen (2002): Searching for Basic Consumption Patterns: Is the Engel Elasticity of Housing Unity?
325  E. Roed Larsen (2002): Consumption Inequality in Norway in the 80s and 90s.
332  M. Greaker (2002): Eco-labels, Production Related Externalities and Trade
333  J. T. Lind (2002): Small continuous surveys and the Kalman filter
334  B. Halvorsen and T. Willumsen (2002): Willingness to Pay for Dental Fear Treatment. Is Supplying Fear Treatment Social Beneficial?
335  T. O. Thoresen (2002): Reduced Tax Progressivity in Norway in the Nineties. The Effect from Tax Changes
340  H. C. Bjørnland and H. Hungnes (2003): The importance of interest rates for forecasting the exchange rate
343  B. Bye, B. Strom and T. Ávitsland (2003): Welfare effects of VAT reforms: A general equilibrium analysis
346  B.M. Larsen and R.Nesbakken (2003): How to quantify household electricity end-use consumption
347  B. Halvorsen, B. M. Larsen and R. Nesbakken (2003): Possibility for hedging from price increases in residential energy demand
349  B. Holtsmark (2003): The Kyoto Protocol without USA and Australia - with the Russian Federation as a strategic permit seller
352  E. Holmøy (2003): Aggregate Industry Behaviour in a Monopolistic Competition Model with Heterogeneous Firms
357  A. Bruvoll T. Bye, J. Larsson og K. Telle (2003): Technological changes in the pulp and paper industry
and the role of uniform versus selective environmental policy.


T. Skjerpen (2004): The dynamic factor model revisited: The identification problem remains


M. Greaker (2004): Industrial Competitiveness and Diffusion of New Pollution Abatement Technology – a new look at the Porter-hypothesis


K. Telle and J. Larssøn (2004): Do environmental regulations hamper productivity growth? How accounting for improvements of firms’ environmental performance can change the conclusion


E. Lund Sagen and F. R. Aune (2004): The Future European Natural Gas Market - are lower gas prices attainable?

A. Langørgen and D. Rønningen (2004): Local government preferences, individual needs, and the allocation of social assistance

K. Telle (2004): Effects of inspections on plants’ regulatory and environmental performance - evidence from Norwegian manufacturing industries

T. A. Galloway (2004): To What Extent Is a Transition into Employment Associated with an Exit from Poverty

J. F. Bjørnstad and E. Yetterstad (2004): Two-Stage Sampling from a Prediction Point of View

A. Bruvoll and T. Fæhn (2004): Transboundary environmental policy effects: Markets and emission leakages


N. Keilman and D. Q. Pham (2004): Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area

G. H. Bjørnnes, D. Rime and H. O. A. Solheim (2004): Liquidity provision in the overnight foreign exchange market


B. Halvorsen (2004): Effects of norms, warm-glow and time use on household recycling


T. Åvitsland and J. Aasness (2004): Combining CGE and microsimulation models: Effects on equality of VAT reforms

M. Greaker and Eirik. Sagen (2004): Explaining experience curves for LNG liquefaction costs: Competition matter more than learning

K. Telle, I. Aslaksen and T. Synnestvedt (2004): ”It pays to be green” - a premature conclusion?


T. Hægeland, O. Rauum and K.G. Salvanes (2004): Pupil achievement, school resources and family background


B. Halvorsen and Runa Neshabaken (2004): Accounting for differences in choice opportunities in analyses of energy expenditure data
