USA Income Distribution Counter-Business-Cyclical Trend

(Estimating Lorenz curve using Continuous L_1 norm estimation)

Bijan Bidabad[1]

Abstract

In this paper, the L_1 norm of continuous functions and corresponding continuous estimation of regression parameters are defined. The continuous L_1 norm estimation problems of linear one and two parameters models are solved. We proceed to use the functional form and parameters of probability distribution function of income to exactly determine the L_1 norm approximation of the corresponding Lorenz curve of the statistical population under consideration. U.S. economic data used to estimate income distribution for the period of 1977-2002. An interesting finding of these calculations is that the distribution of income obeys a counter wise business cycles fluctuations. This finding is a new area for research in realm of the theory and application of income distribution and business cycles interrelationship.

1. Introduction

The skewness of income distribution is persistently exhibited for different populations and in different times. It is discussed that Pearsonian family distributions are rival functions to explain income distribution. Lorenz curve is a method to analyze the skew distributions. There is a relation between the area under the Lorenz curve and the corresponding probability distribution function of the statistical population (see, Kendall and Stuart (1977)). That is, when the probability distribution function is known, we may find the corresponding Gini coefficient as the measure of inequality.

Estimation of the Lorenz curve is confronted with some difficulties. For this estimation, we should define an appropriate functional form which can accept different curvatures (see, Bidabad and Bidabad (1989a,b)). There is another problem, that is, to create the necessary data set for estimating the corresponding parameters of the Lorenz curve, a large amount of computation on raw sample income data is inevitable. Obviously, these problems despite of their computational difficulties, make the significance of the estimated parameters poor (see,

[1]Bijan Bidabad, Professor of Economics & Director of Foreign Exchange Research Dept., Monetary & Banking Research Academy, Centeral Bank of Iran, No.189, Pasdaran Ave., Zarabkhaneh, Tehran 16619, Iran. Tel.: mobile: +98.911.2090164, office: +98.21.2855607-8; fax: +98.21.2850457, Email: bijan_bidabad@msn.com

Bidabad and Bidabad (1989a,b)). To avoid this, we try to estimate the functional form of the Lorenz curve by using continuous information. In this paper we use the probability density function of population income to estimate the Lorenz function parameters. The continuous L_1 norm smoothing method which will be developed for estimating the regression parameters is used to solve this problem. However, we concentrate on two rival probability density functions of Pareto and log-normal. Since, the former is simply integrable, there is no general problem to derive the corresponding Lorenz function and the function is uniquely derived. But in the latter case, the log-normal density function (which has better performance for full income range) than Pareto distribution (which better fits to higher income range, (see, Cramer (1973), Singh and Maddala (1976), Salem and Mount (1974)), is not integrable and we can not determine its corresponding Lorenz function. In this regard we should solve the problem by defining a general Lorenz curve functional form and applying the L_1 norm smoothing to estimate the corresponding parameters.

In this paper continuous L1 norm estimation is developed by using a similar method proposed in Bidabad (1987a,88a,89a,b) for discrete case. Then the method is applied to estimation of the Lorenz curve functional forms which have been proposed by Gupta (1984) and Bidabad and Bidabad (1989,92). At the end, we use our formulation to estimate Gini ratio and Kakwani length indices of inequality for the United States for the period of 1971-1990, based on the assumption that income is distributed log-normally.

2. L_1 norm of continuous functions

Generally, L_p norm of a function f(x) (see, Rice and White (1964)) is defined by,

$$||f(x)||_p = \int_{x\varepsilon I} |f(x)|^p dx)^{1/p} \tag{1}$$

Where, "I" is a closed bounded set. The L_1 norm of f(x) is simply written as,

$$||f(x)||_1 = \int_{x\varepsilon I} |(x)|dx \tag{2}$$

Suppose that the non-stochastic function $f(x,\boldsymbol{\beta})$ of "x", is combined with stochastic disturbance term "u" to form y(x) as follows,

$$y(x) = f(x, \boldsymbol{\beta}) + u \tag{3}$$

Where, $\boldsymbol{\beta}$ is unknown parameters vector. Rewriting u as the residual of $y(x)-f(x,\boldsymbol{\beta})$, for L_1 norm approximation of "$\boldsymbol{\beta}$" we should find "$\boldsymbol{\beta}$" vector such that the L_1 norm of "u" is minimum. That is,

$$\underset{\boldsymbol{\beta}}{\text{Min:}}\ S=||u||_1=||y(x)-f(x,\boldsymbol{\beta})||_1=\int_{x\varepsilon I} |y(x)-f(x,\boldsymbol{\beta})|dx \tag{4}$$

3. Linear one parameter L_1 norm continuous smoothing

Redefine $f(x,\boldsymbol{\beta})$ as βx and y(x) as the following linear function,

$$y(x) = \beta x + u \tag{5}$$

Where, "$\boldsymbol{\beta}$" is a single (non-vector) parameter. Expression (4) reduces to:

$$\underset{\beta}{\text{min:}}\ S = ||u||_1 = ||y(x)- \beta x||_1 = \int_{x\varepsilon I} |y(x)-f(x,\beta)|dx \tag{6}$$

The discrete analogue of (6) is solved by Bidabad (1987a,88a,89a,b). In these papers we proposed applying discrete and regular derivatives to the discrete problem by using a slack variable "t" as a point to distinguish negative and positive residuals. A similar approach is used here to minimize (6). To do so in this case certain Lipschitz conditions are imposed on the functions involved (see, Usow (1967a)). Rewrite (6) as follows,

$$\underset{\beta}{\text{Min:}}\ S = \int_{x\varepsilon I} |x||y(x)/x - \beta|dx \tag{7}$$

For convenience, define "I" as a closed interval [0,1]. The procedure may be applied to other intervals with no major problem (see, Usow (1967a), Hobby and Rice (1965), Kripke and Rivlin (1965)). To minimize this function we should first remove the absolute value sign of the expression after the integral sign. Since "x" belongs to closed interval "I", y(x) (which is a linear function of "x") and also y(x)/x are smooth and continuous. Thus, since y(x)/x is uniformly increasing or decreasing function of "x", a value of tnI can be found to have the following properties,

$$\begin{aligned} y(x)/x &< \beta \quad \text{if } x < t \\ y(x)/x &= \beta \quad \text{if } x = t \\ y(x)/x &> \beta \quad \text{if } x > t \end{aligned} \tag{8}$$

Value of the slack variable "t" actually is the border of negative and positive residuals. If value of "t" were known, from (8) (middle equation) we could calculate optimal value of "β" or inversely. But nor "t" neither "β" are known. To solve this problem, according to (8), we can rewrite (7) as two separate definite integrals with different upper and lower bounds.

$$\min_{\beta}: S = -\int_0^t |x|\,(y(x)/x - \beta)dx + \int_t^1 |x|\,(y(x)/x - \beta)dx \tag{9}$$

Decomposition of (7) into (8) has been done by use of the slack variable "t". Since both "β" and "t" are unknown, to solve (9), we partially differentiate it with respect to "t" and "β" variables.

$$\frac{\delta S}{\delta \beta} = \int_0^t |x|dx - \int_t^1 |x|dx = 0 \tag{10}$$

and using Liebniz' rule to differentiate the integrals with respect to their variable bounds "t", yields,

$$\frac{\delta S}{\delta t} = -|t|\,[\frac{y(t)}{t} - \beta] - |t|\,[\frac{y(t)}{t} - \beta] = 0 \tag{11}$$

Since "x" belongs to [0,1], equation (10) can be written as,

$$\int_0^t x dx - \int_t^1 x dx = 0 \tag{12}$$

or,

$$½\, t^2 - ½ + ½t^2 = 0 \tag{13}$$

Which yields,

$$t = \sqrt{2}/2 \tag{14}$$

Substitute for "t" in equation (11), yields,

$$\beta = \frac{y(\sqrt{2}/2)}{\sqrt{2}/2} \tag{15}$$

Remember that y(t) is function y(x) evaluated at x=t. Value of "β" given by (15) is the optimal solution of (6). The above procedure actually is generalization of Laplace weighted median for continuous case.

Before applying this procedure to Lorenz curve, let us develop the procedure for the two parameters linear model.

4. Linear two parameters L_1 norm continuous smoothing

Now, we try to apply the above technique to the linear two parameters model. Rewrite (4) as,

$$\min_{\alpha,\beta}: S=||u||_1=||y(x)-\alpha-\beta x||_1=\int_{x\varepsilon I} |y(x)-\alpha-\beta x|dx \tag{16}$$

Where, "α" and "β" are two single (non-vector) unknown parameters and y(x) and "x" are as before. According to Rice (1964c), let f(α*,β*,x) interpolates y(x) at the set of canonical points $\{x_i;i=1,2\}$, if y(x) is such that y(x)-f(α*,β*,x) changes sign at these x_i's and at no other points in [0,1], then $f(\alpha^*,\beta^*,x)$ is the best L_1 norm approximation to $y(x)$ (see also, Usow (1967a)). With the help of this rule, if we denote these two points to t_1 and t_2 we can rewrite (16) for I=[0,1] as,

$$S = \int_0^{t_1} [y(x)-\alpha-\beta x]dx - \int_{t_1}^{t_2} [y(x)-\alpha-\beta x]dx + \int_{t_2}^{1} [y(x)-\alpha-\beta x]dx \qquad (17)$$

Since t_1 and t_2 are also unknowns, we should minimize S with respect to α, β, t_1 and t_2. Taking partial derivative of (17) using Liebniz' rule with respect to these variables and equating them to zero, we will have,

$$\frac{\delta S}{\delta \alpha} = - \int_0^{t_1} dx + \int_{t_1}^{t_2} dx - \int_{t_2}^{t_1} dx = 0 \qquad (18)$$

$$\frac{\delta S}{\delta \beta} = - \int_0^{t_1} dx + \int_{t_1}^{t_2} dx - \int_{t_2}^{t_1} dx = 0 \qquad (19)$$

$$\frac{\delta S}{\delta t_1} = 2[y(t_1) -\alpha-\beta t_1] = 0 \qquad (20)$$

$$\frac{\delta S}{\delta t_2} = - 2[y(t_2) -\alpha - \beta t_2] = 0 \qquad (21)$$

Equations (18) through (21) may be solved simultaneously for α, β, t_1 and t_2. Thus, we have the following system of equations,

$$2t_2 - 2t_1 - 1 = 0 \qquad (22)$$
$$t_2^2 - t_1^2 - \tfrac{1}{2} = 0 \qquad (23)$$
$$y(t_1) - \alpha - \beta t_1 = 0 \qquad (24)$$
$$y(t_2) - \alpha - \beta t_2 = 0 \qquad (25)$$

The solutions are,

$$t_1=1/4 \qquad (26)$$
$$t_2=3/4 \qquad (27)$$
$$\alpha = y(3/4)-(3/4)\beta = y(1/4)-(1/4)\beta \qquad (28)$$
$$\beta = 2[y(3/4)-y(1/4)] \qquad (29)$$

This procedure, similar to that of multiple regression model for discrete case may be expanded to include "m" unknown parameters which is not discussed here. Some computational methods for solving the different cases of m parameters model are investigated by Ptak (1958), Rice and White (1964), Rice (1964a,b,c,69,85), Usow (1967a), Lazarski (1975a,b,c,77) (see also, Hobby and Rice (1965), Kripke and Rivlin (1965), Watson (1981)). Now, let us have a look at Lorenz curve and its proposed functional forms.

5. Lorenz curve

The Lorenz curve for a random variable with probability density function f(v) may be defined as the ordered pair[2],

$$\left(P(V|V\le v), \frac{E(V|V\le v)}{E(V)}\right) \qquad v \varepsilon R \qquad (30)$$

Where "P" and "E" stand for probability and expected value operators. For a continuous density function f(v), (30) can be written as,

[2] Taguchi (1972a,b,c,73,81,83,87,88) multiplies the second element of (30) by $P(V|V\le v)$ which is not correct; his definition of (31) is equivalent to ours.

$$\left(\int_{-\infty}^{v} f(w)dw, \frac{\int_{-\infty}^{v} wf(w)dw}{\int_{-\infty}^{+\infty l} wf(w)dw}\right) \equiv (x(v),y(x(v))) \qquad (31)$$

We denote (31) by (x(v),y(x(v))) where x(v) and y(x(v)) are its elements. Therefore, "x" is a function which maps "v" to x(v) and "y" is a function which maps x(v) to y(x(v)). The function y(x(v)) is simply the Lorenz curve function. In recent years some functional forms for Lorenz curve have been introduced. Among different proposed functions we use the forms of Gupta (1984) and Bidabad and Bidabad (1989,92) which benefits from certain properties (see their articles for more explanations). Gupta (1984) proposed the functional form,

$$y=xA^{x-1} \qquad A>1 \qquad (32)$$

Bidabad and Bidabad (1989,92) suggest the following functional form:

$$y=x^{B}A^{x-1} \qquad B\geq 1,\ A\geq 1 \qquad (33)$$

To estimate the above functions by regular estimating method, we should gather discrete data from the statistical population, and manipulate them to construct relevant x and y vectors to estimate "A" of (32) or "A" and "B" of (33). If the probability distribution of income is known, instead of gathering discrete observations, we can estimate the Lorenz curve by using the continuous L_1 norm smoothing method for continuous functions. In the following section we proceed to apply this method to estimate the parameters "A" of (32) and "A" and "B" of (33) by using the information of probability density function of income.

6. Continuous L_1 norm smoothing of Lorenz curve

To estimate the Lorenz curve parameters when income probability density function is known, we can not always take straightforward steps. When the probability density function is easily integrable, there is no major problem in advance. We can find the functional relationship between the two elements of (31) by simple mathematical derivation. But, when integrals of (31) are not obtainable, another procedure should be adopted.

Suppose that income of a society is distributed with probability density function f(w). This density function may be a skewed function such as Pareto or log-normal, as follows

$$f(w)=\theta k\theta w\text{-}\theta\text{-}1, \qquad wrk>0,\ \theta>0 \qquad (34)$$

$$f(w)=[1/w\sigma\sqrt{(2\pi)}]\exp\{-[\ln(w)-\mu]^2/2\sigma^2\}, \quad w\varepsilon(0,\infty),\ \mu\varepsilon(-\infty,+\infty),\ \sigma>0 \qquad (35)$$

These two distributions have been known as good candidates for presenting distribution of personal income.

In the case of Pareto density function of (34), we can simply derive the Lorenz curve function as follows. Let F(w) denote the Pareto distribution function:

$$F(w)=1-(k/w)\theta \qquad (36)$$

with mean equal to,

$$E(w)= \theta^{k}/(\theta-1),\ \ \theta>1 \qquad (37)$$

If we find the function y as stated by (31) as a function of x, the Lorenz function will be derived. Now, proceed as follows. Rearrange the terms of (31) as,

$$x(v) = \int_{-\infty}^{v} f(w)dw \qquad (38)$$

$$y(x(v)) = [1/E(x)]\int_{-\infty}^{tv} wf(w)dw \qquad (39)$$

Substitute Pareto distribution function,

$$x(v) = F(v) = 1-(k/v)^{\theta} \qquad (40)$$

$$y(x(v)) = [(\theta-1)/\theta^{k}]\int_{k}^{v} w\theta k^{\theta}w^{-\theta-1}dw \qquad (41)$$

or,

$$y(x(v)) = 1-(k/v)^{\theta-1} \quad (42)$$

Now, by solving (40) for "v" and substituting in (42), the Lorenz curve for Pareto distribution is derived as,

$$y = 1-(1-x)^{(\theta-1)/\theta} \quad (43)$$

As it was shown in the case of Pareto distribution, formula of Lorenz curve is easily obtained. But, if we select the log-normal density function (35), the procedure may not be the same. Because the integral of log-normal function has not been derived yet. In the following pages, the L_1 norm smoothing technique will be developed to estimate the parameters of given functional forms (32) and (33) by using the continuous probability density function.

According to (30) and (31) independent and dependent variables of (32) and (33) may be written as,

$$x(v) = \int_0^v f(w)dw \quad (44)$$

$$y(x(v)) = [1/E(x)] \int_0^v wf(w)dw \quad (45)$$

Substitute (44) and (45) inside (32) and define random error term u as,

$$[1/E(w)]\int_0^v wf(w)dw = \int_0^v f(w)dw.A^{\int_0^v f(w)dw-1} . e^u \quad (46)$$

or briefly,

$$y(x)=xA^{x-1}e^u \quad (47)$$

Similarly for the model (35),

$$[1/E(w)]\int_0^v wf(w)dw=\{\int_0^v f(w)dw\}^B . A^{\int_0^v f(w)dw-1} . e^u \quad (48)$$

or briefly,

$$y(x)=x^B A^{x-1}e^u \quad (49)$$

Taking natural logarithm of (47) and (49), gives,

$$\ln y(x)=\ln x + (x-1)\ln A + u \quad (50)$$

$$\ln y(x)=B.\ln x + (x-1)\ln A + u \quad (51)$$

With respect to properties of Lorenz curve and probability density function of f(w) and equations (46) to (49), it is obvious that x belongs to the interval [0,1]. Thus the L_1 norm objective function for minimizing (50) or (51) is given by,

$$\min: S = \int_0^1 |u|dx \quad (52)$$

Now, let us deal with L_1 norm estimation of "A" of Lorenz curve functional form (32) (redefined by (50)). The corresponding L_1 norm objective function will be,

$$\min_A: S = \int_0^1 |\ln y(x) - \ln x - (x-1) \ln A|dx \quad (53)$$

or,

$$\min_A: S = \int_0^1 |x-1||[\ln y(x)-\ln x]/(x-1) - \ln A|dx \quad (54)$$

By a similar technique used by (9), we can rewrite (54) as,

$$\min_A: S = \int_0^t |x-1|\{[\ln y(x)-\ln x]/(x-1)-\ln A\}dx - \int_t^1 |x-1|\{[\ln y(x)-\ln x]/(x-1)-\ln A\}dx \quad (55)$$

since, $0\le x\le 1$ we have,

$$\min_{A}: S = -\int_0^t [\ln y(x) - \ln x - (x-1)\ln A]dx + \int_t^1 [\ln y(x) - \ln x - (x-1)\ln A]dx \tag{56}$$

Differentiate (56) partially with respect to "t" and "A" and equate them to zero;

$$\frac{\delta S}{\delta A} = +\int_0^t [(x-1)/A]dx - \text{u}\int_t^1 [(x-1)/A]dx = 0 \tag{57}$$

$$\frac{\delta S}{\delta t} = -2[\ln y(t) - \ln t - (t-1)\ln A] = 0 \tag{58}$$

From equation (57), we have,

$$t = 1\pm\sqrt{2}/2 \tag{59}$$

Since "t" should belong to the interval [0,1], we accept,

$$t = 1-\sqrt{2}/2 \tag{60}$$

Substitute (60) in (58), and solve for "A", gives the L_1 norm estimation for "A" equal to,

$$A = \left[\frac{1-\sqrt{2}/2}{y(1-\sqrt{2}/2)}\right]^{\sqrt{2}} \tag{61}$$

Now, let us apply this procedure to another Lorenz curve functional form of (33) (redefined by (51)). Rewrite L1 norm objective function (52) for the model (51),

$$\min_{A,B}: S = \int_0^1 |\ln y(x) - B\ln x - (x-1)\ln A|dx \tag{62}$$

or,

$$\min_{A,B}: S=\int_0^1 |x-1||[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A|dx \tag{63}$$

The objective function (63) - by some changing on variables - is similar to (16). Thus, by a similar procedure to those of (17) through (29) we can write "S" as,

$$\min_{A,B}: S = \int_0^{t_1} |x-1|\{[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A\}dx$$

$$-\int_{t_1}^{t_2} |x-1|\{[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A\}dx$$

$$+\int_{t_1}^{1} |x-1|\{[\ln y(x)]/(x-1)-(\ln x)/(x-1)-\ln A\}dx \tag{64}$$

Since $0\leq x\leq 1$, then (64) reduces to,

$$\min_{A,B}: S = -\int_0^{t_1} [\ln y(x) - B\ln x - (x-1)\ln A]dx + \int_{t1}^{t_2} [\ln y(x) - B\ln x - (x-1)\ln A]dx$$

$$-\int_{t_2}^{1} [\ln y(x) - B\ln x - (x-1)\ln A]dx \tag{65}$$

Differentiate "S" partially with respect to "A", "B", t_1 and t_2 and equate them to zero,

$$\frac{\delta S}{\delta A} = \frac{1}{A}\left[\int_0^{t_1} (x-1)dx - \int_{t_1}^{t_2} (x-1)dx + \int_{t_2}^{1} (x-1)dx\right] = 0 \tag{66}$$

$$\frac{\delta S}{\delta B} = \int_0^{t_1} \ln(x)dx - \int_{t_1}^{t_2} \ln(x)dx + \int_{t_2}^{1} \ln(x)dx = 0 \tag{67}$$

$$\frac{\delta S}{\delta t_1} = -2\{\ln[y(t_1)] - B\ln(t_1) - (t_1-1)\ln(A)\} = 0 \quad (68)$$

$$\frac{\delta S}{\delta t_2} = 2\{\ln[y(t_2)] - B\ln(t_2) - (t_2-1)\ln(A)\} = 0 \quad (69)$$

The above system of simultaneous equations can be solved for the unknowns t_1, t_2, "A" and "B". Equation (66) is reduced to,

$$t_1^2-t_2^2-2(t_1-t_2)-1/2 = 0 \quad (70)$$

Equation (67) can be written as,

$$t_1(\ln t_1-1) - t_2(\ln t_2-1) – 1/2 = 0 \quad (71)$$

Calculate t1 from (70) as,

$$t_1 = 1 \pm\sqrt{q}\ (t_2^2-2t_2+3/2) \quad (72)$$

Since 0st1s1, we accept,

$$t_1 = 1 - \sqrt{(t_2^2-2t_2+3/2)} \quad (73)$$

Substitute t_1 from (73) into (71), and rearrange the terms, gives;

$$\ln \frac{[1-\sqrt{(t_2^2-2t_2+3/2)}]^{[1-\sqrt{(t_2^2-2t_2+3/2)}]}}{t_2^{t2}} + t_2-3/2+\sqrt{(t_2^2-2t_2+3/2)} = 0 \quad (74)$$

The root of equation (74) may be computed by a suitable numerical algorithm. However, it has been computed and rounded for five digits decimal point as,

$$t_2 = 0.40442 \quad (75)$$

Value of t1 is derived by substituting t2 into (73);

$$t_1 = 0.07549 \quad (76)$$

Values of "B" and "A" are computed from (68) and (69) using t_2 and t_1 given by (75) and (76). Thus,

$$B = \frac{(t_2-1)\ln y(t_1) - (t_1-1)\ln y(t_2)}{(t_2-1)\ln(t_1) - (t_1-1)\ln(t_2)} \quad (77)$$

or,

$$B = -0.84857\ln[y(0.07549)] + 1.31722\ln[y(0.40442)] \quad (78)$$

and,

$$A = [y(0.07549)]^{1.28986}[y(0.40442)]^{-3.68126} \quad (79)$$

Now, let us describe how equation (61) for the model (32) and equations (78) and (79) for the model (33) can be used to estimate the parameters of the Lorenz curve when the probability distribution function is known. In the model (32) we should solve (44) for $x(v)=1-\sqrt{2}/2$. On the other hand, we should find value of "v" such that,

$$x(v) = \int_0^v f(w)dw = 1-\sqrt{2}/2 \quad (80)$$

By substituting this value of "v" into (45), value of $y(1-\sqrt{2}/2)$ is computed. The value $y(1-\sqrt{2}/2)$ is used to compute the parameter "A" given by (61) for model (32).

The procedure for the model (33) is also similar, with the difference that two values of "v" should be computed. Once two different values of "v" are computed as follow,

$$x(v) = \int_0^v f(w)dw = 0.07549 \quad (81)$$

$$x(v) = \int_0^v f(w)dw = 0.40442 \quad (82)$$

Values of "v" are substituted in (45) to find y(0.07549) and y(0.40442). These values of "y" are used to compute the parameters of the model (33) by substituting them into (78) and (79).

The only problem remains is computation of related definite integrals of x(v) defined by (80), (81) and (82) which can be done by appropriate numerical methods such as the enclosed sample computer program coded for MathCAD 11 for a complete example.

7. Income distribution in the Unites States of America

In order to compute the Lorenz curve for the United States we try to apply the above procedure for both (32) and (33) propositions and using log-normal distribution function assumption. The source of data is "the U.S. economic report of president to parliament, different years". Median income and disposable personal income per family report by table 1. The amount of mean and median of income were used to derive the log-normal density function parameters μ and δ. The explained procedure of estimation then applied to the series of data for 1977-2002, and corresponding results are reported in next table 2. The results of Slottje (1989) which are based on quintile data calculations confirm our finding figures partially. Comparisons show the high compatibility of both procedures. An interesting finding of these calculations is that the distribution of income obeys a counter wise business cycles fluctuations. This finding is a new area for research in realm of the theory and application of income distribution and business cycles interrelationship.

A sample computer program is also enclosed at the end of these pages.

Table 1.

Year	Population millions	No. of families millions	Disposable personal income, billions of current $	Per capita disposable income $	Per family disposable income $	Family median Income current $	Gross domestic product billions of $	Real gross domestic product billions of chained (2000) $
1977	220.3	57.2	1435.7	6,517	25,098	16009.0	2,030.9	4,750.5
1978	222.6	57.8	1608.3	7,224	27,825	17639.9	2,294.7	5,015.0
1979	225.1	59.6	1793.5	7,967	30,091	19587.2	2,563.3	5,173.4
1980	227.7	60.3	2009.0	8,822	33,317	21023.2	2,789.5	5,161.7
1981	230.0	61.0	2246.1	9,765	36,820	22387.8	3,128.4	5,291.7
1982	232.2	61.4	2421.2	10,426	39,432	23433.3	3,255.0	5,189.3
1983	234.3	62.0	2608.4	11,131	42,070	24673.9	3,536.7	5,423.8
1984	236.4	62.7	2912.0	12,319	46,446	26433.1	3,933.2	5,813.6
1985	238.5	63.6	3109.3	13,037	48,890	27735.2	4,220.3	6,053.7
1986	240.7	64.5	3285.1	13,649	50,932	29458.2	4,462.8	6,263.6
1987	242.8	65.2	3458.3	14,241	53,042	30970.2	4,739.5	6,475.1
1988	245.1	65.8	3748.7	15,297	56,971	32191.0	5,103.8	6,742.7
1989	247.4	66.1	4021.7	16,257	60,844	34213.1	5,484.4	6,981.4
1990	250.2	66.3	4285.8	17,131	64,643	35353.3	5,803.1	7,112.5
1991	253.5	67.2	4464.3	17,609	66,435	35938.7	5,995.9	7,100.5
1992	256.9	68.2	4751.4	18,494	69,670	36573.1	6,337.7	7,336.6
1993	260.3	68.5	4911.9	18,872	71,709	36929.5	6,657.4	7,532.7
1994	263.5	69.3	5151.8	19,555	74,341	38781.9	7,072.2	7,835.5
1995	266.6	69.6	5408.2	20,287	77,705	40610.6	7,397.7	8,031.7
1996	269.7	70.2	5688.5	21,091	81,033	42300.2	7,816.9	8,328.9
1997	273.0	70.9	5988.8	21,940	84,467	44568.2	8,304.3	8,703.5
1998	276.2	71.6	6395.9	23,161	89,330	46736.8	8,747.0	9,066.9
1999	279.3	73.2	6695.0	23,968	91,461	48789.3	9,268.4	9,470.3
2000	282.5	73.8	7194.0	25,467	97,478	50731.7	9,817.0	9,817.0
2001	285.6	74.3	7469.4	26,156	100,531	51407.4	10,100.8	9,866.6
2002	288.6	75.6	7857.2	27,223	103,932	51680.0	10,480.8	10,083.0
2003	290.5		8039.2	27,675			10,735.8	10,210.4

http://www.gpoaccess.gov/eop/

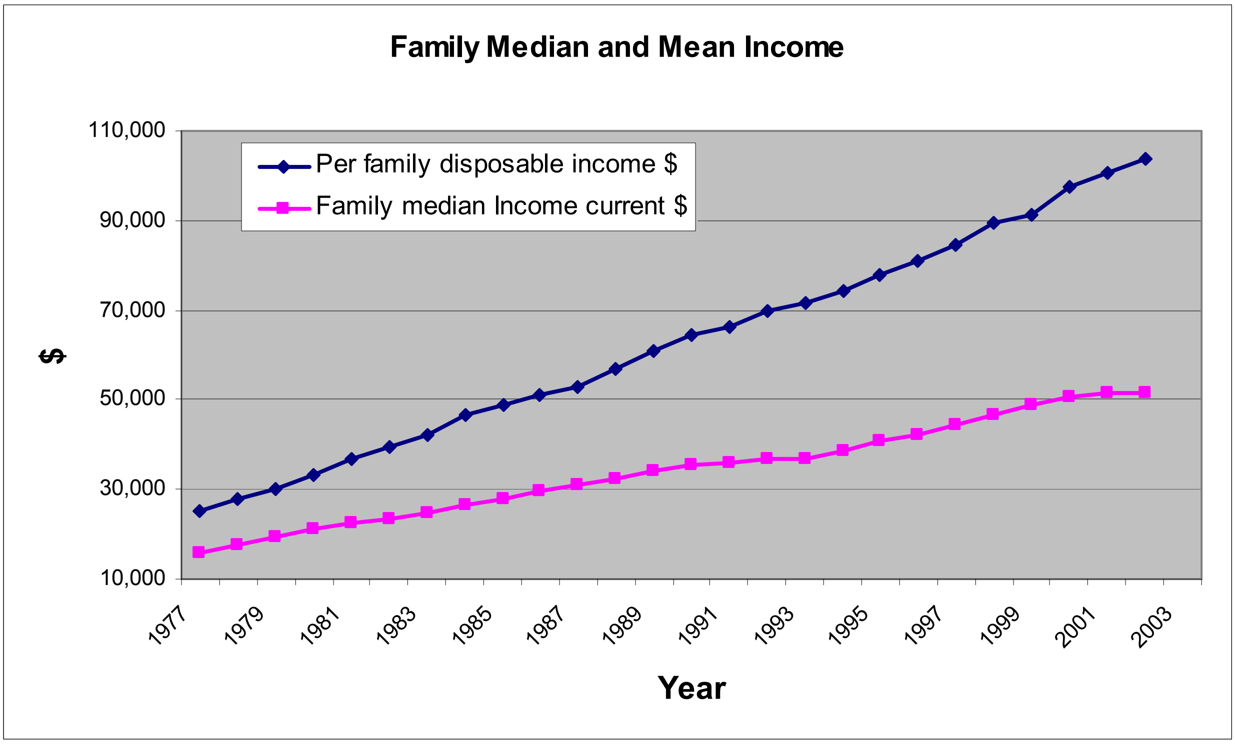
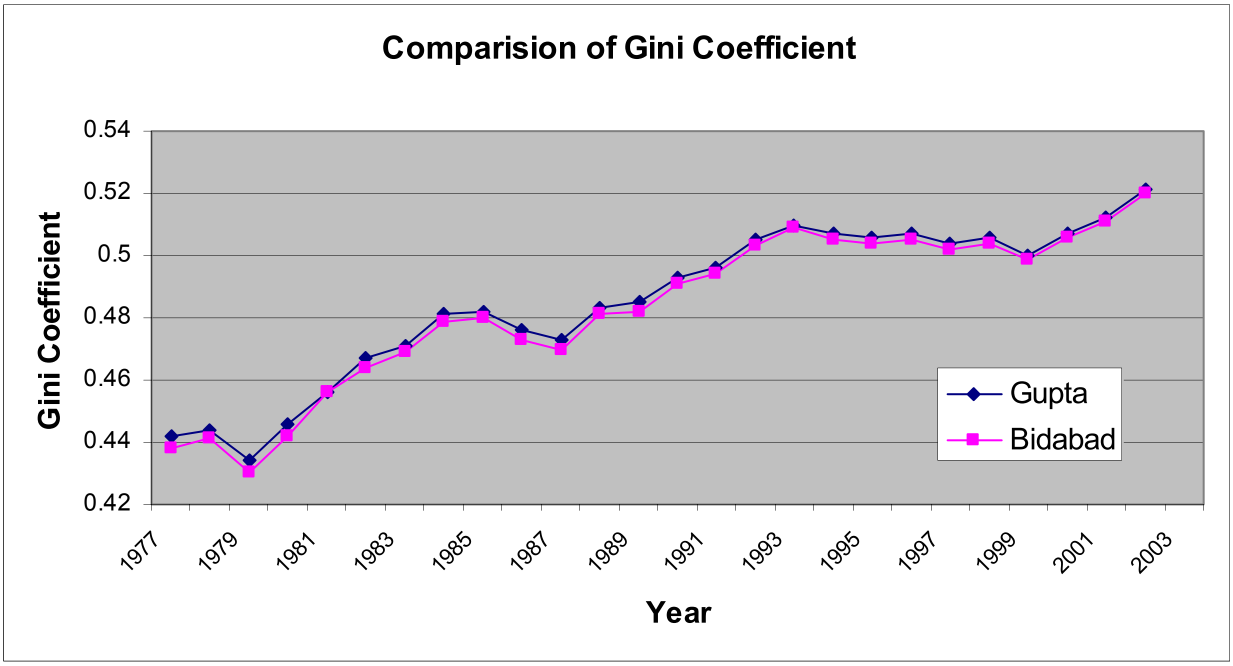
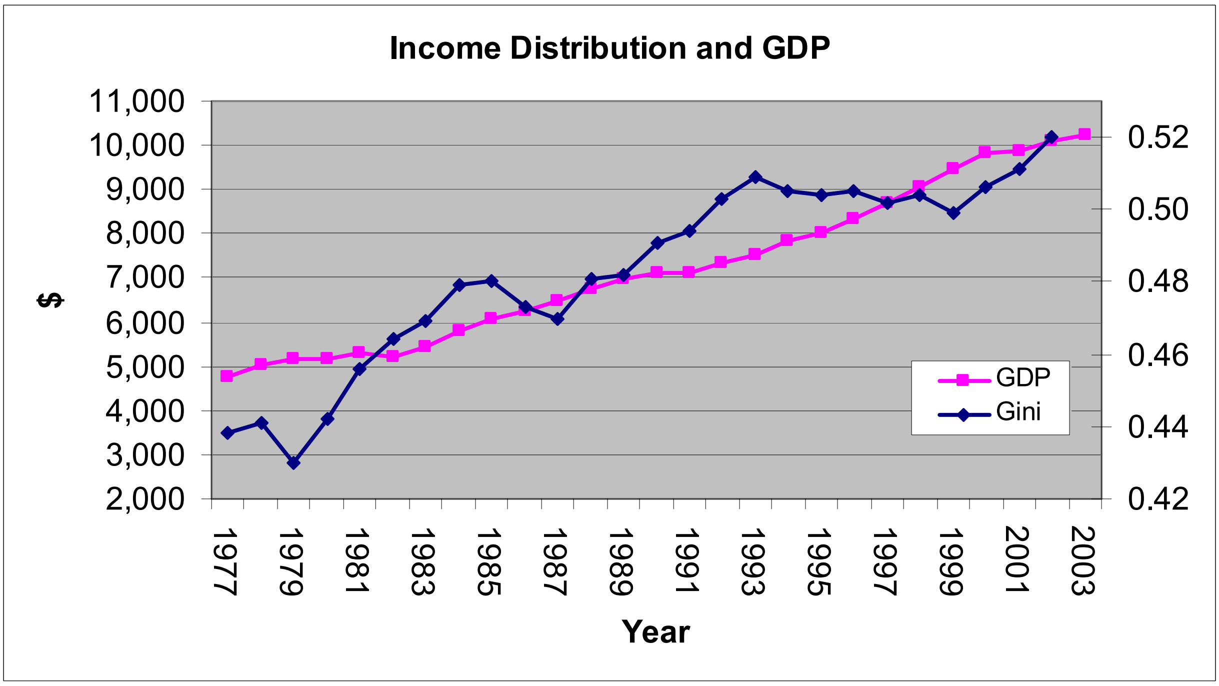


Table 2.

Year	Gupta Model			Bidabad model				Slottje figures	
	A	Gini	Kakwani	A	B	Gini	Kakwani	Gini	Kakwani
1977	7.938	0.442	0.172	5.798	1.214	0.438	0.170	0.426	0.109
1978	8.080	0.444	0.173	5.899	1.214	0.441	0.172	0.427	0.108
1979	7.484	0.434	0.166	5.475	1.212	0.430	0.164	0.427	0.111
1980	8.189	0.446	0.175	5.978	1.215	0.442	0.173	0.428	0.112
1981	9.095	0.456	0.185	6.631	1.218	0.546	0.183	0.435	0.114
1982	9.693	0.467	0.191	7.064	1.220	0.464	0.190	0.447	0.118
1983	10.051	0.471	0.194	7.324	1.221	0.469	0.193	0.447	0.120
1984	10.909	0.481	0.202	7.952	1.222	0.479	0.201	0.449	0.121
1985	11.004	0.482	0.203	8.021	1.223	0.480	0.202		
1986	10.442	0.476	0.198	7.609	1.222	0.473	0.197		
1987	10.175	0.473	0.196	7.416	1.221	0.470	0.194		
1988	11.123	0.483	0.204	8.110	1.223	0.481	0.203		
1989	11.269	0.485	0.205	8.216	1.223	0.482	0.204		
1990	12.137	0.493	0.212	8.858	1.224	0.491	0.211		
1991	12.493	0.496	0.215	9.122	1.225	0.494	0.214		
1992	13.518	0.505	0.222	9.886	1.226	0.503	0.221		
1993	14.207	0.510	0.226	10.403	1.226	0.509	0.226		
1994	13.741	0.507	0.223	10.052	1.226	0.505	0.223		
1995	13.676	0.506	0.223	10.004	1.223	0.504	0.222		
1996	13.717	0.507	0.223	10.034	1.226	0.505	0.222		
1997	13.339	0.504	0.221	9.751	1.226	0.502	0.220		
1998	13.637	0.506	0.223	9.973	1.226	0.504	0.222		
1999	12.962	0.500	0.218	9.472	1.225	0.499	0.217		
2000	13.825	0.507	0.224	10.115	1.226	0.506	0.223		
2001	14.470	0.512	0.229	10.600	1.226	0.511	0.227		
2002	15.759	0.521	0.235	11.573	1.227	0.520	0.235		



The following graph compares the calculated Gini coefficient with real GDP for the period of 1977-2002.



References

- Bidabad B. (1987a) Least absolute error estimation. Submitted to the First International Conference on Statistical Data Analysis Based on the L_1 norm and Related Methods, Neuchatel, Switzerland.
- Bidabad B. (1987b) Least absolute error estimation, part II. Submitted to the First International Conference on Statistical Data Analysis Based on the L_1 norm and Related Methods, Neuchatel, Switzerland.
- Bidabad B. (1988a) A proposed algorithm for least absolute error estimation. Proceedings of the Third Seminar of Mathematical Analysis. Shiraz University, 24-34, Shiraz, Iran.
- Bidabad B. (1988b) A proposed algorithm for least absolute error estimation, part II. Proceedings of the Third Seminar of Mathematical Analysis, Shiraz University, 35-50, Shiraz, Iran.
- Bidabad B. (1989a) Discrete and continuous L_1 norm regressions, proposition of discrete approximation algorithms and continuous smoothing of concentration surface, Ph.D. thesis, Islamic Azad University, Tehran, Iran.
- Bidabad B. (1989b) Discrete and continuous L_1 norm regressions, proposition of discrete approximation algorithms and continuous smoothing of concentration surface, Ph.D. thesis, Islamic Azad University, Tehran, Iran. Persian translation.
- Bidabad B., B. Bidabad (1989) Functional form for estimating the Lorenz curve. Submitted to the Australasian Meeting of Econometric Society, Canberra, Australia.
- Bidabad B., B. Bidabad (1992) Functional form for estimating the Lorenz curve. To be appeared in Economics and Management, quarterly Journal of the Islamic Azad University.
- Cramer J.S. (1973) Empirical econometrics. North-Holland, Amsterdam.
- Gupta M.R. (1984) Functional forms for estimating the Lorenz curve. Econometrica, 52, 1313-1314.
- Hobby C.R., J.R. Rice (1965) A moment problem in L_1 approximation. Proc. Amer. Math. Soc., 16, 665-670.
- Kakwani N.C. (1980) Income inequality and poverty. New York, Oxford University Press.
- Kakwani N.C. (1980) Functional forms for estimating the Lorenz curve: a reply. Econometrica, 48, 1063-64.
- Kakwani N.C., N. Podder (1976) Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations. Econometrica 44, 137-148.
- Kendall M., A. Stuart (1977) The advanced theory of statistics. vol.1, Charles Griffin & Co., London.
- Kripke B.R., T.J. Rivlin (1965) Approximation in the metric of $L_1(X,u)$. Trans. Amer. Math. Soc., 119, 101-22.
- Lazarski E. (1975a) Approximation of continuous functions in the space L_1. Automatika, 487, 85-93.
- Lazarski E. (1975b) The approximation of the continuous function by the polynomials of power functions in L_1 space. Automatika, 487, 95-106.
- Lazarski E. (1975c) On the necessary conditions of the uniqueness of approximation by the polynomials of power functions in L_1 space. Automatika, 487, 107-117.
- Lazarski E. (1977) Approximation of continuous functions by exponential polynomials in the L_1 space. Automatika, 598, 82-87.
- Ptak V. (1958) On approximation of continuous functions in the metric $\int_a^b|x(t)|dt$ Czechoslovak Math. J. 8(83), 267-273.

- Rasche R.H., J. Gaffney, A.Y.C. Koo, N. Obst (1980) Functional forms for estimating the Lorenz curve. Econometrica, 48, 1061-1062.
- Rice J.R. (1964a) On computation of L_1 approximations by exponentials, rationals, and other functions. Math. Comp., 18, 390-396.
- Rice J.R. (1964b) On nonlinear L_1 approximation. Arch. Rational Mech. Anal., 17 61-66.
- Rice J.R. (1964c) The approximation of functions, vol. I, linear theory. Reading Mass:, Addison-Wesley.
- Rice J.R. (1969) The approximation of functions, vol. II, linear theory. Reading Mass:, Addison-Wesley.
- Rice J.R. (1985) Numerical methods, software, and analysis. McGraw-Hill, ch. 11.
- Rice J.R., J.S. White (1964) Norms for smoothing and estimation. SIAM Rev., 6, 243-256.
- Salem A.B.Z., T.D. Mount (1974) A convenient descriptive model of income distribution: the gamma density. Econometrica, 42, 1115-1127.
- Singh S.K., G.S. Maddala (1976) A function for the size distribution of income. Econometrica, 44, 963-970.
- Slottje D.J. (1989) The structure of earnings and the measurement of income inequality in the U.S., North-Holland Publishing Company, Amsterdam.
- Taguchi T. (1972a) On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case-I. Annals of the Inst. of Stat. Math., vol. 24, no.2, 355-381.
- Taguchi T. (1972b) On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case-II. Annals of the Inst. of Stat. Math., vol. 24, no.3, 599-619.
- Taguchi T. (1972c) Concentration polyhedron, two dimensional concentration coefficient for discrete type distribution and some new correlation coefficients etc. The Inst. of Stat. Math., 77-115.
- Taguchi T. (1973) On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case-III. Annals of the Inst. of Stat. Math., vol. 25, no.1, 215-237.
- Taguchi T. (1974) On Fechner's thesis and statistics with norm p. Ann. of the Inst. of Stat. Math., vol. 26, no.2, 175-193.
- Taguchi T. (1978) On a generalization of Gaussian distribution. Ann. of the Inst. of Stat. Math., vol. 30, no.2, A, 211-242.
- Taguchi T. (1981) On a multiple Gini's coefficient and some concentrative regressions. Metron, vol. XXXIX - N.1-2, 5-98.
- Taguchi T. (1983) Concentration analysis of bivariate Paretoan distribution. Proc. of the Inst. of Stat. Math., vol. 31, no.1, 1-32.
- Taguchi T. (1987) On the structure of multivariate concentration. Submitted to the First International Conference on Statistical Data Analysis Based on the L_1 Norm and Related Methods, Neuchatel, Switzerland.
- Taguchi T. (1988) On the structure of multivariate concentration–some relationships among the concentration surface and two variates mean difference and regressions. CSDA, 6,307-334.
- Usow K.H. (1967a) On L_1 approximation: computation for continuous functions and continuous dependence. SIAM J. of Numer. Anal., 4, 70-88.
- Watson G.A. (1981) An algorithm for linear L_1 approximation of continuous functions. IMA J. Num. Anal., 1, 157-167.

CONTINUOUS L NORM ESTIMATION OF LORENZ CURVE

Bijan BIDABAD

(Using Sample Mean and Median)

Calculations for 2002 USA data

This program has been coded for MathCAD 11

Mean = Sample mean of income distribution: $Mean := 10393$

Med = Sample median of income distribution: $Med := 5168$

$$\sigma := \sqrt{2 \cdot \ln\left(\frac{Mean}{Med}\right)}$$

Calculation of Log-Normal density function parameters m and s according to sample mean and median

$$\sigma = 1.18209$$

$$\mu := \ln(Med) \qquad \mu = 10.85283$$

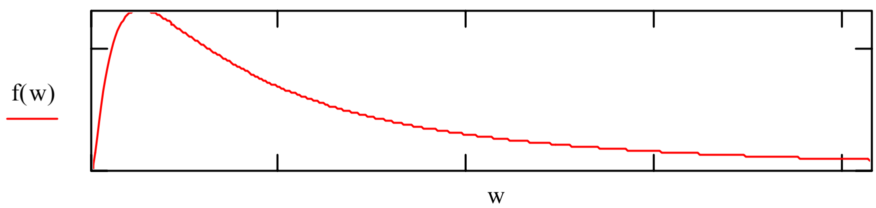
$$f(w) := \left(\frac{1}{w \cdot \sigma \cdot \sqrt{2 \cdot \pi}}\right) \cdot \exp\left[\frac{-(\ln(w) - \mu)^2}{2 \cdot \sigma^2}\right]$$

Log-Normal Probability_Density Function

$$w := 10^{-5}, \frac{Mean}{200} .. 2 \cdot Mean$$

Selective range for_Log-Normal plot, values of_increment and upper bound_may be changed

Log-Normal plot



Precision Tolerance level $TOL := 0.0000$

TOL value should be_ changed for more_ accurate solutions,_(less TOL = higher precision)

(45) $$y(v) := \left(\frac{1}{Mean}\right) \cdot \int_0^v w \cdot f(w)\, dw$$

(44) $$x(v) := \int_{0.00001}^{v} f(w)\, dw$$

Calculation for Gupta model

Initial guess for v. This value should be changed for faster convergence and less iterations

$v := 2000($

$$t_0 := 1 - \frac{\sqrt{2}}{2}$$

(60)

Calculating v for (80) $\quad v := \mathrm{root}\left(x(v) - t_0, v\right)$

Calculated v $\quad v = 27136.6437$

y(t)_ 0 $\quad y(v) = 0.04208 \qquad z_0 := y(v)$

$$A := \left(\frac{t_0}{z_0}\right)^{\sqrt{2}}$$

(61), estimated A: $\quad A = 15.54768$

$$S := \int_0^1 \left|\ln\left(z_0\right) - \ln\left(t_0\right) - \left(t_0 - 1\right)\cdot\ln(A)\right| \, dx$$

(53)

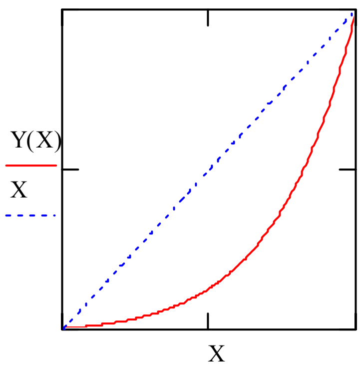
Sum of absolute residuals $\quad S = 0$

Range variable for plotting the Lorenz curves

$X := 0, 0.005 .. 1$

$$Y(X) := X \cdot A^{X-1}$$

Gupta Lorenz curve:



Calculation of Gini coefficient

$$Gini := 1 - 2 \cdot \int_0^1 Y(X) \, dX$$

$Gini = 0.51967$

Calculation of Kakwani length of Lorenz curve

$$Length := \int_0^1 \sqrt{1 + \left[A^{X-1} \cdot (1 + X \cdot \ln(A))\right]^2} \, dX$$

Length of Lorenz curve

$$\text{Length} = 1.5515$$

$$\text{Kakwani} := \frac{\text{Length} - \sqrt{2}}{2 - \sqrt{2}}$$

Kakwani index of length $\text{Kakwani} = 0.23437$

Calculation For Bidabad Model

(76) $t_1 := 0.0754$

Initial guess for v. This value should be changed for faster convergence and less iterations

$$v := 800($$

Calculating v for (81) $v := \text{root}\left(x(v) - t_1, v\right)$

Calculated v $v = 9464.04318$

y(0.07549) $y(v) = 0.00442$ $z_1 := y(v)$

(75) $t_2 := 0.4044$

Initial guess for v. This value should be changed for faster convergence and less iterations

$$v := 2700($$

Calculatig v for (82) $v := \text{root}\left(x(v) - t_2, v\right)$

Calculated v $v = 38826.25803$

y(0.40442) $y(v) = 0.07722$ $z_2 := y(v)$

(79) $A := \left(z_1\right)^{1.28986} \cdot \left(z_2\right)^{-3.68126}$

(78) $B := -0.84857 \ln\left(z_1\right) + 1.31722 \ln\left(z_2\right)$

Estimated A and B: $A = 11.41481$ $B = 1.22709$

(62) $S := \int_0^1 \left|\ln\left(z_1\right) - B \cdot \ln\left(t_1\right) - \left(t_1 - 1\right) \cdot \ln(A)\right| \, dx$

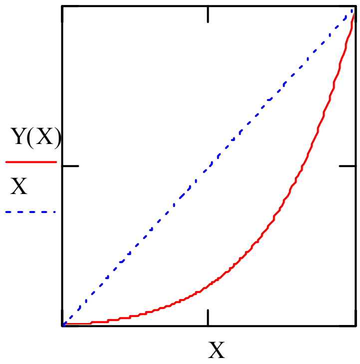
Sum of absolute residuals $S = 0.00002$

Range variable for plotting the Lorenz curves

$$X := 0, 0.005 .. 1$$

Bidabad Lorenz curve $Y(X) := X^B \cdot A^{X-1}$

Calculation of Gini coefficient



$$Gini := 1 - 2\cdot\int_0^1 Y(X)\,dX$$

$$Gini = 0.51834$$

Calculation of Kakwani length of Lorenz curve

$$Length := \int_0^1 \sqrt{1 + \left[A^{X-1}\cdot X^{B-1}\cdot(B + X\cdot\ln(A))\right]^2}\,dX$$

Length of Lorenz curve

$$Length = 1.55118$$

$$Kakwani := \frac{Length - \sqrt{2}}{2 - \sqrt{2}}$$

Kakwani index of length

$$Kakwani = 0.23381$$