

Semi-Heuristic Poverty Measures Used by Economists: Justification Motivated by Fuzzy Techniques

Karen Villaverde
Department of Computer Science
New Mexico State University
Las Cruces, NM 88003, USA
Email: kvillave@cs.nmsu.edu

Nagwa Albehery
and Tonghui Wang
Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA
Email: {albehery,twang}@nmsu.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
Email: vladik@utep.edu

Abstract—To properly gauge the extent of poverty in a country or in a region, economists use semi-heuristic poverty measures such as the Foster-Greer-Thorbecke (FGT) metric. These measures are used because it was empirically shown that they capture the commonsense meaning of the extent of poverty better than previously proposed measures. However, the fact that these measures are better than a few earlier proposed ones does not necessarily mean that these measures are the best possible; so, it is desirable to look for optimal poverty measures. In this paper, we first use fuzzy techniques to provide a commonsense interpretation of FGT poverty measures, and then show that under certain reasonable conditions, these measures are indeed optimal.

I. INTRODUCTION

How poverty is measured now. How do we gauge the extent of poverty in a country or in a region? Usually, there is a threshold z (called *poverty line* or *poverty threshold*), which is the minimum level of income deemed necessary to achieve an adequate standard of living in a given country, so that:

- a person i with an income x_i at or below z is considered poor, while
- a person i whose income is larger than this threshold is not considered to be poor.

At first glance, it may seem natural to simply count the number H of poor people, and to take the proportion $F_0 = \frac{H}{N}$ of poor people (i.e., people whose income is at or below the poverty line z) to the population as a whole as an appropriate measure for the extent of poverty. This proportion is indeed used as a measure of poverty – especially in the media, where the economic hardship is easy to describe by saying, e.g., that 20% of the people live below the poverty line.

However, the proportion F_0 (called the *incidence of poverty* or the *headcount ratio*) may not be the most adequate poverty measure, because it does not distinguish between those who income is close to z (and who are “almost” not poor) and those whose income is much smaller than the poverty line – and who are thus suffering much more from their poverty. To

capture this difference, economists use special Foster-Greer-Thorbecke (FGT) property measures (first introduced in [6]):

$$F_1 = \frac{1}{N} \cdot \sum_{i=1}^H \left(1 - \frac{x_i}{z}\right) \text{ and } F_2 = \frac{1}{N} \cdot \sum_{i=1}^H \left(1 - \frac{x_i}{z}\right)^2,$$

where the sum is over all persons whose income is at or below poverty level. F_1 is called the *intensity* of poverty, and F_2 is called the *severity* of poverty. Both measures give more weight to people whose income is smaller.

Advantages and limitations of the existing poverty measures. Empirically, the poverty measures F_0 , F_1 , and F_2 work well, they are used by economists and governments throughout the world to gauge poverty – and thus, to gauge the success of different measures aimed at reducing poverty. For example, in Mexico, the measure F_2 is officially used as a poverty measure; in Egypt, all three measures are used; see, e.g., [1], [2], [3], [4], [5].

The main problem with these measures is that they are semi-heuristic. There is no precise justification and therefore, there is no guarantee that a slightly modified version of one of these measures would not provide a better description of the extent of poverty. It is therefore desirable to look for poverty measures which are optimal – in some reasonable sense.

What we do in this paper. In this paper, we provide a fuzzy-motivated justification of FGT poverty measures:

- we start with an informal fuzzy motivation, and then
- we describe a precise mathematical derivation of FGT measures motivated by these fuzzy ideas.

II. USING FUZZY TECHNIQUES TO PROVIDE A COMMONSENSE INTERPRETATION OF FGT POVERTY MEASURES

Main idea: using fuzzy logic. The main problem with simply counting the number of poor people is that, like many other things in the world, poverty is a matter of degree: one person can be simply poor, another is somewhat poor, the third person

may be very poor. The fact that many properties are not always absolutely true or absolutely false – but often true to a degree – was the main reason why L. Zadeh introduced *fuzzy logic*; see, e.g., [7], [9], [10]. In fuzzy logic, this degree is usually described by a number from the interval $[0, 1]$:

- the degree 1 means that the property is absolutely true,
- the degree 0 means that the property is absolutely false, and
- the degrees between 0 and 1 mean that the property is true to some extent.

It is therefore reasonable to use fuzzy logic to describe poverty.

Corresponding membership function. We need to have a function that describes, based on the person income x , a degree $\mu(x)$ to which this person is poor. In deriving this function, we can use the following two commonsense facts:

- When the income x is 0 (the smallest possible amount) the person is absolutely poor. In this case, the degree $\mu(0)$ to which this person is poor should be equal to 1.
- When the income x is equal to z (the largest possible amount at which a person is still considered poor), the person is only one cent away from being absolutely not poor. In this case, it is reasonable to require that the degree $\mu(z)$ to which this person is poor should be equal to 0.

Thus, we need to define a function $\mu(x)$ on the interval $[0, z]$ for which $\mu(0) = 1$ and $\mu(z) = 0$.

In principle, there exist many different functions with this property, some simpler, some more complex. The simplest possible functions are linear functions. Thus, it is reasonable to select a linear function $\mu(x)$ for which $\mu(0) = 1$ and $\mu(z) = 0$. (A more mathematical argument for selecting linear or piecewise linear membership functions is given in [8].)

A linear function is uniquely determined by its value at two points $x = 0$ and $x = z$, so we get a uniquely determined function $\mu(x) = 1 - \frac{x}{z}$.

Counting the proportion of number of poor people: three natural possibilities. It is natural to describe the extent of poverty as the ratio between the number of poor people (or, to be more precise, the cardinality of the set of poor people) and the total population.

The first idea, as we have mentioned earlier, is to consider poverty as a crisp property:

- people whose income x is below or at the poverty level z are considered to be poor, while
- people whose income x is above the poverty level z are considered to be not poor.

In this case, the above ratio is the measure F_0 .

A more adequate idea is to take into account that poverty is a fuzzy property, and the set of poor people is a fuzzy set: a person whose income is x belongs to this set with the degree $\mu(x) = 1 - \frac{x}{z}$. To apply the above idea to this fuzzy set, we need to be able to estimate the cardinality of a fuzzy set.

The most widely used way to define a cardinality of a fuzzy set S with a membership function $\mu(x)$ is to define it as the sum of all the membership values $\sum_x \mu(x)$; see, e.g., [7], [9]. For the above membership function, this cardinality is equal to $\sum_{i=1}^H \left(1 - \frac{x_i}{z}\right)$ and therefore, the corresponding ratio of this cardinality to the total number N of people is equal to FGT measure F_1 .

Eliminating poverty is an important ultimate goal. As we have mentioned, some people are somewhat poor, some are very poor. Clearly, from the political viewpoint, it is most important to take care of those who are very poor. From this viewpoint, instead of counting the number of poor people, it may be more reasonable to count the number of people who are very poor. In fuzzy logic, the most widely used way to describe the hedge “very” is to use a squaring operation: the degree to which a property is very satisfied is equal to the square $\mu^2(x)$ of the degree to which this property is satisfied; see, e.g., [7], [9]. From this viewpoint, the degree $\mu_v(x)$ to which a person is very poor is equal to $\mu_v(x) = \left(1 - \frac{x}{z}\right)^2$. Thus, the cardinality of the set of all very poor people is equal to $\sum_{i=1}^H \mu_v(x_i) = \sum_{i=1}^H \left(1 - \frac{x_i}{z}\right)^2$, and the ratio of this cardinality to the total number N of people is equal to the FGT measure F_2 .

Conclusion of this section. All three FGT measures F_0 , F_1 , and F_2 naturally appear in the fuzzy interpretation, as the ratio of the number of poor people to the population as a whole:

- the poverty measure F_0 appears when we consider poverty to be a crisp property, when every person is either poor or not poor;
- the poverty measure F_1 appears when we take into account, when counting the number of poor people, that poverty is a fuzzy property, so that every poor person is poor to a certain degree;
- finally, the poverty measure F_2 appears when instead of simply counting the number of poor people, we count the number of *very* poor people.

III. FROM AN INFORMAL FUZZY JUSTIFICATION TO A PRECISE MATHEMATICAL JUSTIFICATION

Possibility. Since we have shown that the FGT measures naturally come from fuzzy techniques, and fuzzy techniques have a precise mathematical foundation, it makes sense to look for precise mathematical justifications of FGT measures.

Towards a general description of possible poverty measures. We will assume that in the population of N people, there are H poor ones, with incomes x_1, \dots, x_H . The authors of the original paper [6] consider *decomposable* poverty measures, i.e., measures for which, crudely speaking,

- once we know the poverty measures corresponding to two subareas of a given area, and we know the total populations N_1 and N_2 and the total numbers of poor people H_1 and H_2 in these subareas,

- we will be able to compute the poverty measure corresponding to the area as a whole.

They showed that all such measures are proportional to the sums $\sum_{i=1}^H f(x_i)$ for an appropriate function $f(x)$. In view of this conclusion, we will consider poverty measures of this type. Thus, we arrive at the following definition.

Definition 1. By an income, we mean a non-negative real number. By incomes, we mean a tuple of incomes x_1, \dots, x_H .

Definition 2. By a poverty measure, we mean a smooth (twice differentiable) function $f(x)$ from real numbers to real numbers. By the value of a poverty measure $f(x)$ corresponding to incomes x_1, \dots, x_H , we mean the sum

$$v(x_1, \dots, x_H) = \sum_{i=1}^H f(x_i).$$

Comment. Instead of a single poverty measure, we may want to consider several different poverty measures $f_1(x), \dots, f_k(x)$ that describe different aspects of poverty. In this case, for each tuple of incomes x_1, \dots, x_H , instead of a single poverty value v , we have k different poverty values

$$v_j = \sum_{i=1}^H f_j(x_i), \quad j = 1, \dots, k.$$

Important question. Because of the above definition, to select an appropriate poverty measure, we need to select an appropriate function $f(x)$ – or, in case of several poverty measures, select functions $f_1(x), \dots, f_k(x)$ corresponding to all these measures. How can we select these functions?

It is important to select independent poverty measures. Once we have property measures $f_1(x), \dots, f_k(x)$, we can then combine the resulting property values v_1, \dots, v_k . For example, we can take an average $\frac{v_1 + v_2}{2}$, which corresponds to a poverty measure $f(x) = \frac{f_1(x) + f_2(x)}{2}$. It makes no sense to add such “combined” property measures to the original list: indeed, to compute the poverty value corresponding to this measure, we do not need to painstakingly add the values of $f(x_i)$, it is enough to apply the corresponding combination function to the poverty values corresponding to the original poverty measures.

It is therefore reasonable to require that the selected poverty measures are *independent* in the sense that none of these poverty measures can be described as a function of others.

Definition 3. We say that a poverty measure $f(x)$ depends on the poverty measures $f_1(x), \dots, f_k(x)$ if it is possible to uniquely reconstruct the poverty value $v = \sum_{i=1}^H f(x_i)$ based on

the poverty values $v_j = \sum_{i=1}^H f_j(x_i)$ ($j = 1, \dots, k$), i.e., if there exists a function F for which $v = F(v_1, \dots, v_k)$. We say that a set of poverty measures $f_1(x), \dots, f_k(x)$ is independent if none of these measures depends on the others.

Idea. How do we select the set of independent poverty measures? One of the main objectives of using poverty measures is to improve the life of poor people by providing them with immediate financial help. There are two main ways of providing such help:

- One possibility is to allocate a certain fixed amount of money (or goods) a to each poor person. In the US, an example of such allocation is distributing food stamps – approximately the same amount goes to every person below a certain poverty threshold. In this case, the original incomes change from x_i to $x'_i = x_i + a$.
- Another possibility is to provide tax deductions to all the poor people; this is also done in the US. Since taxes are usually proportional to the income, tax deduction means, in effect, that the resulting income x_i of all poor people increases by the same factor, i.e., goes from x_i to $x'_i = \lambda \cdot x_i$, for some constant $\lambda > 0$.

In both cases, instead of the original poverty values $v_j = \sum_{i=1}^H f_j(x_i)$, we get new poverty values $v'_j = \sum_{i=1}^H f_j(x'_i)$ which are equal either to $\sum_{i=1}^H f_j(x_i + a)$ or to $\sum_{i=1}^H f_j(\lambda \cdot x_i)$. Useful poverty measure should enable us to easily predict the effect of both measures on the poverty values. Thus, we arrive at the following definition.

Definition 4. An independent set of poverty measures $f_1(x), \dots, f_k(x)$ is called efficient if the following two properties hold:

- for each incomes tuple x_1, \dots, x_H , once we know all k poverty values $v_j = \sum_{i=1}^H f_j(x_i)$ and the real value a , we can uniquely predict poverty values $v'_j = \sum_{i=1}^H f_j(x_i + a)$;
- for each incomes tuple x_1, \dots, x_H , once we know all k poverty values $v_j = \sum_{i=1}^H f_j(x_i)$ and the real value λ , we can uniquely predict poverty values $v'_j = \sum_{i=1}^H f_j(\lambda \cdot x_i)$.

Comment. One can easily check that the set of three FGT poverty measures $f_0(x) = 1$, $f_1(x) = 1 - \frac{x}{z}$, and $f_2(x) = \left(1 - \frac{x}{z}\right)^2$ is efficient in this sense. Indeed, since these three functions are polynomials of 0-th, 1-st, and 2-nd order, knowing the corresponding sums $v_j = \sum_{i=1}^H f_j(x_i)$ is equivalent to knowing the values $V_0 = \sum_{i=1}^H 1$, $V_1 = \sum_{i=1}^H x_i$, and $V_2 = \sum_{i=1}^H x_i^2$. When we add a , the value V_0 does not change $V'_0 = V_0$, the

value V_1 changes to

$$V'_1 = \sum_{i=1}^H x'_i = \sum_{i=1}^H (x_i + a) = \sum_{i=1}^H x_i + \sum_{i=1}^H a = V_1 + a \cdot V_0,$$

and the value V_2 changes to

$$V'_2 = \sum_{i=1}^H (x'_i)^2 = \sum_{i=1}^H (x_i + a)^2 = \sum_{i=1}^H x_i^2 + 2a \cdot \sum_{i=1}^H x_i + \sum_{i=1}^H a^2 = V_2 + 2a \cdot V_1 + a^2 \cdot V_0.$$

Similarly, when we multiply all the values x_i by a constant λ , the value V_0 does not change $V'_0 = V_0$, the value V_1 changes to

$$V'_1 = \sum_{i=1}^H x'_i = \sum_{i=1}^H \lambda \cdot x_i = \lambda \cdot \sum_{i=1}^H x_i = \lambda \cdot V_1,$$

and the value V_2 changes to

$$V'_2 = \sum_{i=1}^H (x'_i)^2 = \sum_{i=1}^H (\lambda \cdot x_i)^2 = \sum_{i=1}^H \lambda^2 \cdot x_i^2 = \lambda^2 \cdot \sum_{i=1}^H x_i^2 = \lambda^2 \cdot V_2.$$

In both cases, once we know the original values V_0 , V_1 , and V_2 and the corresponding parameter a or λ , we can uniquely determine the new values V'_0 , V'_1 , and V'_2 . It turns out that the FGT poverty measures are the only ones with this property.

Definition 5. We say that two independent sets of poverty measures $f_1(x), \dots, f_k(x)$ and $g_1(x), \dots, g_l(x)$ are equivalent if the following two conditions hold:

- each poverty measure $f_j(x)$ depends on the property measures $g_1(x), \dots, g_l(x)$; and
- each property measure $g_j(x)$ depends on the poverty measures $f_1(x), \dots, f_k(x)$.

Comment. In other words:

- If we know all the poverty values v_1, \dots, v_k corresponding to poverty measures $f_1(x), \dots, f_k(x)$, then we can uniquely reconstruct poverty values w_1, \dots, w_l corresponding to poverty measures $g_1(x), \dots, g_l(x)$.
- Vice versa, if we know all the poverty values w_1, \dots, w_l corresponding to poverty measures $g_1(x), \dots, g_l(x)$, then we can uniquely reconstruct poverty values v_1, \dots, v_k corresponding to poverty measures $f_1(x), \dots, f_k(x)$.

From this viewpoint, the two sets of poverty measures are indeed equivalent.

The following proposition provides an example of equivalent sets of poverty measures.

Proposition. The set of FGT poverty measures is equivalent to the measures 1 , x , and x^2 .

Indeed, e.g., since $f_1(x) = 1 - \frac{x}{z}$, the value F_1 can be represented as

$$F_1 = \frac{1}{N} \cdot \sum_{i=1}^H f_1(x_i) = \frac{1}{N} \cdot \sum_{i=1}^H \left(1 - \frac{x_i}{z}\right) = \frac{1}{N} \cdot \sum_{i=1}^H 1 - \frac{1}{N \cdot z} \cdot \sum_{i=1}^H x_i.$$

Now, we are ready to formulate our main result.

Theorem. Every efficient independent set of poverty measures $f_1(x), \dots, f_k(x)$ is equivalent to the set consisting of the measures 1 , x , x^2 , \dots , x^{k-1} .

The proof of this theorem is given in the next section.

Corollary. Every efficient independent set of poverty measures $f_1(x), f_2(x), f_3(x)$ is equivalent to the set consisting of the three FGT measures.

Comment. Thus, we arrive at the desired justification of the FGT measures.

IV. PROOF

1°. Efficiency means, in particular, that if we change the numbers x_1, \dots, x_H (representing incomes) without changing the values of the original poverty values $v_j = \sum_{i=1}^H f_j(x_i)$, then

the values $v'_j = \sum_{i=1}^H f_j(x_i + a)$ also remain unchanged.

For small changes Δx_i , the resulting change in v_j is equal to

$$\Delta v_j = \sum_{i=1}^H f'_j(x_i) \cdot \Delta x_i + o(\Delta x_i),$$

where f' denote the derivative. Thus, the requirement that $\Delta v_j = 0$ means that

$$\sum_{i=1}^H f'_j(x_i) \cdot \Delta x_i = o(\Delta x_i)$$

for all j from 1 to k .

Under this requirement, the value v'_j should also not change. The corresponding change in v'_j is equal to

$$\Delta v'_j = \sum_{i=1}^H f'_j(x_i + a) \cdot \Delta x_i + o(\Delta x_i),$$

i.e., we should have

$$\sum_{i=1}^H f'_j(x_i + a) \cdot \Delta x_i = o(\Delta x_i).$$

If we take any vector $dx = (dx_1, \dots, dx_n)$ and take $\Delta x_i = \varepsilon \cdot dx_i$, then in the limit $\varepsilon \rightarrow 0$, we arrive at the following conclusion:

- if $\sum f'_j(x_i) \cdot dx_i = 0$ for all j ,
- then $\sum f'_j(x_i + a) \cdot dx_i = 0$ for all j .

2°. The above requirement can be described in the vector form, since the sum $\sum f'_j(x_i) \cdot dx_i$ is a dot (scalar) product $\langle f'_j, dx \rangle$ of two vectors: the vector $f'_j = (f'_j(x_1), \dots, f'_j(x_H))$ and the vector dx . Similarly, the sum $\sum f'_j(x_i+a) \cdot dx_i$ is a dot product $\langle f'_{aj}, dx \rangle$, where

$$f'_{aj} = (f'_j(x_1+a), \dots, f'_j(x_H+a)).$$

In this vector form, the above requirement takes the following form:

- if $\langle f'_j, dx \rangle = 0$ for all j ,
- then $\langle f'_{aj}, dx \rangle = 0$ for all j .

3°. The condition $\langle a, b \rangle = 0$ means that the vectors a and b are orthogonal to each other: $a \perp b$. Thus, the above requirement means that every vector dx which is orthogonal to all the vectors f'_j is also orthogonal to all the vectors f'_{aj} .

As is well known in vector algebra, each vector f'_{aj} can be decomposed into two components: a component c_{\parallel} which belongs to the linear space L generated by the vectors f'_j , and a component c_{\perp} which is orthogonal to this space:

$$f'_{aj} = c_{\parallel} + c_{\perp}.$$

In this case, $\langle c_{\parallel}, c_{\perp} \rangle = 0$ hence

$$\langle f'_{aj}, c_{\perp} \rangle = \langle c_{\perp}, c_{\perp} \rangle = \|c_{\perp}\|^2,$$

where $\|a\|$ denotes the length of the vector a .

If we take $dx = c_{\perp}$, we conclude (by definition of c_{\perp}) that $dx \perp f'_j$ for all j ; thus, we should conclude that $dx \perp f'_{aj}$, i.e., that $\langle f'_{aj}, dx \rangle = 0$. But for $dx = c_{\perp}$, we know that $\langle f'_{aj}, dx \rangle = \|c_{\perp}\|^2$, so we conclude that the length of the vector c_{\perp} is 0, i.e., that $c_{\perp} = 0$. Thus, the vector f'_{aj} belongs to the linear space generated by the vectors f'_1, \dots, f'_n , i.e.,

$$f'_{aj} = c_{j1}(a) \cdot f'_1 + \dots + c_{jk}(a) \cdot f'_k$$

for some coefficients $c_{jk}(a)$ (which may depend on a). In terms of components, we conclude that

$$f'_j(x_i+a) = c_{j1}(a) \cdot f'_1(x_i) + \dots + c_{jk}(a) \cdot f'_k(x_i).$$

In other words, we conclude that the functions $F_j(x) \stackrel{\text{def}}{=} f'_j(x)$ satisfy the following system of equations:

$$F_j(x+a) = c_{j1}(a) \cdot F_1(x) + \dots + c_{jk}(a) \cdot F_k(x).$$

4°. Since the original poverty functions $f_j(x)$ are twice differentiable, their derivatives $D_j(x) = f'_j(x)$ are differentiable. Let us now prove that the coefficients $c_{jl}(a)$ are differentiable as well.

Let us pick any k values x_1, \dots, x_k , then we have

$$c_{j1}(a) \cdot D_1(x_i) + \dots + c_{jk}(a) \cdot D_k(x_i) = D_j(x_i+a).$$

The coefficients $c_{j1}(a), \dots, c_{jk}(a)$ are now a solution to a system of linear equations. We know that, by using Cramer's rule, we can explicitly (and differentially) describe the solution of the system of linear equation in terms of the coefficients and

of the right-hand side. The coefficients do not depend on a at all, the right-hand side depends on a differentially, so we conclude that $c_{jl}(a)$ are indeed differentiable functions of a .

5°. Since the function $D_j(x+a)$ is a differentiable function of a and the coefficients $c_{jl}(a)$ are differentiable functions of a , we can differentiate both sides of the above equality with respect to a :

$$D'_j(x+a) = c'_{j1}(a) \cdot D_1(x) + \dots + c'_{jk}(a) \cdot D_k(x).$$

Substituting $a = 0$ into this formula, we conclude that

$$D'_j(x) = c_{j1} \cdot D_1(x) + \dots + c_{jk} \cdot D_k(x),$$

where $c_{jl} \stackrel{\text{def}}{=} c'_{jl}(0)$.

Thus, the functions $D_1(x), \dots, D_k(x)$ satisfy a system of linear differential equations with constant coefficients. A general solution to such a system is well known: it is a linear combination of the terms $\exp(\lambda \cdot x)$ with possible complex λ (eigenvalues of the matrix c_{jl}) and terms of the type $x^d \cdot \exp(\lambda \cdot x)$ with natural $d = 1, 2, \dots$ corresponding to multiple eigenvalues; see, e.g., [8] and references therein.

6°. Let us now use a second requirement, that we can also uniquely reconstruct the poverty values under a different change $x_i \rightarrow \lambda \cdot x_i$. In this case, if we add a small value Δx_i to the original value x_i , then we will get

$$f_j(\lambda \cdot (x_i + \Delta x_i)) = f_j(x_i) + \lambda \cdot f'_j(\lambda \cdot x_j) \cdot \Delta x_j.$$

Thus, the value v_j is changed by $\lambda \cdot \sum_i f'_j(\lambda \cdot x_i) \cdot \Delta x_i$.

The requirement that when the original poverty values remain intact, the new poverty values will also remain intact mean that if $\langle f'_j, dx \rangle = 0$ for all j then $\langle f'_{\lambda,j}, dx \rangle = 0$ for all j , where $f'_{\lambda,j}$ denotes a vector with components $f'_j(\lambda \cdot x_i)$.

Similarly to the case of a shift $x'_i = x_i + a$, we therefore conclude that

$$D_j(\lambda \cdot x) = c_{j1}(\lambda) \cdot D_1(x) + \dots + c_{jk}(\lambda) \cdot D_k(x).$$

We can similarly prove that all the coefficients $c_{jl}(\lambda)$ are differentiable functions of λ . Thus, we can differentiate both sides of the above equation and conclude that

$$x \cdot D'_j(\lambda \cdot x) = c'_{j1}(\lambda) \cdot D_1(x) + \dots + c'_{jk}(\lambda) \cdot D_k(x).$$

Substituting $\lambda = 1$ into this formula, we conclude that

$$x \cdot D'_j(x) = c_{j1} \cdot D_1(x) + \dots + c_{jk} \cdot D_k(x),$$

where $c_{jl} \stackrel{\text{def}}{=} c'_{jl}(1)$.

Here, $x \cdot \frac{dD}{dx} = \frac{dD}{dx/x} = \frac{dD}{dX}$, where $X \stackrel{\text{def}}{=} \ln(x)$. In terms of X , we can reconstruct x as $x = \exp(X)$. Thus, if we consider new functions $E_j(X) = D_j(\exp(X))$ (for which $D_j(x) = E_j(\ln(x))$), then these new functions satisfy a system of linear differential equations with constant coefficients:

$$E'_j(X) = c_{j1} \cdot E_1(X) + \dots + c_{jk} \cdot E_k(X).$$

Similarly to the shift part, we conclude that a general solution $E_j(X)$ is a linear combination of functions $X^d \cdot \exp(\lambda \cdot X)$. This means that $D_j(x) = G_j(\ln(x))$ is a linear combination of the functions $(\ln(x))^d \cdot \exp(\lambda \cdot \ln(x)) = x^\lambda \cdot (\ln(x))^d$.

7°. Efficiency means that we must be able to reconstruct the new poverty values both for $x'_i = x_i + a$ and for $x'_i = \lambda \cdot x_i$. Thus, each function $D_j(x)$ must be both a linear combination of the functions $x^d \cdot \exp(\lambda \cdot x)$ ($d = 0, 1, \dots$) and a linear combination of functions $x^\lambda \cdot (\ln(x))^d$. If a function contains exponential terms, it cannot be represented in the second form. Thus, it has to be a linear combination of terms x^d with $d = 0, 1, 2, \dots$, i.e., a polynomial.

8°. We have shown that in the efficient set of poverty functions, all the poverty functions $D_1(x) = f'_1(x), \dots, D_k(x) = f'_k(x)$ from this set are polynomials. Let d be the largest degree of all these polynomials. Then, each of the polynomials $D_j(x)$ is of degree $\leq d$ and thus, any linear combination of these polynomials has a degree $\leq d$.

Let us show that the linear space generated L by these polynomials consists of all the polynomials $D_1(x), \dots, D_k(x)$ up to a degree d .

We have shown that the derivative of each of the functions $D_1(x), \dots, D_k(x)$ belongs to this linear space. For each linear combination of the polynomials $D_j(x)$, its derivative is a linear combination of the derivatives $D'_j(x)$ and therefore, is equal to the linear combination of linear combinations of functions $D_j(x)$. Thus, for every function from the linear space L , its derivative also belongs to the same space L .

By definition of the value d , the linear space L contains a polynomial of degree d , i.e., a polynomial of the type

$$a_d \cdot x^d + a_{d-1} \cdot x^{d-1} + \dots$$

with $a_d \neq 0$. Multiplying this polynomial by the value $\frac{1}{a_d}$, we get a new polynomial which also belongs to the linear space L . This new polynomial has the form

$$x^d + b_{d-1} \cdot x^{d-1} + \dots$$

As we have just shown, the derivative of this polynomial also belongs to the space L ; so, L contains a polynomial

$$d \cdot x^{d-1} + \dots$$

By dividing this new polynomial by d , we conclude that the space L contains a polynomial of the type

$$x^{d-1} + c_{d-2} \cdot x^{d-2} + \dots$$

Similarly, by differentiating this polynomial and dividing it by the coefficient at the highest degree term, we conclude that this linear space contains a function

$$x^{d-2} + \dots,$$

a function

$$x^{d-3} + \dots,$$

etc., all the way to the function 1.

The linear space of all polynomials of degree has dimension $d+1$. We have shown that its subspace L contains $d+1$ linearly independent functions from this space, namely, the functions

$$x^d + \dots, x^{d-1} + \dots, \dots, 1.$$

Thus, the subspace L is also $(d+1)$ -dimensional and thus, it indeed coincides with the space of all the polynomials of degree $\leq d$.

These are linear combinations of functions $D_j(x) = f'_j(x)$. Thus, linear combinations of the original functions $f_j(x)$ (which are integrals of the linear combinations of the functions $D_j(x)$) are all polynomials of degree $\leq (d+1)$, and the linear space L_0 of such linear combinations contains polynomials of the type $x^{d+1} + \dots, x^d + \dots, \dots, x + \dots$ (i.e., $x+c$ for some constant c). In particular, for the function $x+c$, we should be able to reconstruct its after-shift value $\sum_{i=1}^H (x_i + a + c)$ from the before-shift value(s) of this (and other) poverty measures. The difference between the after-shift and before-shift poverty values is equal to

$$\sum_{i=1}^H (x_i + a + c) - \sum_{i=1}^H (x_i + c) = a \cdot \sum_{i=1}^H 1.$$

This difference corresponds to poverty measure $f(x) = 1$; thus, the linear space L_0 contains polynomial 1 as well. So, L_0 is a linear subspace of the $(n+2)$ -dimensional linear space of all polynomials of degree $\leq (d+1)$, and it contains $n+2$ linearly independent functions

$$x^{d+1} + \dots, x^d + \dots, \dots, x + \dots, \text{ and } 1.$$

Thus, the linear space L_0 generated by all the poverty measures $f_j(x)$ indeed coincides with the linear space of all polynomials of degree $\leq (d+1)$. So, we can take the functions $1, x, x^2, \dots, x^{d+1}$ as the new basis – for which, as one can easily see, each new poverty value can be uniquely reconstructed from the old poverty values and vice versa.

The theorem is proven.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation grants HRD-0734825 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by Grant 1 T36 GM078000-01 from the National Institutes of Health.

REFERENCES

- [1] N. Albehery and T. Wang, "Statistical inference of poverty measures using U-statistics approach", *Abstracts of the 8th Joint UTEP/NMSU Workshop on Mathematics, Computer Science, and Computational Sciences*, El Paso, Texas, November 13, 2010.
- [2] N. Albehery and T. Wang, "Statistical inference of poverty measures using U-statistics approach", *International Journal of Intelligent Technologies and Applied Statistics*, 2011, Vol. 4, No. 2, pp. 245–263.
- [3] N. Albehery and T. Wang, "Stochastic dominance of poverty measures and its applications", *Abstracts of the 9th Joint NMSU/UTEP Workshop on Mathematics, Computer Science, and Computational Sciences*, Las Cruces, New Mexico, April 2, 2011.
- [4] N. Albehery and T. Wang, "Statistical inference of poverty measures using U-statistics approach", *Abstracts of the Third International Workshop in Sequential Methodologies*, Stanford, California, June 14–16, 2011.

- [5] N. Albehery and T. Wang, "Estimating poverty measures using truncated distributions", *Abstracts of the 10th Joint UTEP/NMSU Workshop on Mathematics, Computer Science, and Computational Sciences*, El Paso, Texas, November 5, 2011.
- [6] J. Foster, J. Greer, and E. Thorbecke, "A class of decomposable poverty measures", *Econometrica*, 1984, Vol. 81, pp. 761–766.
- [7] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
- [8] H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.
- [9] H. T. Nguyen and E. A. Walker, *First Course In Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.
- [10] L. A. Zadeh, "Fuzzy sets", *Information and control*, 1965, Vol. 8, pp. 338–353.