INFORMETRICS 89/90
L. Egghe and R. Rousseau (Editors)



# ELEMENTS OF CONCENTRATION THEORY

Leo EGGHE

LUC, Universitaire Campus, B-3610 Diepenbeek, Belgium (*)
UIA, Universiteitsplein 1, B-2610 Wilrijk, Belgium

Ronald ROUSSEAU

Katholieke Industriële Hogeschool West-Vlaanderen,
Zeedijk 101, B-8400 Oostende, Belgium
UIA, Documentatie- en Bibliotheekwetenschap,
Universiteitsplein 1, B-2610 Wilrijk, Belgium

Abstract

We review some concentration measures proposed in the literature and present a set of principles that good concentration measures must fulfill. We moreover look into some of the consequences of these principles.
The transfer principle is extended to yield a new family of principles, denoted E(p), but a concentration measure can only satisfy E(p) for at most one p. We discuss briefly the issue of sensitivity to transfers and show that Heine's dispersion measures are related to some well-known concentration measures.

## 1. INTRODUCTION

Concentration theory or the measurement of inequality deals with rankings of distributions. In economics and sociology people use this theory to answer questions like the following. Is the distribution of income in country X now more equal than it was in the past? Are third world countries characterized by greater inequality than western countries? Do taxes lead to greater equality in the distribution of wealth? [1]

In informetrics and linguistics one can ask similar questions. Is the distribution of journals that publish papers on physics more concentrated than the distribution of journals dealing with sociology? (cf. the Bradford distribution). Are citations to mathematics papers more unequally distributed than citations to chemical papers? Give a number to characterize the difference in word occurrence between the works of Mark Twain and the works of James Joyce. Are borrowings in a university library more unequally distributed over the available books than in a public library?

This list of questions and areas of application for a general theory of concentration is by no means exhaustive. In general we will use the terminology of sources and items. In the examples taken from sociology the sources are the income classes, the items the people belonging to these classes; according to the point of view taken by the investigator sources might also be people, in which case an item is a certain amount of money.

---

(*) Permanent address.

For the Bradford distribution the sources are journals and the items are papers, but in the case of citation analysis, the sources are papers and the items citations. In the linguistic example sources are words and items occurrences of words. Finally, for the library example, sources are books and items are borrowings. We recall that especially in linguistic studies, one often uses the terms types and tokens instead of sources and items. We will denote the number of items in the i-th source by $x_i$. The total number of sources is denoted $N$, where $N > 1$, and the relative number of items in the i-th source, i.e. $\frac{x_i}{\sum_{k=1}^{N} x_k}$, is denoted $a_i$.

To develop a theory of concentration one has to find a mathematical formulation for the intuitive meaning of "concentration" or "inequality". This means that we will study functions f of N variables : $f(x_1,\ldots,x_N)$. In the spirit of [2] we will develop a set of principles such a function must satisfy in order to be an acceptable measure of concentration. Let us already mention two obvious requirements.

(a) If there is a perfect concentration, i.e. all $x_i = 0$ except one, then $f(x_1,\ldots,x_N)$ attains its maximal value, given a fixed value for $\sum_{i=1}^{N} x_i$.

(b) If all $x_i$ are equal (and different from zero) so that there is a perfect equality, $f(x_1,\ldots,x_N)$ must be zero.

Other natural, but more intricate principles will be investigated further on.

Equivalent to the problem of finding good concentration measures, there is the problem of finding good dispersion measures. Indeed, a dispersion measure can be viewed as the opposite of a concentration measure : it measures the way in which a certain distribution is not concentrated, i.e. is dispersed. If $g(x_1,\ldots,x_N)$ denotes a measure of dispersion, principles (a) and (b) become the following :

(a') If we have no dispersion (i.e. perfect concentration) then $g(x_1,\ldots,x_N) = 0$.

(b') If we have perfect dispersion (i.e. no concentration) then $g(x_1,\ldots,x_N)$ attains its maximal value, given a fixed total of $\sum_{i=1}^{N} x_i$ items.

If the maximal value in condition (a) is taken to be 1 then a first suggestion for a dispersion measure g is to put $g = 1 - f$. This suggestion and related ones will be studied in a more general context.

In the next section we will review some concentration measures, which were proposed in the literature. Section three gives the relation between some of these measures and the discrete Lorenz curve; we will also show that several of these proposed concentration measures are of the form : a measure of deviation (or scatter) about the arithmetic mean $\mu$ divided by $\mu$. In section four we propose a number of principles, i.e. requirements that good concentration measures must fulfill. We also check which measures satisfy these principles. Sections five and six deal with related or equivalent requirements. Section seven considers the important issue of sensitivity to transfers and finally, section eight considers dispersion measures and their relation with concentration measures.

2. CONCENTRATION MEASURES

2.1. It is intuitively clear that the classical notions of standard deviation ($\sigma$) and variance ($\sigma^2$), where

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2 = \frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2$$

$$= \frac{1}{2N^2}\sum_{k=1}^{N}\sum_{\ell=1}^{N}(x_k - x_\ell)^2 \qquad (1)$$

(and $\mu$ denotes the mean of the distribution), bear some relation on the notion of concentration. But, as we will show, they do not satisfy essential conditions for good concentration measures.
The above equalities in (1) are readily seen and are well-known.

2.2. The coefficient of variation

$$V = \frac{\sigma}{\mu} \qquad (2)$$

was introduced to deal with relative values instead of absolute ones (cf. the argumentation in [3], p.164 and following). In a similar way we can consider

$$V^2 = \frac{\sigma^2}{\mu^2} \qquad (3)$$

or Gaston's measure ([4], p.148)

$$Ga = \frac{\sigma^2}{\mu} \qquad (4)$$

or Allison's modified squared variation coefficient [3]

$$A = \frac{\sigma^2 - \mu}{\mu^2} \qquad (5)$$

2.3. From linguistics we consider the Yule characteristic, defined as

$$K = \frac{\sigma^2}{\mu^2 N} = \frac{V^2}{N} \qquad (6)$$

In [5] Johnson advocates the use of Simpson's index [6] in stylistic studies.

$$J = \frac{\sum_{i=1}^{n} i(i-1)\, x_i}{n(n-1)} \qquad (7)$$

where $x_i$ is the number of words that occur i times and n is the total number of words that occur in the text under study. Simpson's index is nothing but the number of identical pairs divided by the number of all possible pairs.

2.4. The Schutz coefficient (relative mean deviation) [7]

$$D = \frac{\frac{1}{N}\sum_{i=1}^{N} |x_i - \mu|}{2\mu} \qquad (8)$$

According to Gastwirth [8], this measure was first proposed by Yntema [9] and Pietra in the 1930's.

2.5. Pratt's measure and the Gini index.

In order to define Pratt's measure we first assume that the $x_i$'s (hence also the $a_i$'s) are ordered decreasingly. Putting

$$q = \sum_{i=1}^{N} i\, a_i \, , \qquad (9)$$

Pratt's measure C is defined as ([10]) :

$$C = \frac{2\left(\frac{N+1}{2} - q\right)}{N-1} \; ; \qquad (10)$$

Gini's index is then

$$G = \frac{N-1}{N} C \, . \qquad (11)$$

It should be recalled however that the Gini index was introduced in econometrics [11] long before Pratt's measure was defined. Relation (11) was established in 1979 by Carpenter [12]. The usual definition of Gini's index uses the so-called Lorenz curve (see further). Using only absolute frequencies ($x_i$'s), Pratt's measure can be rewritten as :

$$C = \frac{(N+1)\sum_{i=1}^{N} x_i - 2\sum_{i=1}^{N} i\, x_i}{(N-1)\left(\sum_{i=1}^{N} x_i\right)} \qquad (12)$$

or also as :

$$C = \frac{N+1}{N-1} - \frac{2}{\mu N(N-1)} \sum_{i=1}^{N} i\, x_i \, . \qquad (13)$$

We also propose the following generalized Pratt measure (see section three for an explanation of this terminology) :

$$P(r) = \frac{\left(\frac{1}{2N(N-1)} \sum_{i=1}^{N}\sum_{j=1}^{N} |x_i - x_j|^r\right)^{1/r}}{\mu} \, , \; r > 0 \qquad (14)$$

2.6. Theil's measure [13].

This inequality measure is defined as :

$$Th = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{x_i}{\mu}\right) \ln \left(\frac{x_i}{\mu}\right) \qquad (15)$$

(cf. the notion of entropy in information theory).

We further remark that in this formula one sets $0.\ln(0) = 0$.

2.7. The variance of the logarithm.

$$L = \frac{1}{N}\sum_{i=1}^{N}\left(\ln(x_i) - \sum_{j=1}^{N}\frac{\ln(x_j)}{N}\right)^2 ,$$

$$= \frac{1}{2N^2}\sum_{k=1}^{N}\sum_{\ell=1}^{N}(\ln x_k - \ln x_\ell)^2 \tag{16}$$

which is only defined if all $x_i \neq 0$.

2.8. Atkinson's index.
In [1] Atkinson introduced a family of concentration measures defined as :

$$A(e) = 1 - \left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\mu}\right)^{1-e}\right)^{\frac{1}{1-e}} , \tag{17}$$

where $e > 0$ and $e \neq 1$.
If all $x_i \neq 0$, $A(1)$ is defined as $\lim_{e\to1} A(e)$, which is nothing but

$$\frac{\mu - GM(x_i)_i}{\mu} \tag{18}$$

as is easily seen.
Here $GM(x_i)_i$ denotes the geometric mean of the $x_i$, $i = 1,\ldots,N$. Remark that $\lim_{e\to0} A(e) = 0$. The formula for the Atkinson index as presented in ([2], p.873) seems to be in error.

2.9. The CON-index [14].
In the authors' own words, this index is the standard deviation of the percentage shares divided by the maximum possible standard deviation in a system of size N.
This yields :

$$CON = \sqrt{\frac{\sum_{i=1}^{N} a_i^2 - \frac{1}{N}}{1 - \frac{1}{N}}} . \tag{19}$$

Since $a_i = \frac{x_i}{\mu N}$, this formula can be rewritten as follows :

$$CON = \frac{1}{\mu}\sqrt{\frac{\sum_{i=1}^{N} x_i^2 - \mu^2 N}{N(N-1)}} = \frac{1}{\mu}\frac{\sigma}{\sqrt{N-1}} = \frac{V}{\sqrt{N-1}} \tag{20}$$

(using (1) and (2)).

Formula (20) shows that CON is only a variant of the coefficient of variation.

2.10. Lotka's $\alpha$.
We finally remark that Rao [15] pointed out that when data follow Lotka's distribution :

$$f(y) = \frac{A}{y^{\alpha}} \tag{21}$$

the exponent $\alpha$ could be used as a measure of concentration. We recall that in (21) $f(y)$ denotes the number of sources with $y$ items.

## 3. THE EXPRESSION 'MEASURE OF DEVIATION' DIVIDED BY THE MEAN. THE LORENZ DISTRIBUTION

Many concentration measures are of the form : a measure of deviation about the mean divided by the mean. This is obvious for the coefficient of variation, Schutz' coefficient and Theil's measure. Here we will show that other measures are also of this form.

### 3.1. Proposition

The Gini index G is equal to

$$\frac{\frac{1}{2N^2}\sum_{k=1}^{N}\sum_{\ell=1}^{N}|x_k - x_\ell|}{\mu} \tag{22}$$

showing that G is also of the special form we are investigating here.

*Proof :*

By (9), (10) and (11)

$$G = \frac{N-1}{N}C = \frac{2\left(\frac{N+1}{2} - \sum_{k=1}^{N} k\, a_k\right)}{N}$$

$$= \frac{\frac{2}{N^2}\left[\frac{N+1}{2}\sum_{k=1}^{N} x_k - \sum_{k=1}^{N} k\, x_k\right]}{\frac{\left(\sum_{k=1}^{N} x_k\right)}{N}}$$

$$= \frac{\frac{1}{N^2}\sum_{k=1}^{N}((N+1) - 2k)\, x_k}{\mu} \qquad (*)$$

Now, $\sum_{k=1}^{N}\sum_{\ell=1}^{N}|x_k - x_\ell| = 2\sum_{k<\ell}(x_k - x_\ell)$ (as $x_k \geq x_\ell$ for $k < \ell$)

$$= 2\left[\sum_{k=1}^{N}((N+1)x_k - 2k\, x_k\right].$$

Substituting this in (*) above gives :

$$G = \frac{\frac{1}{2N^2} \sum_{k=1}^{N} \sum_{\ell=1}^{N} |x_k - x_\ell|}{\mu} .$$

3.2. Remark

The factor $\frac{1}{N^2} \sum_{k=1}^{N} \sum_{\ell=1}^{N} |x_k - x_\ell|$ is known in the literature as the Gini mean difference. It has recently been used [16] as the basis for a measure of association which generalizes the Pearson product-moment correlation coefficient, Kendall's tau and Spearman's rank correlation coefficient.

3.3. Corollary

$$\text{Pratt's measure } C = \frac{\frac{1}{2N(N-1)} \sum_{k=1}^{N} \sum_{\ell=1}^{N} |x_k - x_\ell|}{\mu} . \qquad (23)$$

We remark that formulae (22) and (23) are independent of the special ranking we have used to define G and C.

Proposition 3.1, Corollary 3.3 and the fact that

$$V = \frac{\sigma}{\mu} = \frac{\left(\frac{1}{2N^2} \sum_{k=1}^{N} \sum_{\ell=1}^{N} (x_k - x_\ell)^2\right)^{1/2}}{\mu} \qquad \text{(using (1))}$$

suggest to introduce a generalized Pratt measure (in the same way one could also consider a generalized Gini index) as :

$$P(r) = \frac{\left(\frac{1}{2N(N-1)} \sum_{k=1}^{N} \sum_{\ell=1}^{N} |x_k - x_\ell|^r\right)^{1/r}}{\mu} , \; r > 0$$

(cf. [2], p.870). This gives a whole family of concentration measures depending on the parameter r. For r = 1, P(1) = C; for r = 2, $P(2) = (\frac{N}{N-1})^{1/2} V$.

The next result shows that although A(e) has a somewhat similar form it is more a measure of skewness divided by the mean.

3.4. Proposition

A(e) has the form : arithmetic mean minus the generalized (1-e)-mean, divided by the arithmetic mean. Here the generalized (1-e)-mean is given by

$$\left(\frac{1}{N} \sum_{i=1}^{N} x_i^{1-e}\right)^{1/(1-e)} \qquad (24)$$

*Proof :*

$$A(e) = 1 - \left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{x_i}{\mu}\right)^{1-e}\right)^{\frac{1}{1-e}}$$

$$= 1 - \frac{1}{\mu}\left(\frac{1}{N}\sum_{i=1}^{N} x_i^{1-e}\right)^{\frac{1}{1-e}}$$

$$= \frac{\mu - \left(\frac{1}{N}\sum_{i=1}^{N} x_i^{1-e}\right)^{\frac{1}{1-e}}}{\mu} .$$

Remark that also A(1) has a similar form.

In the sequel of this section we will investigate another relation, namely that between the discrete Lorenz distribution and some of the proposed concentration measures.

### 3.5. The discrete Lorenz distribution

If there are N classes (sources) and there is perfect equality, then every class contains $\frac{1}{N}$ of the total number of items. The cumulative relative frequency distribution of this situation is called the discrete Lorenz distribution of equality and is given by the diagonal points of Fig.1.

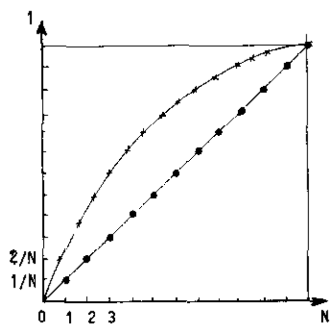


Fig.1

When data points $(x_i)_{i=1,\ldots,N}$ are ordered decreasingly we can also consider the cumulative relative frequency of the $x_i$ (denoted by * in Fig.1). This *distribution is then* the discrete Lorenz distribution of the data. The sum of the differences between these two discrete distributions is then :

$$(a_1 - \frac{1}{N}) + (a_1 + a_2 - \frac{2}{N}) + \ldots + (\sum_{j=1}^{i} a_j - \frac{i}{N}) + \ldots + (1 - \frac{N}{N})$$

$$= N\, a_1 + (N-1)\, a_2 + \ldots + (N-(i-1))\, a_i + \ldots + 1\, a_N - \frac{\sum_{i=1}^{N} i}{N}$$

$$= ((N+1)-1)\, a_1 + \dots + ((N+1)-i)\, a_i + \dots + ((N+1)-N)\, a_N - \frac{N+1}{2}$$

$$= (N+1) \sum_{i=1}^{N} a_i - \sum_{i=1}^{N} i\, a_i - \frac{N+1}{2}$$

$$= \frac{N+1}{2} - q \qquad (*)$$

The greatest difference between the Lorenz distribution of equality and an observed one is obtained when $a_1 = 1$, $a_2 = \dots = a_N = 0$. Then the difference equals

$$\sum_{i=1}^{N} (1 - \frac{i}{N}) = N - \frac{N+1}{2} = \frac{N-1}{2} \quad .$$

Normalizing (*) to obtain a value which is always in the interval [0,1] yields

$$\frac{2\,(\frac{N+1}{2} - q)}{N-1} \quad .$$

This shows the following proposition.

3.6. Proposition

Pratt's measure C is the normalized sum of the differences between the observed Lorenz distribution and the discrete Lorenz distribution of equality.

Also Schutz' coefficient can be related to the Lorenz distribution.

3.7. Proposition (cf. [8], Lemma 3)

Schutz' coefficient is equal to the maximal difference between the observed discrete Lorenz distribution and the Lorenz distribution of equality.

*Proof :*

$$D = \frac{(\sum_{i=1}^{N} |x_i - \mu|)/N}{(2 \sum_{j=1}^{N} x_j)/N} \, ,$$

which should be equal to

$$\max_j \left( \sum_{i=1}^{j} \frac{x_i}{\sum_{k=1}^{N} x_k} - \frac{j}{N} \right) \quad .$$

So, we have to show that

$$\max_j \left[ 2\, (\sum_{i=1}^{j} x_i - j\mu) \right] = \sum_{k=1}^{N} |x_k - \mu| \quad .$$

To show this, we first remark that for every $k \in \{1,2,\ldots,N\}$

$$\sum_{i=1}^{k} (x_i - \mu) = - \sum_{i=k+1}^{N} (x_i - \mu) . \qquad (*)$$

Choose now $j \in \{1,\ldots,N\}$ such that $x_i \geq \mu$ if $i \leq j$ and $x_i < \mu$ if $i > j$. Such a $j$ always exists, if the $x_i$'s are ordered decreasingly (which can always be achieved via a permutation and a relabeling). Then

$$2 \left( \sum_{i=1}^{j} x_i - j\mu \right)$$

$$= 2 \sum_{i=1}^{j} (x_i - \mu)$$

$$= \sum_{i=1}^{j} (x_i - \mu) - \sum_{i=j+1}^{N} (x_i - \mu) \qquad \text{(by (*))}$$

$$= \sum_{i=1}^{j} |x_i - \mu| + \sum_{i=j+1}^{N} |x_i - \mu| \qquad \text{(definition of } j\text{)}$$

$$= \sum_{i=1}^{N} |x_i - \mu| .$$

Hence, the above $\max_j$ is certainly not smaller than $\sum_{k=1}^{N} |x_k - \mu|$. But, using (*) again we have also, for every $i \in \{1,\ldots,N\}$ :

$$\left| 2 \left( \sum_{k=1}^{i} x_k - i\mu \right) \right|$$

$$= \left| \sum_{k=1}^{i} (x_k - \mu) - \sum_{k=i+1}^{N} (x_k - \mu) \right| \leq \sum_{k=1}^{N} |x_k - \mu| .$$

Hence, we have proved that

$$\max_j \left( 2 \left( \sum_{i=1}^{j} x_i - j\mu \right) \right) = \sum_{k=1}^{N} |x_k - \mu| .$$

## 4. CONCENTRATION PRINCIPLES

In this section $f(x_1,\ldots,x_N)$ denotes a general concentration measure.

4.1. (C1) If all $x_i$ are equal, say to $c \neq 0$, then $f(x_1,\ldots,x_N)$ attains its minimal value, equal to 0.

This is a perfectly natural condition, as already explained in the introductory section. Remark also that (C1) implies that a concentration measure is never negative. All measures considered in Section 2 obviously satisfy this condition (Even Lotka's $\alpha$ is then zero). Indeed it is in order to satisfy C1 that one uses Pratt's measure C (formula (10)) instead of q (formula (9)).

4.2. (C2) For every $(x_1,\dots,x_N)$ and every permutation $\pi : \{1,\dots,N\} \to \{1,\dots,N\}$ we require that $f(x_1,\dots,x_N) = f(x_{\pi(1)},\dots,x_{\pi(N)})$.

This principle expresses that e.g. poverty (or richness) of a nation is not a labelled property : it is only determined by the overall configuration. Also this principle is satisfied by all measures considered in Section 2, as is readily seen.

4.3. Scale invariance : (C3)

This principle says that for every $(x_1,\dots,x_N)$ and $c > 0$ :

$$f(cx_1,\dots,cx_N) = f(x_1,\dots,x_N) \quad .$$

It expresses the requirement that a good concentration measure should not be influenced by the units. Returning to the case of income distributions this means that there must not be a difference whether we calculate the income in dollars, yen or rupees.

The standard deviation ($\sigma$) and the variance ($\sigma^2$) do not satisfy this important requirement, neither do Gaston's measure (Ga), Allison's modified squared coefficient of variation (A) and Simpson's index (J). From now on these measures will not be considered anymore.

However, concerning Allison's modified squared coefficient of variation, we should point out that Allison was fully aware of the fact that it is not scale invariant for the observed data (in his case, these data were publication numbers). He shows however that A is scale invariant for the underlying latent rate of publication ([3], [17]).

4.4. (C4) "When the richest source gets richer, inequality rises".

This principle is a very natural one : it has two requirements : the first is the one mentioned above and the second is its dual : when the poorest source gets poorer then too inequality increases. In a mathematical formulation this becomes the following.

(C4a) If $x_i = \max\{x_1,\dots,x_N\}$ and if there exists a $k \neq i$ such that $x_k \neq 0$, then, for $h > 0$,

$$f(x_1,\dots,x_i+h,\dots,x_N) > f(x_1,\dots,x_N) \; .$$

(C4b) If $x_j = \min\{x_1,\dots,x_N\}$ and $0 < h \leq x_j$ then

$$f(x_1,\dots,x_j-h,\dots,x_N) > f(x_1,\dots,x_N) \; .$$

The principles (C4a) and (C4b) can be expressed in a different way as shown in the next propositions.

4.5. Propositions

A. (C4a) is equivalent with (C4a') :
(C4a'). If $x_i = \max\{x_1,\ldots,x_N\}$ and if there exists a $k \neq i$ such that $x_k \neq 0$, then, for $h > 0$, $f(x_1,\ldots,x_i-h,\ldots,x_N) < f(x_1,\ldots,x_N)$ as long as $x_i-h \geq x_t$, for every $t \neq i$.

B. (C4b) is equivalent with (C4b') :
(C4b'). If $x_j = \min\{x_1,\ldots,x_N\}$ and $0 < h$ then $f(x_1,\ldots,x_j+h,\ldots,x_N) < f(x_1,\ldots,x_N)$ as long as $x_j+h \leq x_t$, for every $t \neq j$.

*Proof :*

A. The implication (C4A) → (C4a') is trivial.
Suppose now that (C4a') is satisfied, then

$$f(x_1,\ldots,x_i,\ldots,x_N) = f(x_1,\ldots,(x_i+h)-h,\ldots,x_N)$$
$$< f(x_1,\ldots,x_i+h,\ldots,x_N) \ ,$$

as $x_i = (x_i+h) - h \geq x_t$ for every $t \neq i$. This shows the implication (C4a') → (C4a).

The proof of part B is similar and is left to the reader.

Lotka's $\alpha$ does not satisfy C4a or C4b for a change in one source destroyes the Lotka distribution.

We will now check whether the remaining concentration measures satisfy this princople. For convenience we only check (C4a); (C4b) can be dealt with in a similar way.

4.6. Proposition $V^2$ (hence also V, K and the CON-index) satisfy (C4a).

*Proof :*

*If* $x_i = \max_j (x_j)$, *then we essentially have to show that*

$$\frac{(x_i+h)^2 + \sum_{j\neq i} x_j^2}{(\mu + \frac{h}{N})^2} > \frac{\sum_j x_j^2}{\mu^2}$$

This is equivalent with the requirement that

$$\mu^2N^2 \left(\sum_j x_j^2 + 2\, x_i h + h^2\right) > (\mu^2N^2 + 2\,\mu hN + h^2).\left(\sum_j x_j^2\right)$$

or, using the fact that $\sum_j x_j^2 < \mu^2N^2$, that

$$2\,\mu^2N^2x_ih + \mu^2N^2h^2 \geq 2\,\mu hN.\left(\sum_j x_j^2\right) + h^2\mu^2N^2 \ .$$

As $\mu N x_i = (\sum_j x_j).x_i \geq \sum_j x_j^2$, this proves that $V^2$ satisfies (C4a).

4.7. Proposition. Schutz' coefficient, Pratt's measure and the Gini index satisfy (C4a).

*Proof :*

This follows immediately from their interpretation using the discrete Lorenz curve (Section 3). Indeed, adding h to $x_i$ yields a cumulative relative distribution which lies strictly above the original one. Hence Schutz' coefficient, Pratt's measure and the Gini index increase.

4.8. Proposition. Theil's measure satisfies (C4a).

*Proof :*

We consider the function

$$f(h) = \sum_{k \neq i} \left(\frac{x_k}{\mu + \frac{h}{N}}\right) \ln \left(\frac{x_k}{\mu + \frac{h}{N}}\right) + \frac{x_i + h}{\mu + \frac{h}{N}} \ln \left(\frac{x_i + h}{\mu + \frac{h}{N}}\right)$$

and show it to be increasing in h. For this we calculate f'(h) and show that f'(h) > 0.

$$f'(h) = \sum_{k \neq i} \frac{-x_k N}{(N\mu + h)^2} \ln \left(\frac{Nx_k}{\mu N + h}\right) + \sum_{k \neq i} \frac{Nx_k}{N\mu + h} \cdot \frac{N\mu + h}{Nx_k} \cdot \frac{-Nx_k}{(N\mu + h)^2}$$

$$+ \frac{N(N\mu + h) - N(x_i + h)}{(N\mu + h)^2} \ln \left(\frac{x_i + h}{\mu + \frac{h}{N}}\right) + \frac{N(x_i + h)}{N\mu + h} \cdot \frac{N\mu + h}{N(x_i + h)} \cdot \frac{N(N\mu + h) - N(x_i + h)}{(N\mu + h)^2}.$$

So, we have to show that

$$(N\mu - x_i) \left[1 + \ln \left(\frac{N(x_i + h)}{\mu N + h}\right)\right] > \sum_{k \neq i} x_k \left[1 + \ln \left(\frac{Nx_k}{\mu N + h}\right)\right] .$$

As $\ln (N(x_i + h)) > \ln (Nx_k)$, for every $k \neq i$, this inequality is obviously satisfied.

4.9. Proposition. The variance of logarithms, L, satisfies (C4a).

*Proof :*

$$L = \frac{1}{2N^2} \sum_{k=1}^{N} \sum_{\ell=1}^{N} (\ln x_k - \ln x_\ell)^2 .$$

When $x_i$ is replaced by $x_i + h$ all terms in this sum stay the same or increase, hence L increases.

4.10. Proposition. Atkinson's index A(e) satisfies (C4a).

*Proof :*

Let $e \neq 1$, then we consider the function

$$f(h) = \frac{(\sum_{j\neq i} x_j^{1-e} + (x_i+h)^{1-e})^{1/(1-e)}}{N^{1/(1-e)}(\mu + \frac{h}{N})}$$

and show it to be decreasing for $h \geq 0$.
Taking the derivative yields :

$$f'(h) = \frac{T_1 - T_2}{N^{1/(1-e)}(\mu + \frac{h}{N})^2}$$

with

$$T_1 = (\sum_{j\neq i} x_j^{1-e} + (x_i+h)^{1-e})^{\frac{e}{1-e}} . (x_i+h)^{-e} (\mu + \frac{h}{N})$$

and

$$T_2 = (\sum_{j\neq i} x_j^{1-e} + (x_i+h)^{1-e})^{\frac{1}{1-e}} . \frac{1}{N} .$$

This means we have to show that

$$(\sum_{j\neq i} x_j^{1-e} + (x_i+h)^{1-e})^{-1}(x_i+h)^{-e}(\mu + \frac{h}{N}) - \frac{1}{N} < 0$$

or

$$\mu N + h < (\sum_{j\neq i} x_j^{1-e} + (x_i+h)^{1-e}).(x_i+h)^{e} .$$

Now,

$$\mu N + h = (\sum_j x_j) + h = \sum_{j\neq i} x_j^{1-e}.x_j^{e} + (x_i+h)$$

$$< (x_i+h)^{e} . \sum_{j\neq i} x_j^{1-e} + (x_i+h) \qquad \text{(as } e > 0\text{)}$$

$$= (\sum_{j\neq i} x_j^{1-e} + (x_i+h)^{1-e}).(x_i+h)^{e} .$$

This proves Proposition 4.10 if $e \neq 1$.
For the case $e = 1$, we consider the function

$$g(h) = \frac{(x_1 \dots (x_i+h) \dots x_N)^{1/N}}{\mu + \frac{h}{N}} = \frac{(\prod_j x_j + h \prod_{j\neq i} x_j)^{1/N}}{\mu + \frac{h}{N}} .$$

Then,

$$g'(h) = \frac{\frac{1}{N} (\prod_j x_j + h \prod_{j\neq i} x_j)^{\frac{1}{N}-1} . \prod_{j\neq i} x_j.(\mu + \frac{h}{N}) - (\prod_j x_j + h \prod_{j\neq i} x_j)^{\frac{1}{N}} . \frac{1}{N}}{(\mu + \frac{h}{N})^2} .$$

Now, we have to show that

$$\left(\prod_j x_j + h \prod_{j\neq i} x_j\right)^{-1} . \prod_{j\neq i} x_j . \left(\mu+\frac{h}{N}\right) - 1 < 0$$

or

$$N\mu . \prod_{j\neq i} x_j + h . \prod_{j\neq i} x_j < N \prod_j x_j + hN \prod_{j\neq i} x_j .$$

As $\mu \leq x_i$ and $N > 1$, this is obvious.

This finishes the proof of Proposition 4.10.

Finally, we have to show that the generalized Pratt measure satisfies(C4a). As the proof we will present is rather long we postpone it to the appendix. Here we only state the proposition.

4.11. Proposition. The generalized Pratt measure, P(r) satisfies(C4a), for $r \geq 1$.

An important consequence of (C4a) is that if f satisfies (C4a), f depends explicitly on all sources. Indeed, if f would be independent of the i-th source, one could add items to it untill it becomes the richest (without changing the value of f). An additional increase in items would still yield the same value for f, but this would contradict principle (C4a).

4.12. (C5) The principle of nominal increase.
This principle requires that an equal, nominal increase in each source strictly decreases the global inequality. More formally this becomes : for every $(x_1,\ldots,x_N)$, where not all $x_i$ are equal, and $h > 0$ :
$f(x_1+h,\ldots,x_N+h) < f(x_1,\ldots,x_N)$.

The principle of nominal increase is equivalent with the requirement that the function $F(h) = f(x_1+h,\ldots,x_N+h)$ is strictly decreasing on $[0,+\infty[$.

Indeed, if an inequality measure satisfies this requirement then :

$$\forall h > 0 : f(x_1+h,\ldots,x_N+h) < f(x_1,\ldots,x_N) .$$

Conversely, if $h_1 < h_2$ and if f satisfies the principle of nominal increase, then

$$f(x_1+h_2,\ldots,x_N+h_2)$$
$$= f(x_1+h_1+(h_2-h_1),\ldots,x_N+h_1+(h_2-h_1)) < f(x_1+h_1,\ldots,x_N+h_1) .$$

It is easy to see that V, $V^2$, K, CON and D satisfy (C5). Adding a fixed value to each class again destroyes a Lotka distribution, so that Lotka's $\alpha$ does not satisfy this principle. Consequently, this measure will not be considered anymore in the following.

The next propositions show that P(r) (and hence also C and G and once more V), Th, L and A(e) satisfy(C5).

4.13. Proposition. The generalized Pratt measure satisfies the principle of nominal increase.

*Proof :*

As

$$P(r) = \frac{\left(\frac{1}{2N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} |x_i - x_j|^r\right)^{1/r}}{\mu}$$

we see that adding h to each $x_k$ does not change the nominator. The denominator on the other hand, increases from $\mu$ to $\mu + h$. This shows that P(r) satisfies (C5).

4.14. Proposition. Theil's measure satisfies the principle of nominal increase.

*Proof :*

It suffices to show that the function

$$F(h) = \sum_{i=1}^{N} \frac{x_i + h}{\mu + h} \ln \left(\frac{x_i + h}{\mu + h}\right)$$

is decreasing in h. Hence we will show that $F'(h) < 0$, for every $h > 0$.

$$F'(h) = \sum_{i=1}^{N} \frac{\mu - x_i}{(\mu + h)^2} \left(\ln \left(\frac{x_i + h}{\mu + h}\right) + 1\right) .$$

We suppose now that not all $x_i$ are equal and that they are ordered increasingly. Let j be the largest index in {1,...,N} such that $x_j \le \mu$ (so $j < N$). Since $\sum_{i=1}^{N} (\mu - x_i) = 0$ we have :

$$\sum_{i=1}^{j} (\mu - x_i) = - \sum_{i=j+1}^{N} (\mu - x_i) . \qquad (*)$$

But

$$(\mu + h)^2 . F'(h) = \sum_{i=1}^{j} (\mu - x_i)\left(\ln \left(\frac{x_i + h}{\mu + h}\right) + 1\right) + \sum_{i=j+1}^{N} (\mu - x_i)\left(\ln \left(\frac{x_i + h}{\mu + h}\right) + 1\right) .$$

Hence, by the definition of j, we have the following majorization :

$$(\mu + h)^2 . F'(h) \le \sum_{i=1}^{j} (\mu - x_i)\left(\ln \left(\frac{x_j + h}{\mu + h}\right) + 1\right) + \sum_{i=j+1}^{N} (\mu - x_i)\left(\ln \left(\frac{x_{j+1} + h}{\mu + h}\right) + 1\right)$$

with equality only if $x_1 = \dots = x_j < x_{j+1} = \dots = x_N$.

Using (*) this gives :

$$(\mu+h)^2 F'(h) \le \sum_{i=1}^{j} (\mu-x_i)(\ln(\frac{x_j+h}{\mu+h}) - \ln(\frac{x_{j+1}+h}{\mu+h})) < 0 ,$$
for every $h > 0$.

This finishes the proof of the proposition.

4.15. Proposition. The variance of logarithms satisfies the principle of nominal increase.

*Proof :*

We have to show that :

$$\frac{1}{N}\sum_{i=1}^{N} (\ln(x_i+h) - \sum_{j=1}^{N} \frac{\ln(x_j+h)}{N})^2$$

$$< \frac{1}{N}\sum_{i=1}^{N} (\ln(x_i) - \sum_{j=1}^{N} \frac{\ln x_j}{N})^2 , \quad h > 0 ,$$

or

$$\sum_{i=1}^{N}\sum_{j=1}^{N} (\ln(x_i+h) - \ln(x_j+h))^2$$

$$< \sum_{i=1}^{N}\sum_{j=1}^{N} (\ln x_i - \ln x_j)^2 , \quad h > 0 .$$

This inequality is obviously satisfied as ln is an increasing concave function. Finally also Atkinson's index satisfies the principle of nominal increase.

4.16. Proposition. $A(e)$, $e > 0$, satisfies the principle of nominal increase.

*Proof :*

We will show that the function

$$F(h) = (\frac{1}{N}\sum_{i=1}^{N} (\frac{x_i+h}{\mu+h})^{1-e})^{\frac{1}{1-e}} = \frac{(\frac{1}{N}\sum_{i=1}^{N} (x_i+h)^{1-e})^{\frac{1}{1-e}}}{\frac{1}{N}\sum_{i=1}^{N} (x_i+h)}$$

is increasing in $h$ ($h \ge 0$). This will prove the proposition for $e \ne 1$.

Calculating $F'(h)$ yields :

$$F'(h) = \frac{N}{N^{\frac{1}{1-e}}} \frac{T_1 - T_2}{(\sum_{i=1}^{N} (x_i+h))^2}$$

with

$$T_1 = \frac{1}{1-e}\left(\sum_{i=1}^{N}(x_i+h)^{1-e}\right)^{\frac{e}{1-e}}.\left(\sum_{i=1}^{N}(1-e)(x_i+h)^{-e}\right).\left(\sum_{i=1}^{N}(x_i+h)\right)$$

and

$$T_2 = \left(\sum_{i=1}^{N}(x_i+h)^{1-e}\right)^{\frac{1}{1-e}}.N$$

Then :

$$F'(h) = \frac{N^{-\frac{e}{1-e}}\left(\sum_{i=1}^{N}(x_i+h)^{1-e}\right)^{\frac{1}{1-e}}}{\left(\sum_{i=1}^{N}(x_i+h)\right)^2}\cdot\left[\frac{\sum_{i=1}^{N}(x_i+h)^{-e}.\sum_{i=1}^{N}(x_i+h)}{\sum_{i=1}^{N}(x_i+h)^{1-e}} - N\right].$$

Now, to prove that $F(h)$ is increasing, we have to show that $F'(h) > 0$, or :

$$\sum_{i=1}^{N}(x_i+h)^{-e}.\sum_{i=1}^{N}(x_i+h) > N.\sum_{i=1}^{N}(x_i+h)^{1-e}.$$

*This follows from Tchebycheff's inequality ([18], p.43)*, but for the reader's convenience, we include a complete proof for the particular case we need.

$$N\sum_{j=1}^{N}(x_j+h)^{1-e} - \sum_{i=1}^{N}(x_i+h)\,.\sum_{j=1}^{N}(x_j+h)^{-e}$$

$$= \sum_{i=1}^{N}\sum_{j=1}^{N}\left((x_j+h)^{1-e} - (x_i+h)(x_j+h)^{-e}\right)$$

$$= \sum_{j=1}^{N}\sum_{i=1}^{N}\left((x_i+h)^{1-e} - (x_j+h)(x_i+h)^{-e}\right)$$

$$= \frac{1}{2}\sum_{i,j}\left((x_j+h)^{1-e} - (x_i+h)(x_j+h)^{-e} + (x_i+h)^{1-e} - (x_j+h)(x_i+h)^{-e}\right)$$

$$= \frac{1}{2}\sum_{i,j}\left((x_i+h) - (x_j+h)\right).\left((x_i+h)^{-e} - (x_j+h)^{-e}\right).$$

This expression is only smaller than 0 for every $(x_i+h)_i$ if the power of $(x_i+h)$ is such that it reverses the order, i.e. when $-e < 0$ or $e > 0$. As this is always the case, this proves the proposition for $e \neq 1$.

*For the case $e = 1$ we have to show that the function* $\dfrac{N\prod_{i=1}^{N}(x_i+h)^{1/N}}{\sum_{i=1}^{N}(x_i+h)}$ increases, or that its derivative :

$$\frac{\prod_{i=1}^{N} (x_i + h)^{1/N} . \left[ \sum_{i=1}^{N} (x_i + h) . \sum_{j=1}^{N} (\frac{1}{x_j + h}) - N^2 \right]}{(\sum_{i=1}^{N} (x_i + h))^2} > 0 .$$

This again is a consequence of Tchebycheff's inequality ([18], p.43).

4.17. (C6) The transfer principle

This principle, originating from Dalton ([19]) states that if we make a strictly positive transfer from a poorer source to a richer, this must lead to a strictly positive increase in the index of inequality. Formulated in a precise mathematical way this is : if $x_i \leq x_j$ and $0 < h \leq x_i$, then

$$f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N) < f(x_1,\ldots,x_i-h,\ldots,x_j+h,\ldots,x_N) \qquad (25)$$

We remark that such a transfer leaves the arithmetic mean unchanged.

Again it is easy to show that $\sigma$, hence $\sigma^2$, V, $V^2$, K and CON satisfy the principle of transfers. The fact that the generalized Pratt measure P(r) (and hence also C, G and again V) satisfies the transfer principle will follow from our investigations in Section 6. Here we will show that Schutz' coefficient and the variance of logarithms do not satisfy(C6).Theil's measure and the Atkinson index on the other hand do satisfy the transfer principle.

Schutz' coefficient does not satisfy the transfer principle for changes between $x_i$ and $x_j$ such that $x_j + h$ is smaller than $\mu$, or, such that $x_i - h$ is larger than $\mu$, obviously leave D invariant. So, the relative mean deviation is not a good measure of concentration.

Also, the variance of logarithms does not always satisfy the transfer principle.

4.18. Example

Take N = 4; $x_1 = 1$, $x_2 = 2$, $x_3 = 24$, $x_4 = 25$ and i = 3, j = 4, h = 1. Then L = 2.0936 and after the transfer L becomes 2.0929, so that in this case L decreases instead of showing an increase.

Allison ([2], p.868) writes that this effect happens when $x_i$ and $x_j$ are both larger than e ($\approx$ 2.718) times the geometric mean. We were unable to verify this assertion.

We also remark that for N = 2, L does satisfy the principle of transfers (see appendix). But, of course, this remark is unimportant in practice.

4.19. Proposition. Theil's measure satisfies the principle of transfers.

*Proof :*

We have to show that

$$(x_i - h) \ln (\frac{x_i - h}{\mu}) + (x_j + h) \ln (\frac{x_j + h}{\mu}) > x_i \ln (\frac{x_i}{\mu}) + x_j \ln (\frac{x_j}{\mu}) ,$$

or $(x_i - h)\ln(x_i - h) + (x_j + h)\ln(x_j + h) > x_i \ln x_i + x_j \ln x_j$.
This follows immediately from the fact that the function $t \ln t$ is convex.

4.20. Proposition. A(e) satisfies the transfer principle, for every $e > 0$.

*Proof :*

For every $0 < e \leq 1$, the function $H(t) = t^{1-e}$ is increasing and concave, hence, for every $h > 0$, and $x_i \leq x_j$

$$H(x_i) - H(x_i - h) \geq H(x_j + h) - H(x_j) \qquad (*)$$

As

$$A(e) = \frac{(N\mu^{1-e})^{\frac{1}{1-e}} - (\sum_{k=1}^{N} x_k^{1-e})^{\frac{1}{1-e}}}{N^{\frac{1}{1-e}}.\mu}$$

(*) and the fact that $J(t) = t^{1/(1-e)}$ is increasing shows that such a transfer diminishes the second term of the nominator, hence A(e) itself increases.

If $e > 1$, the function $H_1(t) = t^{1-e}$ is decreasing and convex, hence, for $h > 0$ and $x_i \leq x_j$

$$H_1(x_i - h) - H_1(x_i) \geq H_1(x_j) - H_1(x_j + h) \qquad (**)$$

This shows that the term $(\sum_{k=1}^{N} x_k^{1-e})$ increases, but as $J_1(t) = t^{\frac{1}{1-e}}$ is decreasing in this case, $(\sum_{k=1}^{N} x_k^{1-e})^{\frac{1}{1-e}}$ decreases as a whole, which shows that A(e) increases.

Finally, as $(x_i - h)(x_j + h) < x_i x_j$ also A(1) satisfies the transfer principle.

In the next sections we will study some consequences and extensions of the transfer principle. Here however, we already note one important consequence.

4.21. Theorem

If f satisfies the transfer principle (C6), then

$$f(n,0,\ldots,0) = \max_{\sum_{k=1}^{N} x_k = n} f(x_1,\ldots,x_N)$$

Indeed, one can transfer one unit at the time from a poorer source to a richer one. By the transfer principle, the function f increases during this process. It stops when one source contains all the items. As a consequence, it is at that moment that f attains its maximal value.

4.22. It could also be argued that a good measure of concentration should vary between 0 and 1 :

Principle (B) : For all $(x_1,\dots,x_N)$

$$0 \le f(x_1,\dots,x_N) \le 1$$

However, this principle is only a mathematical convenience. It should not imply any preference, as simple transformations can produce any desired bounds. If a measure f is positive and does not satisfy the requirement that $f \le 1$, then we can use the transformation

$$f \to \frac{f}{1+f}$$

This yields an increasing function of f with values in the interval [0,1]. The transformed function satisfies (C1) to (C6) if f does.

4.23. We conclude that the following measures satisfy all our concentration principles and hence might be considered to be good concentration measures in the case the number of sources stays fixed : the coefficient of variation (V) and its square ($V^2$), the Yule characteristic (K), the CON-index, Pratt's measure (C) and the Gini index (G), the generalized Pratt index, P(r), for $r \ge 1$, Theil's measure and Atkinson's index.

4.24. Remark

The principles we have studied in this section can be described in a more abstract mathematical framework. Then they are a consequence of the fact that good concentration measures must be strictly Schur-convex and scale invariant. For the reader interested in this mathematical theory we refer to [20].

## 5. REQUIREMENTS RELATED TO THE TRANSFER PRINCIPLE

### 5.1. An equivalent formulation

Instead of taking from a poorer source to give to a richer in order to increase the inequality, one may also consider the opposite. Does giving to the poor what has been taken from the rich diminish the inequality? And, is this in some sense equivalent with the transfer principle as expressed in (25)? The exact answer is given in Theorem 5.2. We thank professor I.K. Ravichandra Rao for suggesting us to investigate this matter.

### 5.2. Theorem

If f satisfies(C2)then the transfer principle :

$$0 < x_i \le x_j;\ 0 < h \le x_i \Rightarrow \tag{25}$$
$$f(x_1,\dots,x_i,\dots,x_j,\dots,x_N) < f(x_1,\dots,x_i-h,\dots,x_j+h,\dots,x_N)$$

is equivalent with the following

$$0 < x_i \le x_j,\ |(x_i+h) - (x_j-h)| < x_j - x_i \Rightarrow \tag{26}$$
$$f(x_1,\dots,x_i+h,\dots,x_j-h,\dots,x_N) < f(x_1,\dots,x_i,\dots,x_j,\dots,x_N)$$

*Proof :*

A. (25) $\Rightarrow$ (26)

( i) If $0 \leq (x_j+h) - (x_i+h) < x_j - x_i$ then $x_i+h \leq x_j-h$ and $h > 0$. Hence (25) implies that

$$f(x_1,\ldots,x_i+h,\ldots,x_j-h,\ldots,x_N) < f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N) .$$

(ii) If $0 \leq (x_i+h) - (x_j-h) < x_j - x_i$ then obviously $x_i+h \geq x_j-h$. Moreover, $(x_i+h) - (x_j-h) < x_j - x_i$ implies that $h < x_j - x_i$ or $0 < x_j - h - x_i$. Then (25) implies that

$$f(x_1,\ldots,x_i+h,\ldots,x_j-h,\ldots,x_N) < f(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_N) ,$$

(with $(x_j-h-x_i)$ in the role of $h$).

This proves part A. (Remark that we have used the fact that f satisfies (C2)).

B. (26) $\Rightarrow$ (25)

If $0 < h \leq x_i$, and $x_i \leq x_j$ then

$$|x_i-x_j| = x_j - x_i < (x_j+h) - (x_i-h) .$$

By (26), this implies that

$$\begin{aligned} &f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N) \\ &= f(x_1,\ldots,(x_i-h)+h,\ldots,(x_j+h)-h,\ldots,x_N) \\ &< f(x_1,\ldots,x_i-h,\ldots,x_j+h,\ldots,x_N) . \end{aligned}$$

This proves part B.

5.3. Remark

It is easy to see that (25) and (26) are also equivalent with

$$0 < x_i \leq x_j;\ 0 \leq (x_j-h) - (x_i+h) < x_j - x_i \tag{27}$$
$$\Rightarrow f(x_1,\ldots,x_i+h,\ldots,x_j-h,\ldots,x_N) < f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N)$$

Indeed, (26) $\Rightarrow$ (27) (trivial) and (27) $\Rightarrow$ (25) as in part B of the preceding proof. Finally, (25) $\Rightarrow$ (26) (part A of Theorem 5.2).

The transfer principle is also equivalent with the following.

5.4. Proposition. The transfer principle is equivalent with the requirement that for $x_i \leq x_j$ the function

$$\Delta(h) = f(x_1,\ldots,x_i-h,\ldots,x_j+h,\ldots,x_N)$$

is strictly increasing on $[0,x_i]$.

*Proof :*

If f satisfies the transfer principle and $h_1 < h_2$ then

$$f(x_1,\dots,x_i - h_2,\dots,x_j + h_2,\dots,x_N)$$
$$= f(x_1,\dots,x_i - h_1 - (h_2-h_1),\dots,x_j + h_1 + (h_2-h_1),\dots,x_N)$$
$$> f(x_1,\dots,x_i - h_1,\dots,x_j + h_1,\dots,x_N) \ .$$

Conversely, if $\Delta(h)$ is strictly increasing then $\forall h \in ]0,x_i]$

$$f(x_1,\dots,x_i - h,\dots,x_j + h,\dots,x_N) > f(x_1,\dots,x_i,\dots,x_j,\dots,x_N) \ .$$

The transfer principle entails the following interesting consequences.

5.5. Proposition

( i) If f satisfies the transfer principle then

$$f(x_1 + h, x_2 - \frac{h}{N-1},\dots,x_N - \frac{h}{N-1}) > f(x_1,\dots,x_N)$$

where $0 < h \le (N-1).\min\{x_2,\dots,x_N\}$ and

$$x_1 = \max\{x_1,\dots,x_N\} \qquad (28)$$

(ii) If f satisfies the transfer principle then

$$f(x_1 - h, x_2 + \frac{h}{N-1},\dots,x_N + \frac{h}{N-1}) > f(x_1,\dots,x_N)$$

where $x_1 = \min\{x_1,\dots,x_N\}$ and $0 < h \le x_1$ (29)

*Proof :*

We will only show (i), (ii) follows in a similar way.
By the transfer principle, we have to following inequalities :

$$f(x_1,\dots,x_N) < f(x_1 + \frac{h}{N-1},\ x_2 - \frac{h}{N-1},\ x_3,\dots,x_N)$$

$$f(x_1 + \frac{h}{N-1},\ x_2 - \frac{h}{N-1},\ x_3,\dots,x_N) < f(x_1 + \frac{2h}{N-1},\ x_2 - \frac{h}{N-1},\ x_3 - \frac{h}{N-1},\dots,x_N)$$

$$\dots$$

$$f(x_1 + \frac{(N-2)h}{N-1},\ x_2 - \frac{h}{N-1},\dots,x_{N-1} - \frac{h}{N-1}, x_N)$$
$$< f(x_1 + h,\ x_2 - \frac{h}{N-1},\dots,x_N - \frac{h}{N-1}) \ .$$

Combining these (N-1) inequalities yields the required result.

It is remarkable that, although in practice, (28) is almost the same as (C4a), the proof that P(r) satisfies (C4a) is not trivial while the proof that P(r) satisfies (28) is easy, as can be seen from the following direct proof.

5.6. Proposition. The generalized Pratt measure P(r) satisfies (28).

*Proof :*

We have to show that

$$\sum_{i=2}^{N}\sum_{j=2}^{N}|x_i - x_j|^r + 2\sum_{i=2}^{N}(x_1 + h\frac{N}{N-1} - x_i)^r > \sum_{i=1}^{N}\sum_{j=1}^{N}|x_i - x_j|^r \quad ,$$

which yields

$$\sum_{i=1}^{N}(x_1 + h\frac{N}{N-1} - x_i)^r > \sum_{i=1}^{N}(x_1 - x_i)^r \quad .$$

This inequality is obviously satisfied.

The main difficulty in proving (C4a) for P(r) (or any other measure) lies mainly in the fact that in this principle $\sum_{i=1}^{N} x_i$ does not remain constant (as opposed to the situation in (28)). This is however a very natural situation (in econometrics : the total wealth of a country is not constant in time; in bibliometrics : the total number of articles in a bibliography over a fixed time interval, is not constant in time, and so on).

## 6. THE EXTENDED TRANSFER PRINCIPLE

In this section we introduce a family of principles, related to, and in fact extending the transfer principle.

### 6.1. The E(p) principle

If $(x_1,\ldots,x_N)$ is transformed into $(x_1',\ldots,x_N')$ such that

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i'$$

and

$$\sum_{i=1}^{N}\sum_{j=1}^{N}|x_i - x_j|^p < \sum_{i=1}^{N}\sum_{j=1}^{N}|x_i' - x_j'|^p \ , \ p \geq 1$$

then $f(x_1,\ldots,x_N) < f(x_1',\ldots,x_N')$.

V, $V^2$, K and CON obviously satisfy E(2); C and G satisfy E(1) (by Proposition 3.1 and Corollary 3.3); P(r) trivially satisfies E(r), for every $r \geq 1$.

The next proposition shows that E(p) is indeed a generalization of the transfer principle.

6.2. Proposition. If a concentration measure f satisfies E(p) for some $p \geq 1$, then it satisfies the transfer principle.

*Proof :*

By induction on the number of sources (N).
Suppose that $(x_k)_{k=1,\ldots,N}$ is ordered decreasingly and let $(x_k')_{k=1,\ldots,N}$

denote the sequence $(x_1,\ldots,x_i+h,\ldots,x_j-h,\ldots,x_N)$ where $i < j$ and $0 < h \leqq x_j$. We have to show that

$$\sum_{k=1}^{N}\sum_{=1}^{N} |x_k - x_\ell|^p < \sum_{k=1}^{N}\sum_{=1}^{N} |x_k' - x_\ell'|^p \qquad (*)$$

For $N = 2$, this inequality becomes

$$(x_1 - x_2)^p < (x_1 - x_2 + 2h)^p ,$$

which is obviously satisfied.

We suppose now that (*) is satisfied for N and we will show that then (*) is also satisfied when N is replaced by N+1. We suppose the vectors to be decreasing.

a) If $x_1' = x_i + h$ and $x_{N+1}' = x_j - h$, then (*) is trivially satisfied for N+1.

b) We now suppose that $x_1' \neq x_i + h$ or $x_{N+1}' \neq x_j - h$.

b1) Suppose $x_{N+1}' \neq x_j - h$, hence $x_{N+1}' < x_j - h$ (and $x_{N+1} = x_{N+1}'$).
Deleting $x_{N+1}$ gives the following N-sequence :
$(x_1,\ldots,x_i+h,\ldots,x_j-h,\ldots,x_N)$ (denoted $(x_k')_{k=1,\ldots,N}$). Then

$$\sum_{k=1}^{N+1}\sum_{\ell=1}^{N+1} |x_k - x_\ell|^p$$

$$= \sum_{k=1}^{N}\sum_{\ell=1}^{N} |x_k - x_\ell|^p + 2\sum_{k=1}^{N} (x_k - x_{N+1})^p$$

$$< \sum_{k=1}^{N}\sum_{\ell=1}^{N} |x_k' - x_\ell'|^p + 2\,[(x_1 - x_{N+1})^p + \ldots + (x_i - x_{N+1})^p + \ldots + (x_j - x_{N+1})^p + \ldots + (x_N - x_{N+1})^p] \qquad (**)$$

(using the induction hypothesis).

Consider the vectors $A = (x_i - x_{N+1},\ x_j - x_{N+1})$ and $B = (x_i + h - x_{N+1},\ x_j - h - x_{N+1})$. Note that all components are non-negative. It is easy to see that [18], p.89 is applicable to A and B and hence (since $\varphi(\cdot) = |\cdot|^p$ is convex)

$$(x_i - x_{N+1})^p + (x_j - x_{N+1})^p \leqq (x_i + h - x_{N+1})^p + (x_j - h + x_{N+1})^p$$

Substituting this in (**) gives

$$\sum_{k=1}^{N+1}\sum_{\ell=1}^{N+1} |x_k - x_\ell|^p$$

$$< \sum_{k=1}^{N}\sum_{\ell=1}^{N} |x_k' - x_\ell'|^p + 2\,[(x_1 - x_{N+1})^p + \ldots +$$

$$\ldots + (x_i + h - x_{N+1})^p + \ldots + (x_j - h - x_{N+1})^p + \ldots + (x_N - x_{N+1})^p]$$

$$= \sum_{k=1}^{N+1} \sum_{\ell=1}^{N+1} |x_k' - x_\ell'|^p \ .$$

*b2) Suppose* $x_i' \neq x_i + h$.
This proof is exactly the same as the one of (b1).

This finishes the proof of Proposition 6.2.

6.3. Corollary

The generalized Pratt measure, P(r), r ≥ 1, satisfies the transfer principle and hence, so does V, C and G.

6.4. Remark

There does not exist a function $f(x_1,\ldots,x_N)$ that satisfies E(p) for every $p \in \mathbb{N}_0$. *In fact, we can show [21] that if a function f satisfies* E(p) *for some* fixed p, it can not satisfy E(q), for all q ≠ p, at least if N ≥ 3, which is always the case in practical applications. (If N = 3 then the above is also true except for (p,q) = (2,4) or (p,q) = (4,2); indeed : if N = 3 then E(2) = E(4) : see [21]).

6.5. *Definition*

As a kind of limiting property for the principles E(p) we also formulate a principle E(∞) as follows :

If $(x_1,\ldots,x_N)$ is transformed into $(x_1',\ldots,x_N')$ such that

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i'$$

and

$$\max_{i,j} |x_i - x_j| < \max_{i,j} |x_i' - x_j'|$$

then

$$f(x_1,\ldots,x_N) < f(x_1',\ldots,x_N') \ .$$

Now, one might conjecture that also E(∞) implies the transfer principle. That this is not so can be seen as follows. We define a measure P(∞) by

$$P(\infty) = \frac{1}{\mu} \max_{i,j} |x_i - x_j|$$

It is easily seen that the measure P(∞) is a limiting measure of the measures P(r) since

$$\lim_{r\to\infty} \left( \sum_{i=1}^{N} |x_i|^r \right)^{1/r} = \max_i |x_i| \ .$$

This argument also explains why the notation $E(\infty)$ is used above.

For this measure $P(\infty)$ we show the following.

6.6. Proposition. $P(\infty)$ satisfies $E(\infty)$ but it does not satisfy the transfer principle, if $N \geq 4$.

*Proof :*

$P(\infty)$ satisfies $E(\infty)$ trivially. To see that $P(\infty)$ does not satisfy the transfer principle, an example suffices.
Take $N = 4$ :

$$x_1 = 1,\ x_2 = 2,\ x_3 = 3,\ x_4 = 5$$

$$x_1' = 1,\ x_2' = 1,\ x_3' = 4,\ x_4' = 5$$

(hence $i = 2$, $j = 3$, $h = 1$). Then

$$P(\infty)(1,2,3,5) = \frac{4}{\mu} = P(\infty)(1,1,4,5)\ .$$

6.7. Remark 1

It is easily seen that, if $N = 3$, $E(\infty) = E(1)$ (see also [21]). Hence only in this trivial case, $E(\infty)$ implies the transfer principle, since $E(1)$ does.

Remark 2

The facts that $E(p)$ implies the transfer principle and $E(\infty)$ does not (if $N \geq 3$), do not contradict each other. Indeed, although $E(\infty)$ is a limiting case of the $E(p)$'s, this does not mean that $E(\infty)$ inherits all properties of the $E(p)$'s. The problem lies in the fact that if

$$\max_{i,j} |x_i - x_j| < \max_{i,j} |x_i' - x_j'|$$

then there is a $p > 0$ such that

$$\sum_{i,j} |x_i - x_j|^p < \sum_{i,j} |x_i' - x_j'|^p$$

but this $p$ depends on the particular difference $\max_{i,j} |x_i' - x_j'| - \max_{i,j} |x_i - x_j|$.
So, the only thing that can be said is, that if $f$ satisfies $E(p)$ for every $p$ larger than some fixed $p_0$ then $f$ satisfies $E(\infty)$. However no function of this kind exists (cf. Remark 6.4).

It is possible that a measure satisfies the transfer principle and, in fact, all other principles C1 - C6, without satisfying any of the $E(p)$'s, $p \geq 1$. This is shown - several times - by the next results.

6.8. Proposition. Theil's measure does not satisfy any of the $E(p)$ principles, $p \geq 1$.

*Proof :*

Let $N = 3$, $X = (x_1,x_2,x_3) = (2,47,134)$ and $X' = (x_1',x_2',x_3') = (8,34,141)$. We will first show that, for every $r \geq 1$ :

$$\sum_{i,j=1}^{3} |x_i - x_j|^r < \sum_{i,j=1}^{3} |x'_i - x'_j|^r \quad . \qquad (*)$$

For this, it suffices to show that, for every $r \geq 1$ :

$$45^r + 132^r + 87^r < 26^r + 133^r + 107^r \ .$$

We put :

$$A = (a_1, a_2, a_3) = (132, 87, 45)$$

and

$$B = (b_1, b_2, b_3) = (133\alpha, 107\alpha, 26\alpha) \quad ,$$

where $\alpha = \frac{264}{266}$ . The parameter $\alpha$ ensures that $\sum_{i=1}^{3} a_i = \sum_{i=1}^{3} b_i$.

Furthermore : $a_1 = b_1$ and $a_1 + a_2 < b_1 + b_2$. As also $a_1 \geq a_2 \geq a_3$, $b_1 \geq b_2 \geq b_3$ and $r \geq 1$, we can apply [18], p.89 once more, showing that, for every $r \geq 1$ :

$$132^r + 87^r + 45^r \leq (133\alpha)^r + (107\alpha)^r + (26\alpha)^r \ .$$

Hence

$$132^r + 87^r + 45^r < 133^r + 107^r + 26^r \ ,$$

which shows that (*) is satisfied. We also have $\sum_{i=1}^{3} x_i = \sum_{i=1}^{3} x'_i = 183$, but

$$Th(X) = 0.472 > Th(X') = 0.448 \ ,$$

showing that Th does not satisfy E(p), for every $p \geq 1$.

6.9. Proposition. Atkinson's index A(e) does not satisfy E(p), for every $p \geq 1$.

*Proof :*

We will show this only for e = 0.5, e = 1, e = 2 and e = 3. We begin with the same example as in Proposition 6.8. Since

$$\sqrt{2} + \sqrt{47} + \sqrt{134} = 19.846 < \sqrt{8} + \sqrt{34} + \sqrt{141} = 20.534$$

we see that A(0.5) does not satisfy E(p), for every $p \geq 1$.

Also, since

$$(2).(47).(134) = 12596 < (8).(34).(141) = 38352 \ ,$$

we see that A(1) does not satisfy E(p), for every $p \geq 1$.

Finally, choose $X = (x_i)_{i=1}^{N}$ and $X' = (x'_i)_{i=1}^{N}$, such that, for every $p \geq 1$ :

$$\sum_{i,j=1}^{N} |x_i - x_j|^p < \sum_{i,j=1}^{N} |x'_i - x'_j|^p \qquad (*)$$

and such that

$$\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x'_i \qquad (**)$$

(and no $x_i$ or $x'_i$ is zero).

When N = 2, this can be realized by taking X = (x,x) and X' on the line segment joining (2x,0) and (0,2x), but X' ≠ X and not on the x-axis or y-axis, see figure 2.

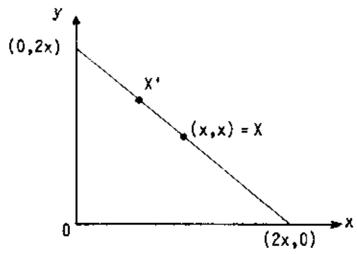


Fig.2

Since (*) is also valid for p = 2, we have, by (1) in 2.1, that (continuing with N = 2) :

$$x_1^2 + x_2^2 < x_1'^2 + x_2'^2 \qquad (***)$$

We now show that (**) and (***) imply that

$$\frac{1}{x_1^2} + \frac{1}{x_2^2} < \frac{1}{x_1'^2} + \frac{1}{x_2'^2} .$$

Indeed :

$$\frac{1}{x_1^2} + \frac{1}{x_2^2} = \frac{x_1^2 + x_2^2}{x_1^2 x_2^2} < \frac{x_1'^2 + x_2'^2}{x_1^2 x_2^2}$$

But $x_1 + x_2 = x'_1 + x'_2$, hence

$$x_1^2 + 2x_1x_2 + x_2^2 = x_1'^2 + 2x'_1x'_2 + x_2'^2 .$$

So, (***) implies

$$x_1x_2 > x'_1x'_2 . \qquad (****)$$

This yields

$$\frac{1}{x_1^2} + \frac{1}{x_2^2} < \frac{x_1'^2 + x_2'^2}{x_1'^2 x_2'^2} = \frac{1}{x_1'^2} + \frac{1}{x_2'^2} .$$

This, in turn, together with (*) and (**) yields that A(3) does not satisfy E(p) for every p ≥ 1. Also,

$$\frac{1}{x_1} + \frac{1}{x_2} = \frac{x_1 + x_2}{x_1x_2} < \frac{x'_1 + x'_2}{x'_1x'_2} = \frac{1}{x'_1} + \frac{1}{x'_2}$$

by (**) and (****). This shows, together with (*) and (**) that A(2) does not satisfy E(p) for every p ≥ 1.

6.10. Remark

In [21], it is shown that, when N = 3, E(1) = E(∞). Consequently, we have that Th and A(e) do not satisfy E(∞) also.

## 7. SENSITIVITY TO TRANSFERS

7.1. The transfer principle describes an increase of f under the transformation

$$(x_1,\dots,x_i,\dots,x_j,\dots,x_N) \to (x_1,\dots,x_i-h,\dots,x_j+h,\dots,x_N)$$

where $x_i \le x_j$ and $0 < h \le x_i$, but it does not say anything about the degree of increase in function of other parameters. One such parameter could be the difference between $x_j$ and $x_i$ or the place where the transfer occurs (when the N sources are ordered in some natural way). Considerations on this kind of sensitivity are given e.g. by Atkinson ([1]) and Allison ([2], [3]).

In this context we first offer the following proposition.

7.2. Proposition. If, for every i and j, $G_{i,j}(t) = f(x_1,\dots,x_i,\dots,x_j,\dots,x_N)$ is only function of $t = x_j - x_i$, $x_j \ge x_i$, then the following are equivalent :

(a) the functions $G_{i,j}(t)$ are strictly increasing, for every i,j;

(b) f satisfies the transfer principle.

*Proof :*

If the functions $G_{i,j}(t)$ are strictly increasing then obviously

$$f(x_1,\dots,x_i-h,\dots,x_j+h,\dots,x_N) > f(x_1,\dots,x_i,\dots,x_j,\dots,x_N) .$$

Conversely, if $0 \le t_1 < t_2$, take $x_i$ and $x_j$ such that $x_i \ge \frac{t_2 - t_1}{2}$ and $x_j - x_i = t_1$ and put $h = \frac{t_2 - t_1}{2}$, then, by the transfer principle :

$$G_{i,j}(t_1) = f(x_1,\dots,x_i,\dots,x_j,\dots,x_N)$$
$$< f(x_1,\dots,x_i-\frac{t_2-t_1}{2},\dots,x_j+\frac{t_2-t_1}{2},\dots,x_N) = G_{i,j}(t_2) .$$

This proves the proposition.

7.3. The expression $f(x_1,\dots,x_i-h,\dots,x_j+h,\dots,x_N)$ can also be studied as a function of "the place" where the transfer occurs. As there are actually two places where changes occur, the simplest way to proceed is to consider f as a function of $\frac{i+j}{2}$ . This approach makes the most sense when the sources have some intrinsic ordering (i.e. are not just ordered from largest to smallest), as is the case when studying the income distribution of a country, where sources are income classes.

The problem of sensitivity of a concentration measure depends largely on the specific situation to which the measure is applied. For instance, Allison [3] argues - in the case of publications or citations - in favor of a measure that is equally sensitive to differences at all levels of the distribution. The coefficient of variation satisfies this requirement.

We will now review the sensitivity of some of the good concentration measures we have found in section 4. We leave an exhaustive investigation for further research. In what follows, $\Delta f$ denotes the difference between $f(x_1,\ldots,x_i - h,\ldots,x_j + h,\ldots,x_N)$ and $f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_N)$, $0 < h \le x_i$; $x_i \le x_j$.

7.4. Sensitivity of $V^2$

$$\Delta V^2 = \frac{1}{N\mu^2}\,[(x_i - h - \mu)^2 + (x_j + h - \mu)^2 - (x_i - \mu)^2 - (x_j - \mu)^2]$$

$$= \frac{2h}{N\mu^2}\,((x_j - x_i) + h) \quad .$$

This expression depends linearly on the difference between $x_j$ and $x_i$, but is independent of the place where the transfer occurs.

7.5. Sensitivity of Theil's measure

$$\Delta Th = \frac{1}{N\mu}\,[(x_i - h)\ln(x_i - h) + (x_j + h)\ln(x_j + h) - x_i \ln x_i - x_j \ln x_j]$$

$$= \frac{1}{N\mu}\,(x_i \ln\frac{x_i - h}{x_i} + x_j \ln\frac{x_j + h}{x_j} + h \ln\frac{x_j + h}{x_i - h}) \quad .$$

To estimate the meaning of this difference we will show that

$$\lim_{h\to 0} \frac{\Delta Th}{\frac{1}{N\mu}\, h \ln(\frac{x_j}{x_i})} = 1 \quad .$$

$$\lim_{h\to 0} \frac{x_i \ln(\frac{x_i - h}{x_i}) + x_j \ln(\frac{x_j + h}{x_j}) + h \ln(\frac{x_j + h}{x_i - h})}{h \ln(\frac{x_j}{x_i})}$$

= (using L'Hôpital's rule)

$$\lim_{h\to 0} \frac{\frac{-x_i^2}{(x_i - h).x_i} + \frac{x_j^2}{(x_j + h).x_j} + \ln\frac{x_j + h}{x_i - h} + h\,\frac{(x_i - h)(x_i + x_j)}{(x_j + h)(x_i - h)^2}}{\ln(\frac{x_j}{x_i})}$$

$$= \frac{-1 + 1 + \ln\left(\frac{x_j}{x_i}\right) + 0}{\ln\left(\frac{x_j}{x_i}\right)} = 1 .$$

Hence, for small h, $\Delta Th$ depends linearly on the difference between $\ln x_j$ and $\ln x_i$.

7.6. Sensitivity of Pratt's measure

Let $j < i$, hence $x_i < x_j$. Then

$$C_1 = \frac{N+1}{N-1} - \frac{2}{\mu N\,(N-1)} \sum_{k=1}^{N} k x_k \qquad \text{(formula (13))}$$

and $C_2$, after a transfer of h from $x_i$ to $x_j$ :

$$C_2 = \frac{N+1}{N-1} - \frac{2}{\mu N\,(N-1)} \sum_{k=1}^{N} k x'_k \quad ,$$

where $(x'_k)_k$ is the transformed, ordered sequence consisting of $x_i - h$, $x_j + h$ and $x_k$, $k \neq i$, $k \neq j$.

Suppose now that $x_i - h$ has a rank equal to $i+s$ and that the rank of $x_j + h$ is $j-r$ $(r,s \neq 0)$, then

$$\Delta C = C_2 - C_1$$

$$= \frac{2}{\mu N\,(N-1)} \{[(j-r)x_{j-r} + (j-r+1)x_{j-r+1} + \ldots + (j-1)x_{j-1} + j\, x_j]$$

$$- [(j-r)(x_j + h) + (j-r+1)x_{j-r} + \ldots + j\, x_{j-1}]$$

$$+ [i\, x_i + (i+1)x_{i+1} + \ldots + (i+s)x_{i+s}]$$

$$- [i\, x_{i+1} + (i+1)x_{i+2} + \ldots + (i+s-r)x_{i+s} + (i+s)(x_i - h)]\}$$

$$= \frac{2}{\mu N\,(N-1)} [h(i+s-j+r) - x_{j-r} - x_{j-r+1} - \ldots - x_{j-1} + r\, x_j + x_{i+1} + \ldots + x_{i+s} - s\, x_i] \quad .$$

In the special case where there is no shift in rank

$$\Delta C = \frac{2h\,(i-j)}{\mu N\,(N-1)} \quad ,$$

as is readily seen.

For typically shaped income distributions the Pratt measure (and the Gini index) tends to be most sensitive to transfers around the middle of the distribution. On the other hand, for publication patterns, the Pratt measure is most sensitive among low producers.

The Atkinson index A(e) and the generalized Pratt index P(r) are in fact families of concentration measures. According to the value of the parameter e or r the sensitivity changes. They have the advantage that the value of the parameter can be adapted to the specific case under consideration.

## 8. DISPERSION MEASURES

8.1. Roughly speaking, dispersion measures are the opposite of concentration measures. Making adequate changes to the principles for concentration measures (such as, using < instead of >) one obtains a set of principles good dispersion measures must fulfill.

8.2. A general strategy to construct dispersion measures from concentration measures is the following. Let f be a concentration measure taking values in the interval [0,1]. Then h = 1 - f is a dispersion measure. So, implicitly, we already know a large set of dispersion measures.

8.3. Independently of the above remarks, Heine ([22]) studied some measures of dispersion. In our notation, the three measures he introduced are given as follows. Let $(x_1,\ldots,x_N)$ be ordered in increasing order (this assumption is also made in [22]).

A. The adapted Gini index $D_G$.

If $N \neq 1$ and $N \neq \sum_{i=1}^{N} x_i$, then, with $a_j = \dfrac{x_j}{\sum_{k=1}^{N} x_k}$,

$$D_G = 1 - \sum_{i=1}^{N-1} \left(\frac{i}{N} - \sum_{j=1}^{i} a_j\right)\left(\frac{2\sum_{k=1}^{N} x_k}{(N-1)\left(\sum_{k=1}^{N} x_k - N\right)}\right) \tag{30}$$

B. Singleton's index $D_S$.

$$\text{If } N \neq 1,\ D_S = \frac{2}{N-1}\sum_{i=1}^{N-1} (N-i)\, a_i \tag{31}$$

(It can readily be verified that formula (31) is the same as Heine's more intricate formula (7) in [22]).

C. The normalized entropy index $D_E$.

If $N \neq \sum_{i=1}^{N} x_i$ and $N+1 \neq \sum_{i=1}^{N} x_i$, then

$$D_E = \frac{E - E_{min}}{E_{max} - E_{min}} \tag{32}$$

with

$$E = -\sum_{i=1}^{N} a_i \log_2 a_i \tag{33}$$

$$E_{min} = \log_2 \left(\sum_{i=1}^{N} x_i\right) - \left(1 - \frac{N-1}{\sum_{i=1}^{N} x_i}\right) \log_2 \left(N+1 - \sum_{i=1}^{N} x_i\right) \tag{34}$$

and

$$E_{max} = -(N-K) \cdot \frac{\sum_{i=1}^{N} x_i - K}{N \cdot \sum_{i=1}^{N} x_i} \log_2 \left(\frac{\sum_{i=1}^{N} x_i - K}{N \sum_{i=1}^{N} x_i}\right) -$$

$$K \cdot \frac{\sum_{i=1}^{N} x_i - K + N}{N \cdot \sum_{i=1}^{N} x_i} \log_2 \left(\frac{\sum_{i=1}^{N} x_i - K + N}{N \cdot \sum_{i=1}^{N} x_i}\right) \tag{35}$$

where $K = \sum_{i=1}^{N} x_i \pmod N$, i.e. K is the rest of $\sum_{i=1}^{N} x_i$ after division by N. Although intricate, $D_E$ can immediately be determined from the raw data $(x_1,\ldots,x_N)$. The use of $\log_2$ shows its relation with information theory (but this is not essential).

8.4. In [22], Heine gives properties of the measures $D_G$, $D_S$ and $D_E$. We note however that $D_G$, $D_S$ and $D_E$ are not new measures. Instead, we can relate them to well-known concentration measures.

8.5. Proposition. $D_S = 1 - C$.

*Proof :*

By (10)

$$C = \frac{2\left(\frac{N+1}{2} - q\right)}{N-1}$$

where q is now equal to $\sum_{i=1}^{N} (N-i+1)\, a_i$, since $(x_1,\ldots,x_N)$ is now ordered in increasing order.
So

$$C = \frac{N+1-2\sum_{i=1}^{N}(N-i)a_i - 2}{N-1}$$

$$= 1 - \frac{2}{N-1}\sum_{i=1}^{N}(N-i)a_i$$

$$= 1 - \frac{2}{N-1}\sum_{i=1}^{N-1}(N-i)a_i = 1 - D_S$$

The next proposition shows that $D_G$ and $D_S$ are essentially the same.

8.6. Proposition

If $N \neq \sum_{i=1}^{N} x_i$ ,

$$D_G = -\frac{N}{\sum_{i=1}^{N} x_i - N} + \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N} \cdot D_S$$

*Proof :*

$$D_G = 1 - \left[\sum_{i=1}^{N-1} \frac{i}{N} - \sum_{i=1}^{N-1} \sum_{j=1}^{i} a_j\right]\left[\frac{2 \sum_{i=1}^{N} x_i}{(\sum_{i=1}^{N} x_i - N)(N-1)}\right]$$

$$= 1 - \left[\frac{N-1}{2} + \sum_{i=1}^{N-1} (N-i)a_i\right]\left[\frac{2 \sum_{i=1}^{N} x_i}{(\sum_{i=1}^{N} x_i - N)(N-1)}\right]$$

$$= 1 - \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N} + D_S \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N}$$

$$= -\frac{N}{\sum_{i=1}^{N} x_i - N} + D_S \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N}$$

8.7. Corollary. $D_G \leq D_S$, with equality only if all $x_i$ are equal.

*Proof :*

Since $0 \leq C \leq 1$, we have also $0 \leq D_S \leq 1$. Hence

$$\frac{N}{\sum_{i=1}^{N} x_i - N} \geq D_S \frac{N}{\sum_{i=1}^{N} x_i - N} = D_S \left(\frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N} - 1\right)$$

with equality only if $D_S = 1$, i.e. $C = 0$, i.e. when all $x_i$ are equal.
Hence

$$D_S \geq D_S \left(\frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N}\right) - \frac{N}{\sum_{i=1}^{N} x_i - N} = D_G .$$

The following corollaries are easy consequences from Propositions 8.5 and 8.6 and the relation between Pratt's measure and the Gini index.

8.8. Corollary

$$D_G = 1 - \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N} C \text{ , if } C \neq 0$$

and

$$D_G = 1 \Leftrightarrow C = 0$$

8.9. Corollary

$$D_G = 1 - \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N} \cdot \frac{N}{N-1} G$$

8.10. Corollary. $C \leq 1 - D_G$, with equality only if $C = 0$.

*Proof :*

$$1 - D_G = \frac{\sum_{i=1}^{N} x_i}{\sum_{i=1}^{N} x_i - N} C > C \text{ , if } C \neq 0 \text{ .}$$

When $C = 0$, then all $x_i$ are equal and hence $D_G = 1$. In this case : $C = 1 - D_G$.

The previous results show that $D_G$ nor $D_S$ are new measures, hence their properties can be deduced from Sections 3 and 4 : no special investigation is needed. Furthermore, the same holds for $D_E$.

8.11. Proposition

$$Th = \ln N - E \text{ ,}$$

hence

$$D_E = \frac{\ln N - Th - E_{min}}{E_{max} - E_{min}}$$

*Proof :*

$$Th = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{x_i}{\mu}\right) \ln \left(\frac{x_i}{\mu}\right)$$

$$= \sum_{i=1}^{N} a_i \ln (a_i N)$$

$$Th = \ln N - (-\sum_{i=1}^{N} a_i \ln a_i)$$

$$= \ln N - E .$$

So we conclude that Heine [22] is only dealing with the dispersion version of C (or G) and Th. As we have shown before that these are good concentration measures, $D_G$, $D_S$ and $D_E$ are good dispersion measures.

## 9. CONCLUSION; SUGGESTIONS FOR FURTHER RESEARCH

In this paper we have reviewed some concentration measures proposed in the literature. We have presented a set of principles - (C1) to (C6) - that good concentration measures must fulfill and have investigated some of the consequences of these principles. The transfer principle was extended to yield a new family of principles, denoted E(p), but a concentration measure can only satisfy E(p) for at most one p. We discussed briefly the issue of sensitivity to transfers and have shown how Heine's dispersion measures are related to some well-known concentration measures.

Suggestions for further research :

1. The principles we have investigated are stated with respect to a fixed number of sources. What then is the behavior of a good concentration measure with respect to a varying number of sources? (cf. [14]).

2. In [23], Egghe calculated and interpreted Pratt's measure for some classical bibliometric distributions, including the geometric distribution. A similar investigation, using other concentration measures might be interesting.

3. What are the implications of using different concentration measures and what is the "most desirable" level of inequality, in particular in connection with science policy decisions (cf. [24]).

4. A more refined study of sensitivity to transfers is certainly needed.

5. In [25] Stephen Cole raises the issue that the Gini index (hence also Pratt's measure) simultaneously measures two concepts. One is consensus in a field and the other is dispersion of recognition. Is this remark also valid with respect to other uses of the Gini index than in the field of sociology of science. Does it apply to any concentration measure?

6. The measures of concentration we have studied in this paper are only one-dimensional representations of the complex notion "concentration". Does there exist useful, interpretable, more-dimensional extensions of the measures we know?

## REFERENCES

[1] Atkinson, A.B., On the measurement of inequality. Journal of Economic Theory, 2 : 244-263, 1970.

[2] Allison, P.D., Measures of inequality. American Sociological Review, 43 : 865-880, 1978.

[3] Allison, P.D., Inequality and scientific productivity. Social Studies of Science, 10 : 163-179, 1980.

[4] Gaston, J., The reward system in British and American science. New York : Wiley, 1978.

[5] Johnson, R.L., Measures of vocabulary diversity. In : Ager D.E., Knowles, F.E., Smith J., editors. Advances in computer-aided literary and linguistic research. Birmingham : AMLC, 1979 : 213-227.

[6] Simpson, E.H., Measurement of diversity. Nature, 163 : 688, 1949.

[7] Schutz, R.R., On the measurement of income inequality. American Economic Review, 41 : 107-122, 1951.

[8] Gastwirth, J., The estimation of the Lorenz curve and Gini index. Review of Economics and Statistics, 54 : 306-316, 1972.

[9] Yntema, D., Measures of the inequality in the personal distribution of wealth or income. Journal of the American Statistical Association, 28 : 423-433, 1933.

[10] Pratt, A.D., A measure of class concentration in bibliometrics. Journal of the American Society for Information Science, 28 : 285-292, 1977.

[11] Gini, C., Il diverso accrescimento delle classi sociali e la concentrazione della ricchezza. Giornale degli Economisti, serie 11, 37, 1909.

[12] Carpenter, M.P., Similarity of Pratt's measure of class concentration to the Gini index. Journal of the American Society for Information Science, 30 : 108-110, 1979.

[13] Theil, H., Economics and information theory. Chicago : Rand McNally, 1967.

[14] Ray, J.L., Singer, J.D., Measuring the concentration of power in the international system. Sociological Methods & Research, 1 : 403-437, 1973.

[15] Rao I.K. Ravichandra, Probability distributions and inequality measures for analyses of circulation data. In : Egghe, L., Rousseau, R., editors. Informetrics 87/88, Amsterdam : Elsevier, 1988 : 231-248.

[16] Schechtman, E., Yitzhaki, S., A measure of association based on Gini's mean difference. Communications in Statistics - Theory and Methods, 16 : 207-231, 1987.

[17] Allison, A., The reliability of variables measured as the number of events in an interval of time. In : Schuessler, K.F., editor. Sociological methodology 1978, San Francisco : Jossey-Bass, 1977 : 238-253.

[18] Hardy, G., Littlewood, J.E., Polya, G., Inequalities. Cambridge : University Press, 1988.

[19] Dalton, H., The measurement of the inequality of incomes. Economic Journal, 30 : 248-361, 1920.

[20] Marshall, A., Olkin, I., Inequalities : theory of majorization and its applications. New York : Academic Press, 1979.

[21] Egghe, L., A norm problem, its solution and its application in econometrics and bibliometrics. Preprint 1988.

[22] Heine, M.H., Indices of literature dispersion based on qualitative attributes. Journal of Documentation, 34 : 175-188, 1978.

[23] Egghe, L., Pratt's measure for some bibliometric distributions and its relation with the 80/20 rule. Journal of the American Society for Information Science, 38 : 288-297, 1987.

[24] Chapman, I.D., Farina, C., Concentration of resources : the National Research Council's (Canada) grants in aid of research : 1964-1974. Scientometrics, 4 : 105-117, 1982.

[25] Cole, S., The hierarchy of the sciences? American Journal of Sociology, 89 : 111-139, 1983.

APPENDIX

1. If N = 2, L satisfies the transfer principle.

*Proof :*

$$L(x_1,x_2)$$
$$= \frac{1}{2} [(\ln x_1 - \frac{1}{2} (\ln x_1 + \ln x_2))^2 + (\ln x_2 - \frac{1}{2} (\ln x_1 + \ln x_2))^2]$$
$$= \frac{1}{8} [\ln^2 x_1 - 2 \ln x_1 . \ln x_2 + \ln^2 x_2 + \ln^2 x_2 - 2 \ln x_1 . \ln x_2 + \ln^2 x_1]$$
$$= \frac{1}{4} (\ln x_1 - \ln x_2)^2$$
$$= \frac{1}{4} \ln^2 (\frac{x_1}{x_2}) .$$

Similarly, (taking $x_1 \leq x_2$)

$$L(x_1 - h, x_2 + h) = \frac{1}{4} \ln^2 (\frac{x_1 - h}{x_2 + h}) .$$

As $\frac{x_1}{x_2} > \frac{x_1 - h}{x_2 + h}$, also $\ln (\frac{x_1}{x_2}) > \ln (\frac{x_1 - h}{x_2 + h})$, hence ($\ln (\frac{x_1}{x_2}) < 0$!)

$$\ln^2 (\frac{x_1}{x_2}) < \ln^2 (\frac{x_1 - h}{x_2 + h})$$

*or*

$$L(x_1,x_2) < L(x_1 - h, x_2 + h) .$$

2. The generalized Pratt measure, P(r), satisfies (C4a), for $r \geq 1$

The following proof is due to Q. Burrell who provided us with a simpler proof than the one we had originally. We thank him for his help.

*Proof :*

We note first that if a function z(t) is a quotient, say $\frac{u(t)}{v(t)}$, then the sign of z'(t) is the same as that of v(t)u'(t) - v'(t)u(t).

According to our principle (C2) we may order the sources in non-decreasing order of productivity :

$$0 \leq x_1 \leq x_2 \leq \ldots \leq x_N .$$

Then (C4a) becomes : if $x_{N-1} \neq 0$, then, for h > 0

$$f(x_1,\ldots,x_{N-1},x_N + h) > f(x_1,\ldots,x_N)$$

Now :

$$P(r) = \frac{\left(\frac{1}{2N(N-1)} \sum_{k=1}^{N} \sum_{\ell=1}^{N} |x_k - x_\ell|^r\right)^{1/r}}{\mu}$$

$$= \frac{\left(\frac{1}{N(N-1)} \sum_{1\le\ell<k\le N} (x_k - x_\ell)^r\right)^{1/r}}{\mu}$$

$$= \frac{\left(\frac{1}{N(N-1)} \left\{ \sum_{1\le\ell<k\le N-1} (x_k - x_\ell)^r + \sum_{\ell=1}^{N-1} (y - x_\ell)^r\right\}\right)^{1/r}}{\frac{1}{N}\left(\sum_{\ell=1}^{N-1} x_\ell + y\right)}$$

$$= F(y) \text{ , where } y = x_N \text{ .}$$

Now, $P(r)$ satisfies (C4a) if $F(y)$ is an increasing function of $y$, or equivalently, if

$$G(y) = \frac{\sum_{1\le\ell<k\le N-1} (x_k - x_\ell)^r + \sum_{\ell=1}^{N-1} (y - x_\ell)^r}{\left(\sum_{\ell=1}^{N-1} x_\ell + y\right)^r}$$

is increasing in $y$.

Then, by the remark made in the beginning of this proof,

$$\text{sgn } G'(y) = \text{sgn } \left\{\left(\sum_{1}^{N-1} x_\ell + y\right)^r \sum_{1}^{N-1} (y - x_\ell)^{r-1} r \right.$$

$$\left. - \left(\sum_{1\le\ell<k\le N-1} (x_k - x_\ell)^r + \sum_{1}^{N-1} (y - x_\ell)^r\right) r \left(\sum_{1}^{N-1} x_\ell + y\right)^{r-1}\right\}$$

$$= \text{sgn } \left\{\left(\sum_{1}^{N-1} x_\ell + y\right) \sum_{1}^{N-1} (y - x_\ell)^{r-1} - \left(\sum_{1\le\ell<k\le N-1} (x_k - x_\ell)^r \right.\right.$$

$$\left.\left. + \sum_{\ell=1}^{N-1} (y - x_\ell)^r\right)\right\} \text{ .}$$

Reverting to $y = x_N$, the expression within the braces can be rewritten as :

$$\left(\left(\sum_{k=1}^{N} x_k\right) \sum_{\ell=1}^{N-1} (x_N - x_\ell)^{r-1}\right) - \sum_{1\le\ell<k\le N} (x_k - x_\ell)^r \text{ .}$$

That this is positive is most easily seen by writing the two expressions in arrays as follows :

$$x_N(x_N - x_{N-1})^{r-1} + x_N(x_N - x_{N-2})^{r-1} + \ldots + x_N(x_N - x_2)^{r-1} + x_N(x_N - x_1)^{r-1}$$

$$+ x_{N-1}(x_N - x_{N-1})^{r-1} + \ldots$$

$$+ \ldots$$

$$+ x_1(x_N - x_{N-1})^{r-1} + \ldots \qquad + x_1(x_N - x_1)^{r-1};$$

while the second is best written as :

$$\begin{array}{l} (x_N - x_{N-1})^r + (x_N - x_{N-2})^r + \quad \ldots \qquad + (x_N - x_2)^r + (x_N - x_1)^r \\ \qquad + (x_{N-1} - x_{N-2})^r + \ldots \qquad + (x_{N-1} - x_1)^r \\ \qquad\qquad \ldots \\ \qquad\qquad + (x_3 - x_2)^r + (x_3 - x_1)^r \\ \qquad\qquad\qquad + (x_2 - x_1)^r . \end{array}$$

Because the $x_i$'s are ordered non-decreasingly, all of the above terms are non-negative and every term in the top array is at least as large as the one in the corresponding position below (blanks in the lower array being taken as zeroes), e.g. if $k > j$ then

$$x_k(x_N - x_j)^{r-1} \geq (x_k - x_j)(x_N - x_j)^{r-1} \geq (x_k - x_j)^r .$$

This shows that $G(y)$, hence $F(y)$ is increasing in $y$, so that $P(r)$ satisfies (C4a).