Central tendency for compositional data sets: an axiomatic approach

Thierry Marchant

April 21, 2010

Abstract

There is no unanimity in the literature about which measure of central tendency should be used for compositional data sets. In this paper, we use an axiomatic approach to this problem and we provide some support to a family of central tendency measures including the arithmetic mean.

1 Introduction

Suppose we measure the composition of n rock samples in k components. Each sample is then represented by k real numbers (a data point) between 0 and 1, adding up to 1. If we want to characterize this data set by a measure of central tendency, a simple and intuitive solution consists in computing the arithmetic mean. This definition of central tendency has been widely used in the literature. Some authors [most notably Aitchison, 1986] have argued that the arithmetic mean is not a suitable measure of central tendency because the coordinates of a data point are not independent of each other: they must satisfy the constant sum property. Since then, many authors have adopted Aitchison's view and use the logratio means. Yet many authors have different opinions and still use the arithmetic mean [Woronow, 1997, e.g.]. Recently, Sharp [2006] has noted that the logratio mean sometimes behaves strangely and he has proposed another measure of central tendency: the graph median.

So, today, there is no unanimity about the correct definition of the central tendency. That is why we think it is wise to pause a while and to give a deep thought to this concept. What is the central tendency or what does it ought to be? We will try to answer this question using an axiomatic approach. That is, we will define some axioms (properties that we think the central tendency must satisfy) and we will then see which definitions of the central tendency are compatible with these axioms. And if our axioms are strong enough, we will eventually identify a unique central tendency definition compatible with them. This way, we will avoid any ad hoc argument (e.g., 'Now variances and covariances of ratios are awkward to manipulate, and as any lecturer in statistics must grow weary of telling his or her students, when stuck by complicated products and quotients: take logarithms' [Aitchison, 1986, p.65]).

It will turn out that only the arithmetic mean is compatible with the axioms that we will propose. Some might then question our axioms and wonder whether other axioms would lead to other measures of central tendency. They are right: our axioms are not dogmas, they can be criticized (even if we consider some of them as inescapable). But they provide a very clear, explicit and formal justification for the arithmetic mean. They constitute a basis for a dialog between proponents of different means. If anyone finds our axioms not compelling, she can propose other axioms and find which definition of central tendency is compatible with her axioms.

Before proceeding, let us mention that the arithmetic mean has already been axiomatically characterized [e.g., Kolmogoroff, 1930] but only for single-dimensional data sets.

2 Problem statement and axioms

We consider data sets of size n where each data point is a k-dimensional vector expressing a composition, so that each coordinate is non-negative and the sum of the coordinates of each data point is equal to 1. We will consider data sets of different sizes but the dimension of the data points will not vary. Let $K = \{1, \ldots, k\}$, $N(n) = \{1, \ldots, n\}$ and $X = \{u \in \mathbb{R}^k : u_i \geq 0 \ (i \in K)$ and $\sum_{i \in K} u_i = 1\}$. The set X represents all possible data points. The set X^n represents all possible data sets of size n. Let $\mathcal{X} = \bigcup_{n \in \mathbb{N}_0} X^n$. The set $\mathcal{X}$ represents all possible data sets of any size.

We want to define a mapping $M : \mathcal{X} \to X : (x^1, \ldots, x^n) \to M(x^1, \ldots, x^n)$ so that, for any data set $(x^1, \ldots, x^n) \in \mathcal{X}$, the composition $M(x^1, \ldots, x^n)$ represents the central tendency in $(x^1, \ldots, x^n)$ and $M_i(x^1, \ldots, x^n)$ is the ith coordinate of the central tendency.

We are now able to formally define some popular means. The arithmetic mean M^a is defined by $M_i^a(x^1, \ldots, x^n) = \sum_{j=1}^n x_i^j/n$ for all $i \in N$. The logratio mean M^l is defined, following Aitchison [1986], by

$$M_i^l(x^1, \ldots, x^n) = \frac{\exp\left(\sum_{j=1}^n \frac{\ln(x_i^j/x_k^j)}{n}\right)}{\sum_{l=1}^k \exp\left(\sum_{j=1}^n \frac{\ln(x_l^j/x_k^j)}{n}\right)} \tag{1}$$

for all $i \in N$. Notice that, since $\ln(x_i^j/x_k^j) = \ln(x_i^j) - \ln(x_k^j)$ and $\frac{1}{n} \ln x_i^j = \ln(x_i^j)^{1/n}$, we can rewrite (1) as

$$M_i^l(x^1, \ldots, x^n) = \frac{\exp\left(\frac{1}{n} \sum_{j=1}^n \ln x_i^j\right)}{\sum_{l=1}^k \exp\left(\frac{1}{n} \sum_{j=1}^n \ln x_l^j\right)} \tag{2}$$

or the much simpler

$$M_i^l(x^1, \ldots, x^n) = \frac{\prod_{j=1}^n (x_i^j)^{1/n}}{\sum_{l=1}^k \prod_{j=1}^n (x_l^j)^{1/n}}. \tag{3}$$

Most authors define the logratio mean only for strictly positive compositions because $\ln x_i^j$ is undefined when $x_i^j = 0$. But, actually, this is not a problem because, even if $\ln x_i^j$ is undefined, $\exp\left(\frac{1}{n}\sum_{j=1}^{n}\ln x_i^j\right)$ is defined: it is equal to zero. So, if $x_i^j = 0$, then $M_i^l(x^1,\ldots,x^n) = 0$ and, for any p such that $x_p^j \neq 0\,\forall j \in N(n)$, $M_p^l(x^1,\ldots,x^n)$ is given by (3). The only serious difficulty with zero occurs when, for each $i \in K$, there is a data point with the ith component equal to zero. In that case, the denominator of the right hand side of (3) is zero. So, $M_p^l(x^1,\ldots,x^n) = 0/0$ and is indeterminate. In that case, for each data point, we add $s\varepsilon$ to each zero value and we subtract $(k-s)\varepsilon$ from each non-zero value, where s is the number of non-zero values in the data point. After this modification, the new vector is still a data point (i.e., with non-negative coordinates adding up to 1). We then compute the limit of M^l when ε goes to zero using (3). It is easy to see that, when we add $s\varepsilon$ or subtract $(k-s)\varepsilon$ to some of the values, the numerator and the denominator of (3) become polynomials in ε, both without constant term. The term with the lowest degree in the numerator has degree γ_p equal to the number of zeroes before adding/subtracting ε. The term with the lowest degree in the denominator has degree λ_p at most equal to γ_p. So, the limit of of M_p^l (when ε goes to zero) exists; it is either zero (if $\gamma_i > \lambda_i$) or positive (if $\gamma_i = \lambda_i$). So, when we use the logratio mean, there is no need to restrict ourselves to strictly positive compositions.

The graph median mean M^g is defined as follows. We first search all minimum spanning trees of $(x^1,\ldots,x^n)$ induced by the half-taxi metric. For each such tree, we identify the graph median (a single data point or a pair). Since there can be several minimum spanning trees (Sharp [2006] did not mention this) and each one can yield one or two graph medians, there can be many points satisfying the definition of the graph median given by Sharp. We therefore need to complement this definition by a tie-breaking rule: if several data points are possible graph medians, use their arithmetic mean as central tendency $M^g(x^1,\ldots,x^n)$.

We now present some conditions (or axioms) that, we think, should be satisfied by any mapping M representing the central tendency. We begin with a very compelling axiom.

A 1 Idempotence. *For all $x \in X$, $M(x,x,\ldots,x) = x$.*

In other words, if the composition of our n datapoints is the same, then the central tendency should be equal to that composition. The arithmetic mean, the logratio mean and the graph median satisfy Idempotence. The next axiom is about the role played by the different points.

A 2 Symmetry. *For any $n \in \mathbb{N}_0$, any permutation σ of $N(n)$ and any data set $(x^1,\ldots,x^n)$, $M(x^1,\ldots,x^n) = M(x^{\sigma(1)},\ldots,x^{\sigma(n)})$.*

The effect of the permutation σ in the statement of this axiom is simple: it just reorders the data points. So, our mapping satisfies Symmetry if and only if the order in which we collect or store our data does not matter: all data points play the same role. The arithmetic mean, the logratio mean and the graph median satisfy Symmetry. Let us now look at what happens when a point in our data set changes.

A 3 Strict Monotonicity. *For all* $i \in N(n)$, $x_i^m > x_i'^m$ *implies* $M_i(x^1, \ldots, x^m, \ldots, x^n) > M_i(x^1, \ldots, x'^m, \ldots, x^n)$.

Suppose we observe a data set $(x^1, x^2, \ldots, x^m, \ldots, x^n)$ and we compute $M(x^1, x^2, \ldots, x^m, \ldots, x^n)$. We then realize that the mth data point is wrong: it should be x'^m instead of x^m. Suppose also $x_i^m > x_i'^m$. We then recompute the central tendency with the correct composition. Since $x_i^m > x_i'^m$, we obviously expect the central tendency to decrease in the ith component. That is precisely what Strict Monotonicity expresses.

The arithmetic mean and the logratio mean satisfy Monotonicity. The graph median does not. Indeed, consider the following data set: $x^1 = (0.6, 0.2, 0.2)$, $x^2 = (0.2, 0.6, 0.2)$ and $x^3 = (0.2, 0.2, 0.6)$. In this case, there are three possible minimum spanning trees: the first one is the set $\{(x^1, x^2), (x^2, x^3)\}$, the second one $\{(x^1, x^2), (x^1, x^3)\}$ and the last one $\{(x^1, x^3), (x^2, x^3)\}$. Each of these minimum spanning trees leads to a different median: x^2, x^1 and x^3 respectively. So, using our tie-breaking rule: we choose $(1/3, 1/3, 1/3)$ as central tendency. Suppose now we slightly modify x^2 so as to obtain $x'^2 = (0.2+\varepsilon, 0.6-2\varepsilon, 0.2+\varepsilon)$ with $\varepsilon > 0$. The unique minimum spanning tree is then $\{(x^1, x'^2), (x'^2, x^3)\}$ with x'^2 as median. So, although $x_2'^2 < x_2^2$ decreases, $M_2(x^1, x'^2, x^3) > M_2(x^1, x^2, x^3)$. The next condition is very similar but focuses on the components that do not change from x^m to x'^m.

A 4 Consistency. *For all* $i \in N(n)$, $x_i^m = x_i'^m$ *implies* $M_i(x^1, \ldots, x^m, \ldots, x^n) = M_i(x^1, \ldots, x'^m, \ldots, x^n)$.

Since $x_i^m = x_i'^m$, we might expect the central tendency to remain the same in the ith component. Notice that this conditon is not satisfied by the logratio mean, as shown by Shurtz [2000]. It is also violated by the graph median (use $x^1 = (0.6, 0.2, 0.2)$, $x^2 = (0.2, 0.6, 0.2)$, $x^3 = (0.2, 0.2, 0.6)$ and $x'^2 = (0.2 + \varepsilon, 0.6 - \varepsilon, 0.2)$ to check this) but it is satisfied by the arithmetic mean. We now turn to an axiom stating that any m data points can be replaced by their mean.

A 5 Decomposability. *Let* $\alpha = M(x^1, \ldots, x^m)$. *Then* $M(x^1, \ldots, x^m, x^{m+1} \ldots, x^n) = M(\underbrace{\alpha, \ldots, \alpha}_{m \text{ times}}, x^{m+1} \ldots, x^n)$.

In other words, since α summarizes the central tendency of the points $x^1, \ldots, x^m$, we can replace each of them by α when we want to compute the central tendency of $x^1, \ldots, x^m, x^{m+1} \ldots, x^n$. The arithmetic mean and the logratio mean satisfy Decomposability but the graph median does not. Indeed, let $x^1 = (0.6, 0.2, 0.2)$, $x^2 = (0.3, 0.4, 0.3)$, $x^3 = (0.2, 0.2, 0.6)$. We have $M^g(x^1, x^2, x^3) = x^2$ and $M^g(x^1, x^3) = (0.4, 0.2, 0.4) = \alpha$ but $M(\alpha, x^2, \alpha) = \alpha \neq x^2$. We now conclude this section with a stability condition.

A 6 Continuity. *M is a continuous function of each of its arguments.*

This standard condition implies that the central tendency can not make an abrupt change when we only slightly modify the composition of a data point.

It is easy to check that the arithmetic mean satisfies Continuity. The logratio mean is almost continuous: the only data sets where it is not continuous are those such that, for each component, there is at least one data point with a zero coordinate. The graph median exhibits a lot of discontinuities. This can easily be seen using the same example as the one used for proving that the graph median is not strictly monotonic. We have $M^g(x^1, x^2, x^3) = (1/3, 1/3, 1/3)$ while $M^g(x^1, x'^2, x^3) = x'^2$ for any $\varepsilon > 0$ even if ε is very small.

3 Result and proofs

In this section, we prove that only one mapping M (the arithmetic mean) satisfies all axioms presented so far.

Theorem 1 *Assume $k \geq 3$. The mapping M satisfies Idempotence, Symmetry, Strict Monotonicity, Decomposability, Consistency and Continuity if and only if, for every $p \in K$,*

$$M_p(x^1, \ldots, x^n) = \frac{x_p^1 + \ldots + x_p^n}{n}. \quad (4)$$

3.1 Proofs

First a few lemmas. Given $p, q \in K$, let K_{pq} denote $K \setminus \{p, q\}$.

Lemma 1 *Let us assume Idempotence, Symmetry, Strict Monotonicity, Decomposability and Continuity. Consider n points $x^1, \ldots, x^n$ and suppose $x_i^l = 0$ for all $i \in K_{pq}$ and all $l \in N(n)$. Then*

$$M_p(x^1, \ldots, x^n) = f_p^{-1}\left(\frac{f_p(x_p^1) + \ldots + f_p(x_p^n)}{n}\right)$$

for some continuous and strictly monotonic mapping $f : [0,1] \to \mathbb{R}$.

Proof. For each l, we have $x_p^l + x_q^l = 1$. So, the composition of each point x^l is completely described by x_p^l and, consequently, $M(x^1, \ldots, x^n) = F(x_p^1, \ldots, x_p^n)$ for some $F : [0,1]^n \to X$. In particular, $M_p(x^1, \ldots, x^n) = F_p(x_p^1, \ldots, x_p^n)$ for some $F_p : [0,1]^n \to [0,1]$. Notice that, for each $l \in N(n)$, x_p^l can vary in the same interval $[0,1]$. It is easily seen that the four conditions used by Kolmogoroff [1930] are satisfied by F_p (they are straightforward consequences of Idempotence, Symmetry, Strict Monotonicity, Decomposability and Continuity). We can therefore apply Kolmogoroff's Theorem and we find

$$F_p(x_p^1, \ldots, x_p^n) = f_p^{-1}\left(\frac{f_p(x_p^1) + \ldots + f_p(x_p^n)}{n}\right) \quad (5)$$

for some continuous and strictly monotonic mapping $f_p : [0,1] \to \mathbb{R}$. $\square$

This lemma tells us what M_p is when all points have zero coordinates on all components but two. The next lemma generalizes this to arbirary data sets.

Lemma 2 *Let us assume Idempotence, Symmetry, Strict Monotonicity, Consistency, Decomposability and Continuity. Consider n points $x^1, \ldots, x^n$. Then, for every $p \in K$,*

$$M_p(x^1, \ldots, x^n) = f_p^{-1}\left(\frac{f_p(x_p^1) + \ldots + f_p(x_p^n)}{n}\right) \tag{6}$$

for some continuous and strictly monotonic mapping $f_p : [0,1] \to \mathbb{R}$.

Proof. Consider n points $y^1, \ldots, y^n$ such that, for every $l \in N(n)$, $y_p^l = x_p^l$ and $y_i^l = 0 \, \forall i \in K_{pq}$. By Consistency, $M_p(x^1, \ldots, x^n) = M_p(y^1, \ldots, y^n)$ and by Lemma 1, $M_p(x^1, \ldots, x^n) = f_p^{-1}\left(\frac{f_p(x_p^1)+\ldots+f_p(x_p^n)}{n}\right)$. $\square$

The form obtained for M_p in this lemma is still very general. Indeed, if $f_p(u) = u$ in (5), then F_p is the arithmetic mean; if $f_p(u) = 1/u$, then F_p is the harmonic mean; if $f_p(u) = \ln u$, then F_p is the geometric mean; if $f_p(u) = u^\alpha$ (for some real α), then F_p is the power mean; and there are many other possible choices for f_p. In the next result, we will show that, actually, we do not have much freedom in the choice of f_p.

Lemma 3 *Assume $k \geq 3$. Assume also Idempotence, Symmetry, Strict Monotonicity, Decomposability, Consistency and Continuity. Then, for every $p \in K$,*

$$M_p(x^1, \ldots, x^n) = f^{-1}\left(\frac{f(x_p^1) + \ldots + f(x_p^n)}{n}\right) \tag{7}$$

for some continuous and strictly monotonic mapping $f : [0,1] \to [0,1]$ such that $f(0) = 0$ and, for every $t \in [0,1], f(t) = 1 - f(1-t)$.

Proof. Choose any $p, q \in K$. Consider two points x, y with $x_i = 0 = y_i$ for all $i \in K_{pq}$. Using Lemma 2, we find $M_i(x,y) = 0$ for all $i \in K_{pq}$, $M_p(x,y) = f_p^{-1}\left(\frac{f_p(x_p)+f_p(y_p)}{2}\right)$ and $M_q(x,y) = f_q^{-1}\left(\frac{f_q(x_q)+f_q(y_q)}{2}\right)$. Since $\sum_{i \in K} M_i(x,y) = 1$, $y_q = 1 - y_p$ and $x_q = 1 - x_p$, we have

$$f_p^{-1}\left(\frac{f_p(x_p) + f_p(y_p)}{2}\right) + f_q^{-1}\left(\frac{f_q(1-x_p) + f_q(1-y_p)}{2}\right) = 1 \tag{8}$$

for all $x_p, y_p \in [0,1]$. Let us solve this functional equation with two unknowns: f_p and f_q. Clearly,

$$f_q^{-1}\left(\frac{f_q(1-x_p) + f_q(1-y_p)}{2}\right) = 1 - f_p^{-1}\left(\frac{f_p(x_p) + f_p(y_p)}{2}\right)$$

and
$$\frac{f_q(1-x_p)+f_q(1-y_p)}{2} = f_q\left(1 - f_p^{-1}\left(\frac{f_p(x_p)+f_p(y_p)}{2}\right)\right).$$

Let $u = f_p(x_p)$, $v = f_p(y_p)$ and $\phi(u) = f_q(1-f_p^{-1}(u))$. We can then rewrite the previous equation as
$$\frac{\phi(u)+\phi(v)}{2} = \phi\left(\frac{u+v}{2}\right).$$

This is the well-known Jensen's equation [Aczél, 1966, p.43]. Since f_p and f_q are continuous, ϕ is also continuous and Jensen's equation has then as unique solution $\phi(u) = au + b$ for some real-valued a, b. So, $f_q(1 - f_p^{-1}(u)) = au + b$. This proves that
$$f_p(u) = af_q(1-u) + b. \tag{9}$$

Since our reasoning holds for any pair of dimensions and not just p, q, we also have $f_p(u) = a'f_r(1-u) + b'$ and $f_r(u) = a''f_q(1-u) + b''$. So, $f_p(u) = a'(a''f_q(u)+b'')+b' = a'''f_q(u)+b'''$. This proves that f_p and f_q are related by an affine transformation. It is simple to check that using f_p or an affine transformation thereof in (6) does not matter. So, we can use f_p or f_q or any other f_i since our reasoning holds for any pair of dimensions and not just p, q. Pick any of them and call it f. We have then $M_i(x^1, \ldots, x^n) = f^{-1}\left(\frac{f(x_i^1)+\ldots+f(x_i^n)}{n}\right)$, for any $i \in K$, as in (7). Equation (9) can therefore be rewritten as
$$f(1-t) = af(t) + b. \tag{10}$$

Set $t = 0$; then $f(1) = af(0) + b$. Set $t = 1$; then $f(0) = af(1) + b$. Hence, $f(1) - af(0) = b = f(0) - af(1)$ or $f(1)(1+a) = f(0)(1+a)$. This equation has two solutions: (i) $f(1) = f(0)$ or (ii) $1 + a = 0$. The first one is impossible for f is strictly increasing. So, $a = -1$ and, since $f(0) = af(1) + b$, we have $b = f(0) + f(1)$. Let us replace in (10):$f(1-t) = -f(t) + f(0) + f(1)$. Set $t = 0.5$ and we obtain $f(0.5) = \frac{f(0)+f(1)}{2}$. As we have seen previously, using f_p or an affine transformation thereof in (6) does not matter. So, we can always arbitrarily set $f(0) = 0$ and $f(1) = 1$. Then $f(0.5) = 0.5$ and, for every $t \in [0,1]$, $f(t) = 1 - f(1-t)$. □

We are now ready to prove our main result.

Proof of Theorem 1. Necessity of the conditions. It is simple to check that (4) indeed satisfies Idempotence, Symmetry, Strict Monotonicity, Decomposability, Consistency and Continuity.

Sufficiency of the conditions. Choose any $p, q, r \in K$. Consider two points x, y with $x_i = 0 = y_i$ for all $i \in K_{pqr}$. Using Lemma 3, we find $M_i(x, y) = 0$ for all $i \in K_{pqr}$, $M_p(x,y) = f^{-1}\left(\frac{f(x_p)+f(y_p)}{2}\right)$, $M_q(x,y) = f^{-1}\left(\frac{f(x_q)+f(y_q)}{2}\right)$, and $M_r(x,y) = f^{-1}\left(\frac{f(x_r)+f(y_r)}{2}\right)$. Since $\sum_{i\in K} M_i(x,y) = 1$, $y_r = 1 - y_p - y_q$ and

$x_r = 1 - x_p - x_q$, we have

$$f^{-1}\left(\frac{f(x_p)+f(y_p)}{2}\right) + f^{-1}\left(\frac{f(x_q)+f(y_q)}{2}\right) + f^{-1}\left(\frac{f(1-x_p-x_q)+f(1-y_p-y_q)}{2}\right) = 1. \quad (11)$$

Let us freeze y and define $g(x_p) = f^{-1}\left(\frac{f(x_p)+f(y_p)}{2}\right)$, $h(x_q) = f^{-1}\left(\frac{f(x_q)+f(y_q)}{2}\right)$ and $l(x_p+x_q) = 1 - f^{-1}\left(\frac{f(1-x_p-x_q)+f(1-y_p-y_q)}{2}\right)$. We then have

$$g(x_p) + h(x_q) = l(x_p + x_q) \quad (12)$$

for all x_p, x_q such that $x_p + x_q \leq 1$. In particular, it holds for all $x_p, x_q \in [0, 1/2]$. Equation (12) is a Pexider functional equation [Aczél, 1966] and has as unique solution (on the interval $[0, 1/2]$) $g(u) = au + b$ for some positive real number a. So, $f^{-1}\left(\frac{f(x_p)+f(y_p)}{2}\right) = a(y_p)x_p + b(y_p)$. Since the left-hand side of this equation is symmetric in x_p and y_p, the right-hand side must also be symmetric. Hence, $a(y_p)x_p + b(y_p) = a(x_p)y_p + b(x_p)$ for all $x_p, x_q \in [0, 1/2]$. Setting $x_p = 0$ yields $b(y_p) = a(0)y_p + b(0)$. Setting $x_p = 1/2$ yields $a(y_p)/2 + a(0)y_p + b(0) = a(1/2)y_p + b(0)$ and, after some elementary algebra, $a(y_p) = 2(a(1/2) - a(0))y_p$. Let $\alpha = 2(a(1/2) - a(0))$ and $\beta = b(1/2) - b(0)$. Then $a(y_p) = \alpha y_p + \beta$. So, for all $x_p, x_q \in [0, 1/2]$, we have $f^{-1}\left(\frac{f(x_p)+f(y_p)}{2}\right) = (\alpha y_p + \beta)x_p + a(0)y_p + b(0) = \alpha y_p x_p + \beta x_p + a(0)y_p + b(0)$. Using again a symmetry argument, we find $\beta = a(0)$. By Idempotence, for all $u \in [0, 1/2]$, we have $f^{-1}\left(\frac{f(u)+f(u)}{2}\right) = u = \alpha u^2 + 2\beta u + b(0)$. This is possible only if $b(0) = 0 = \alpha$ and $\beta = 1/2$. In other words, for all $x_p, x_q \in [0, 1/2]$, we have $f^{-1}\left(\frac{f(x_p)+f(y_p)}{2}\right) = (x_p + y_p)/2$. This is once more Jensen's functional equation, with as unique solution $f(u) = cu + d$ for all $u \in [0, 1/2]$. By Lemma 3, $f(u) = 1 - f(1-u)$ for all $u \in [0, 1]$. Choose any $u \in [1/2, 1]$ so that $1 - u \in [0, 1/2]$. Then $f(u) = 1 - f(1-u) = 1 - c(1-u) - d = cu + 1 - c - d$. Since $f(1) = 1$, we find $d = 0$. Since $1/2$ belongs to $[0, 1/2]$ and $[1/2, 1]$, we have $f(1/2) = c/2 = c/2 + 1 - c$. This implies $c = 1$. Summarizing, we have $f(u) = u$ for all $x \in [0, 1]$ and, hence, $M_p(x^1, \ldots, x^n) = \frac{x_p^1 + \ldots + x_p^n}{n}$, for all $(x^1, \ldots, x^n) \in \mathcal{X}$. □

4 Discussion

Theorem 1 characterizes the arithmetic mean by a set of very compelling conditions. Our result therefore provides a justification for the arithmetic mean.

The graph median violates most of the axioms used in Theorem 1. It therefore seems reasonable to discard the graph median as a potential measure of central tendency. The logratio mean violates Consistency and, only in very special configurations, Continuity. Since Consistency is probably the least compelling of the conditions of Theorem 1, our result is not a clear point against the logratio mean. An axiomatic characterization of the logratio mean would be useful for comparing with the arithmetic mean.

References

J. Aczél. *Lectures on functional equations and their applications.* Academic Press, 1966.

J. Aitchison. *The statistical analysis of compositional data.* Chapman and Hall, 1986.

A. Kolmogoroff. Sur la notion de la moyenne. *Atti delle Reale Accademia Nazionale dei Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez.*, 12:323–343, 1930.

W. E. Sharp. The graph median—a stable alternative measure of central tendency for compositional data sets. *Mathematical Geology*, 38:221–229, 2006.

R. F. Shurtz. Comment on "Logratios and natural laws in compositional data analysis" by J. Aitchison. *Mathematical Geology*, 32:645–647, 2000.

A. Woronow. Regression and discrimination analysis using raw compositional data—is it really a problem? In *Proceedings of IAMG97, The Third Annual Conference of the International Association for Mathematical Geology: Universitat Politècnica de Catalunya, Barcelona*, pages 129–132, 1997.