Axiomatic characterization of the χ^2 dissimilarity measure

Denis Bouyssou, Thierry Marchant and Marc Pirlot

Abstract. We axiomatically characterize the χ^2 dissimilarity measure. To this end, we solve a new generalization of a functional equation discussed in Aczel (Lectures on functional equations and their applications, Academic Press, 1966).

Mathematics Subject Classification (2010). Primary 39B22; Secondary 62G10.

Keywords. Chi squared, dissimilarity measure.

1. Introduction

Let $N = \{1, 2, \ldots, n\}$ be a set of categories (with $n \geq 2$). The vector $x = (x_1, \ldots, x_n)$ represents the respective numbers of observations in each category and the total number of observations is denoted by $s(x) = \sum_{i \in N} x_i$. We want to measure the dissimilarity between the observed distribution x and a reference distribution $\pi = (\pi_1, \ldots, \pi_n)$, with $\sum_{i \in N} \pi_i = 1$ and $\pi_i \in \mathbb{Q}_{++}$ for all $i \in N$, where $\mathbb{Q}_{++}$ is the set of positive rational numbers. We exclude reference distributions with null components because the χ^2 dissimilarity measure is not defined when a component is zero. The set of all observed distributions is $X = \mathbb{N}^N$, i.e. the set of all mappings from N to $\mathbb{N}_0$, where $\mathbb{N}_0$ is the set of non-negative integers. The set Π of all reference distributions is defined by $\Pi = \{\pi \in \mathbb{Q}^N_{++} : \sum_{i \in N} \pi_i = 1\}$.

A dissimilarity measure f is a mapping from $X \times \Pi$ to $\mathbb{R}_+$ (the set of non-negative real numbers) satisfying $f(x, \pi) = 0$ iff $x/s(x) = \pi$. It measures how far the observed distribution is from the reference. In this paper, we axiomatically characterize the χ^2_1 dissimilarity measure defined by

$$\chi^2_1(x, \pi) = \sum_{i \in N} \frac{(s(x)\pi_i - x_i)^2}{s(x)\pi_i}$$

Authors are listed alphabetically and have contributed equally.

and frequently used in statistics as a measure of goodness of fit.

The dissimilarity measure χ_0^2 defined by $\chi_0^2(x, \pi) = \chi_1^2(x, \pi)/s(x)$ has been characterized in [7] and we will also provide a new characterization thereof. It is popular in ecology [6], sociology [8], economics [9], and so on.

While we consider in our paper that the number n of categories is given and fixed, [7] considers that n can vary. Depending on the context, one or the other assumption can be more relevant. For instance, when we use Pearson's χ^2 test, we have a sample distributed over n categories and the p-value is computed conditional on a theoretical probability distribution with the same number n of categories. If we repeat the experiment and draw other samples, we obtain other p-values always based on the same theoretical probability distribution with the same number n of categories. It therefore makes sense to consider n as given.

A common feature of [7] and our paper is that we use a framework in which π can vary and such that comparisons of the dissimilarity measure across different reference distributions are relevant. Yet, unlike [7], we also consider the case in which the reference distribution π is fixed (as in our Pearson's χ^2 example).

For characterizations of other dissimilarity measures, in the context of political sciences, see [3]. See [2] for a characterization of a wide class of dissimilarity measures. While we consider dissimilarity measures in this paper, it is also interesting to consider dissimilarity rankings as in [4].

Section 2 presents our main conditions and results. Section 3 shows the independence of the conditions used in our results. Section 4 concludes the discussion. All the proofs are gathered in Section 5.

2. Axioms and results

The dissimilarity measures χ_0^2 and χ_1^2 are homogeneous of degree 0 and 1, respectively, where homogeneity is defined as follows.

A 1. Homogeneity of degree ω. For all positive integers λ and $x \in X$, $f(\lambda x, \pi) = \lambda^\omega f(x, \pi)$.

In statistics, it seems unanimously accepted that a dissimilarity measure (used as a goodness-of-fit statistic) should be homogeneous of degree 1, but in ecology, many researchers seem to favour homogeneity of degree 0. Indeed, when they measure the dissimilarity between the species distribution in an ecosystem and a reference distribution, they want the dissimilarity to be independent of the size of the ecosystem. It is easy to see that Homogeneity of degree 0 (resp. 1) is satisfied by χ_0^2 (resp. χ_1^2). Indeed, we have

$$\chi_0^2(\lambda x, \pi) = \sum_{i \in N} \frac{(\pi_i - \lambda x_i / s(\lambda x))^2}{\pi_i} = \sum_{i \in N} \frac{(\pi_i - x_i / s(x))^2}{\pi_i} = \chi_0^2(x, \pi)$$

and

$$\chi_1^2(\lambda x, \pi) = \sum_{i \in N} \frac{(s(\lambda x)\pi_i - \lambda x_i)^2}{s(\lambda x)\pi_i} = \lambda \sum_{i \in N} \frac{(s(x)\pi_i - x_i)^2}{s(x)\pi_i} = \lambda \chi_1^2(x, \pi).$$

Suppose the dissimilarity between a distribution x and π is zero. This implies $x = k\pi$ for some positive integer k. The next condition states that, when we modify $k\pi$ by moving a single individual from category l to j, then the dissimilarity measure is inversely proportional to the harmonic mean of π_j and π_l. Let $\mathbf{1}^i \in X$ be a vector such that $\mathbf{1}^i_i = 1$ and $\mathbf{1}^i_j = 0$ for all $j \neq i$.

A 2. Inverse Effects. If $k\pi, k\pi' \in X$, then, for all $j, l, r, s \in N$, with $j \neq l$ and $r \neq s$,

$$\frac{f(k\pi + \mathbf{1}^j - \mathbf{1}^l, \pi)}{f(k\pi' + \mathbf{1}^r - \mathbf{1}^s, \pi')} = \frac{\frac{1}{\pi_j} + \frac{1}{\pi_l}}{\frac{1}{\pi'_r} + \frac{1}{\pi'_s}}.$$

In our first result, we will use a restricted variant of Inverse Effects in which $\pi = \pi'$. This weaker condition is named Restricted Inverse Effects and is trivially satisfied when $n = 2$. We now prove that Inverse Effects is satisfied by χ_0^2:

$$\chi_0^2(k\pi + \mathbf{1}^j - \mathbf{1}^l, \pi) = \frac{(-1/k)^2}{\pi_j} + \frac{(1/k)^2}{\pi_l} = \frac{1}{k^2}\left(\frac{1}{\pi_j} + \frac{1}{\pi_l}\right).$$

The proof for χ_1^2 is similar.

Let x and y be two observed distributions of size k. The deviation between x and $k\pi$ is $x - k\pi$. The corresponding deviation for y is $y - k\pi$. If we add these two vectors of deviations, we obtain $x + y - 2k\pi$ and the corresponding observed distribution is $x + y - 2k\pi + k\pi = x + y - k\pi$ (provided all components are non-negative). Hence, $f(x + y - k\pi, \pi)$ represents the dissimilarity corresponding to the additive combination of two deviations: between x (resp. y) and $k\pi$. Similarly, $f(x - y + k\pi, \pi)$ corresponds to the subtractive combination of the same two deviations. Finally, $f(x + y - k\pi, \pi) + f(x - y + k\pi, \pi)$ corresponds in some sense to four deviations (two x- and two y-deviations) combined once additively and once subtractively. Our next condition states that this must be equal to $2f(x, \pi) + 2f(y, \pi)$, which is another way to combine the same four deviations.

A 3. Deviations Balancedness. For all $x, y \in X$ with $s(x) = s(y) = k$, if $x + y - k\pi \in X$ and $x - y + k\pi \in X$, then

$$f(x + y - k\pi, \pi) + f(x - y + k\pi, \pi) = 2\big(f(x, \pi) + f(y, \pi)\big).$$

This condition is inspired by [5], in which they characterize the Euclidean distance in $\mathbb{R}^n$. Let us prove that χ_1^2 satisfies Deviations Balancedness.

We have

$$\begin{aligned}\chi_1^2(x+y-k\pi,\pi) &= \sum_{i\in N}\frac{\big(s(x+y-k\pi)\pi_i-(x_i+y_i-k\pi_i)\big)^2}{s(x+y-k\pi)\pi_i}\\ &= \sum_{i\in N}\frac{(2k\pi_i-x_i-y_i)^2}{k\pi_i}\end{aligned}$$

and

$$\begin{aligned}\chi_1^2(x-y+k\pi,\pi) &= \sum_{i\in N}\frac{\big(s(x-y+k\pi)\pi_i-(x_i-y_i+k\pi_i)\big)^2}{s(x-y+k\pi)\pi_i}\\ &= \sum_{i\in N}\frac{(x_i-y_i)^2}{k\pi_i}.\end{aligned}$$

Hence, $\chi_1^2(x+y-k\pi,\pi)+\chi_1^2(x-y+k\pi,\pi)$ is equal to

$$\begin{aligned}&\sum_{i\in N}\frac{(2k\pi_i-x_i-y_i)^2}{k\pi_i}+\sum_{i\in N}\frac{(x_i-y_i)^2}{k\pi_i}\\ =&\sum_{i\in N}\frac{2(k^2\pi_i^2+x_i^2-2k\pi_ix_i)+2(k^2\pi_i^2+y_i^2-2k\pi_iy_i)}{k\pi_i}\\ =&2\chi_1^2(x,\pi)+2\chi_1^2(y,\pi).\end{aligned}$$

We are now ready to state our first result in which we consider that π is given and does not vary.

Theorem 2.1. *Assume π is given. For $\omega\in\{0,1\}$, a dissimilarity measure f satisfies Homogeneity of degree ω, Deviations Balancedness and Restricted Inverse Effects iff $f=\gamma\chi_\omega^2$, for some positive $\gamma\in\mathbb{R}$. Restricted Inverse Effects is not required when $n=2$.*

Notice that Theorem 2.1 does not hold when π is not fixed. Indeed, for any $\phi:\Pi\to\mathbb{R}_+$ with ϕ not constant, the dissimilarity measure

$$f_\phi(x,\pi)=\phi(\pi)\sum_{i\in N}\frac{(\pi_i-x_i/s(x))^2}{\pi_i}$$

satisfies Homogeneity of degree 1, Deviations Balancedness and Restricted Inverse Effects but is not of the form $\gamma\chi_0^2$ or $\gamma\chi_1^2$. In order to characterize the dissimilarity measure χ^2 when π varies, we need the full power of Inverse Effects.

Theorem 2.2. *For $\omega\in\{0,1\}$, a dissimilarity measure f satisfies Homogeneity of degree ω, Deviations Balancedness and Inverse Effects iff $f=\gamma\chi_\omega^2$, for some positive $\gamma\in\mathbb{R}$.*

3. Independence of the axioms

In order to prove the independence of the conditions characterizing χ_0^2 with variable π , we provide three examples of dissimilarity measures violating only one of the three conditions in Theorem 2.2.

The dissimilarity measure χ_1^2 violates Homogeneity of degree 0 but satisfies Deviations Balancedness and Inverse Effects. The dissimilarity measure

$$f(x,\pi) = \sum_{i \in N} \frac{|\pi_i - x_i/s(x)|}{\pi_i}$$

violates Deviations Balancedness but satisfies Homogeneity of degree 0 and Inverse Effects. The dissimilarity measure

$$f(x,\pi) = \sum_{i \in N} (\pi_i - x_i/s(x))^2$$

violates Inverse Effects but satisfies Homogeneity of degree 0 and Deviations Balancedness.

Our examples are easily adapted to prove the independence of the conditions characterizing χ_1^2 with variable π . Finally, our examples can also be used for Theorem 2.1 since it involves the same conditions as Theorem 2.2 except for Restricted Inverse Effects which is weaker than Inverse Effects.

4. Discussion

Theorems 2.1 and 2.2 characterize the dissimilarity measures χ_0^2 and χ_1^2 up to a multiplication by a positive real number γ. We could easily add a condition characterizing exactly χ_0^2 or χ_1^2. For instance, the extra condition $f\big(\mathbf{1}^1, (1/n, \ldots, 1/n)\big) = n - 1$ is enough to force $\gamma = 1$ in both characterizations. Yet, unlike [7], we consider that such a normalization is not really interesting. Indeed χ_1^2 and $\gamma\chi_1^2$ (with $\gamma \neq 1$) convey exactly the same information, just like a distance measurement in meters or yards. In particular, if we want to perform a Pearson's χ^2 test, we are free to use Pearson's statistic (i.e. χ_1^2) and to compute the p-value using the χ^2 density or to use $\gamma\chi_1^2$ (with an arbitrary γ) and to compute the p-value using the corresponding density. The resulting p-value will of course be identical. The same holds for χ_0^2 and $\gamma\chi_0^2$.

5. Proofs

We need a few lemmas before proving Theorem 2.1.

Lemma 1. Let $f(x,\pi) = s(x)f'(x,\pi)$. Then f satisfies Homogeneity of degree 1 iff f' satisfies Homogeneity of degree 0. And f satisifies Deviations Balancedness (resp. Inverse Effects) iff f' satisfies Deviations Balancedness (resp. Inverse Effects).

Proof. Since f satisfies Homogeneity of degree 1, we have $f(\lambda x, \pi) = \lambda f(x, \pi)$ for all positive integers λ. We thus have $\lambda s(x) f'(\lambda x, \pi) = \lambda s(x) f'(x, \pi)$. Hence $f'(\lambda x, \pi) = f'(x, \pi)$ and f' is homogeneous of degree 0. The proof of the reverse implication is similar. The rest of the proof is left to the reader. □

Lemma 2. Suppose π is fixed. If a dissimilarity measure f satisfies Homogeneity of degree 0, then $f(x, \pi) = F(x/s(x))$, for some mapping $F : \Pi \to \mathbb{R}_+$.

Proof. Since π is fixed, we can define a mapping $g : X \to \mathbb{R}_+$ such that $f(x, \pi) = g(x)$. Define now the mapping $F : \Pi \to \mathbb{R}_+$ as follows. For any $p \in \Pi$, $F(p) = g(x)$ if there is $x \in X$ such that $p = x/s(x)$. The mapping F is defined everywhere because p has rational components and, hence, there is always $x \in X$ such that $p = x/s(x)$. The mapping F is well defined. Indeed, suppose now there are x, y such that $p = x/s(x)$ and $p = y/s(y)$. By Homogeneity of degree 0, $f(x, \pi) = f(y, \pi)$. Therefore, $F(p) = g(x) = g(y)$. □

We say that a set S in $\mathbb{Q}^k$ is rational convex if whenever $u, v \in S$, then $\alpha u + (1-\alpha)v \in S$ for all rational $\alpha \in [0, 1]$.

Lemma 3. Let S be a rational convex subset of $\mathbb{Q}^2$ such that S is full-dimensional. Let $g : S \to \mathbb{R}_+$ be a mapping such that the graph of g is a parabola on any line segment $r \subset S$. Then $g(u, v) = \rho u^2 + \sigma v^2 + \tau uv + \mu u + \nu v + \xi$ for some real $\rho, \sigma, \tau, \mu, \nu, \xi$.

Proof. Since S is full-dimensional, the interior of S is not empty and we can suppose without loss of generality that $(0, 0) \in \operatorname{int} S$. Let us consider the line defined by $(\alpha t, (1-\alpha)t)$ for some $\alpha \in \mathbb{Q}$ and all $t \in \mathbb{Q}$. The intersection of this line with S defines a line segment r_α passing by the origin. The graph of g on r_α is a parabola. We can express this by means of the following polynomial of degree 2 in t:

$$g(\alpha t, (1-\alpha)t) = k_\alpha t^2 + l_\alpha t + m_\alpha, \tag{5.1}$$

where k_α, l_α and m_α are real numbers.

Let us now consider the line defined by $(\alpha t, (1-\alpha)t)$ for some $t \in \mathbb{Q}$ and all $\alpha \in \mathbb{Q}$. The intersection of this line with S defines a line segment s_t. We can express that the graph of g on s_t is a parabola by means of a polynomial of degree 2 in α:

$$g(\alpha t, (1-\alpha)t) = \alpha^2 \beta_t + \alpha \gamma_t + \delta_t, \tag{5.2}$$

where β_t, γ_t and δ_t are real numbers. Setting $t = 0$ in (5.2) yields, $g(0, 0) = \alpha^2 \beta_0 + \alpha \gamma_0 + \delta_0$. Since this must be true for all α, we must have $\beta_0 = \gamma_0 = 0$.

Equating (5.1) and (5.2) yields

$$k_\alpha t^2 + l_\alpha t + m_\alpha = \alpha^2 \beta_t + \alpha \gamma_t + \delta_t. \tag{5.3}$$

Setting $t = 0$, $t = 1$ and $t = 2$ in (5.3) yields

$$\begin{aligned} m_\alpha &= \delta_0 \\ k_\alpha + l_\alpha + m_\alpha &= \alpha^2 \beta_1 + \alpha \gamma_1 + \delta_1 \\ 4k_\alpha + 2l_\alpha + m_\alpha &= \alpha^2 \beta_2 + \alpha \gamma_2 + \delta_2. \end{aligned}$$

The solution of this system is

$$\begin{aligned} m_\alpha &= \delta_0 \\ k_\alpha &= \alpha^2 \, \frac{\beta_2 - 2\beta_1}{2} + \alpha \frac{\gamma_2 - 2\gamma_1}{2} + \frac{\delta_2 - 2\delta_1 + \delta_0}{2} \\ l_\alpha &= \alpha^2 \, \frac{4\beta_1 - \beta_2}{2} + \alpha \, \frac{4\gamma_1 - \gamma_2}{2} + \frac{4\delta_1 - \delta_2 - 3\delta_0}{2}. \end{aligned}$$

Let us rewrite (5.1):

$$\begin{aligned} g(\alpha t, (1-\alpha)t) &= \left(\alpha^2 \, \frac{\beta_2 - 2\beta_1}{2} + \alpha \, \frac{\gamma_2 - 2\gamma_1}{2} + \frac{\delta_2 - 2\delta_1 + \delta_0}{2}\right) t^2 \\ &\quad + \left(\alpha^2 \, \frac{4\beta_1 - \beta_2}{2} + \alpha \, \frac{4\gamma_1 - \gamma_2}{2} + \frac{4\delta_1 - \delta_2 - 3\delta_0}{2}\right) t + \delta_0. \end{aligned}$$

Letting $\alpha t = u$, $(1-\alpha)t = v$, $\alpha = u/(u+v)$ and $t = u+v$, we find that $g(u,v)$ is equal to

$$\begin{aligned} & u^2 \, \frac{\beta_2 - 2\beta_1}{2} + u(u+v) \, \frac{\gamma_2 - 2\gamma_1}{2} + (u+v)^2 \, \frac{\delta_2 - 2\delta_1 + \delta_0}{2} \\ = \; & + \frac{u^2}{u+v} \, \frac{4\beta_1 - \beta_2}{2} + u \, \frac{4\gamma_1 - \gamma_2}{2} + (u+v) \, \frac{4\delta_1 - \delta_2 - 3\delta_0}{2} + \delta_0. \quad (5.4) \end{aligned}$$

The graph of $g(u,v)$ must be a parabola on the line segment corresponding to $v = u+1$. That is,

$$\begin{aligned} & u^2 \, \frac{\beta_2 - 2\beta_1}{2} + u(2u+1) \, \frac{\gamma_2 - 2\gamma_1 + \gamma_0}{2} + (2u+1)^2 \, \frac{\delta_2 - 2\delta_1 + \delta_0}{2} \\ & + \frac{u^2}{2u+1} \, \frac{4\beta_1 - \beta_2}{2} + u \, \frac{4\gamma_1 - \gamma_2 - 3\gamma_0}{2} + (2u+1) \, \frac{4\delta_1 - \delta_2 - 3\delta_0}{2} + \delta_0 \end{aligned}$$

must be a parabola in u. This is possible only if $4\beta_1 - \beta_2 = 0$. We have therefore reached the conclusion that (5.4) can be written as in the statement of the lemma. □

Let $K = \{1, 2, \ldots, k\}$ and $K^* = \{1, 2, \ldots, k-1\}$.

Lemma 4. Let S be a rational convex subset of $\mathbb{Q}^k$ such that S is full-dimensional. Let $g : S \to \mathbb{R}_+$ be a mapping such that the graph of g is a parabola on any line segment $r \subset S$. Suppose the restriction of g to the hyperplane defined by $\sum_{i \in K} u_i = t$ (for all $t \in \mathbb{R}$ such that the hyperplane intersects S) has the form $g(u_1, \ldots, u_{k-1}, t - \sum_{i \in K^*} u_i) = \sum_{i \in K^*} \sigma_{ii} u_i^2 + \sum_{i,j \in K^*, i<j} \sigma_{ij} u_i u_j + \sum_{i \in K^*} \sigma_i u_i + \sigma_0$ for some real $\sigma_{ii}, \sigma_{ij}, \sigma_i, \sigma_0$.

Then $g(u_1, \ldots, u_k) = \sum_{i \in K} \rho_{ii} u_i^2 + \sum_{i,j \in K, i<j} \rho_{ij} u_i u_j + \sum_{i \in K} \rho_i u_i + \rho_0$ for some real $\rho_{ii}, \rho_{ij}, \rho_i, \rho_0$.

Proof. Since S is full-dimensional, there is $u \in \operatorname{int} S$ and we can suppose without loss of generality that $u = (0, \ldots, 0)$. Let us consider the line defined by $(\alpha_1 t, \alpha_2 t, \ldots, \alpha_{k-1} t, (1 - \sum_{i \in K^*} \alpha_i) t)$ for some $\alpha_1, \ldots, \alpha_{k-1} \in \mathbb{R}$ and all $t \in \mathbb{R}$. The intersection of this line with S defines a line segment r_α passing by the origin. The graph of g on r_α is a parabola. We can express this by means of the following polynomial of degree 2 in t:

$$g(\alpha_1 t, \alpha_2 t, \ldots, \alpha_{k-1} t, (1 - \sum_{i \in K^*} \alpha_i) t) = k_\alpha t^2 + l_\alpha t + m_\alpha, \tag{5.5}$$

where k_α, l_α and m_α are real numbers.

Let us now consider the hyperplane defined by $(\alpha_1 t, \alpha_2 t, \ldots, \alpha_{k-1} t, (1 - \sum_{i \in K^*} \alpha_i) t)$ for some $t \in \mathbb{R}$ and $\alpha_i \in \mathbb{R}, \forall i \in K^*$. We assumed in the statement of the lemma,

$$g(\alpha_1 t, \alpha_2 t, \ldots, \alpha_{k-1} t, (1 - \sum_{i \in K^*} \alpha_i) t) = \sum_{i \in K^*} \sigma_{ii}^t \alpha_i^2 + \sum_{i,j \in K^*, i<j} \sigma_{ij}^t \alpha_i \alpha_j + \sum_{i \in K^*} \sigma_i^t \alpha_i + \sigma_0^t. \tag{5.6}$$

Setting $t = 0$ in (5.6) yields, $g(0, \ldots, 0) = \sum_{i \in K^*} \sigma_{ii}^0 \alpha_i^2 + \sum_{i,j \in K^*, i<j} \sigma_{ij}^0 \alpha_i \alpha_j + \sum_{i \in K^*} \sigma_i^0 \alpha_i + \sigma_0^0$. Since this must be true for all $\alpha_i \in \mathbb{R}, i \in K^*$, we must have $\sigma_{ii}^0 = \sigma_{ij}^0 = \sigma_i^0 = 0$, for all $i, j \in K^*$.

Equating (5.5) and (5.6) yields

$$k_\alpha t^2 + l_\alpha t + m_\alpha = \sum_{i \in K^*} \sigma_{ii}^t \alpha_i^2 + \sum_{i,j \in K^*, i<j} \sigma_{ij}^t \alpha_i \alpha_j + \sum_{i \in K^*} \sigma_i^t \alpha_i + \sigma_0^t. \tag{5.7}$$

Setting $t = 0$, $t = 1$ and $t = 2$ in (5.7) yields

$$\begin{aligned}
m_\alpha &= \sigma_0 \\
k_\alpha + l_\alpha + m_\alpha &= \sum_{i \in K^*} \sigma_{ii}^1 \alpha_i^2 + \sum_{i,j \in K^*, i<j} \sigma_{ij}^1 \alpha_i \alpha_j + \sum_{i \in K^*} \sigma_i^1 \alpha_i + \sigma_0^1 \\
4k_\alpha + 2l_\alpha + m_\alpha &= \sum_{i \in K^*} \sigma_{ii}^2 \alpha_i^2 + \sum_{i,j \in K^*, i<j} \sigma_{ij}^2 \alpha_i \alpha_j + \sum_{i \in K^*} \sigma_i^2 \alpha_i + \sigma_0^2.
\end{aligned}$$

The solution of this system is

$$\begin{aligned}
m_\alpha &= \sigma_0 \\
k_\alpha &= \sum_{i \in K^*} \alpha_i^2 \, \frac{\sigma_{ii}^2 - 2\sigma_{ii}^1}{2} + \sum_{i,j \in K^*, i<j} \alpha_i \alpha_j \, \frac{\sigma_{ij}^2 - 2\sigma_{ij}^1}{2} \\
&\quad + \sum_{i \in K^*} \alpha_i \, \frac{\sigma_i^2 - 2\sigma_i^1}{2} + \frac{\sigma_0^2 - 2\sigma_0^1 + \sigma_0^0}{2} \\
l_\alpha &= \sum_{i \in K^*} \alpha_i^2 \, \frac{4\sigma_{ii}^1 - \sigma_{ii}^2}{2} + \sum_{i,j \in K^*, i<j} \alpha_i \alpha_j \, \frac{4\sigma_{ij}^1 - \sigma_{ij}^2}{2} \\
&\quad + \sum_{i \in K^*} \alpha_i \, \frac{4\sigma_i^1 - \sigma_i^2}{2} + \frac{4\sigma_0^1 - \sigma_0^2 - 3\sigma_0^0}{2}.
\end{aligned}$$

Let us rewrite (5.5) : $g(\alpha_1 t, \alpha_2 t, \ldots, \alpha_{k-1} t, (1 - \sum_{i\in K^*} \alpha_i)t) =$

$$\left(\sum_{i\in K^*} \alpha_i^2 \,\frac{\sigma_{ii}^2 - 2\sigma_{ii}^1}{2} + \sum_{i,j\in K^*, i<j} \alpha_i\alpha_j \,\frac{\sigma_{ij}^2 - 2\sigma_{ij}^1}{2} + \sum_{i\in K^*} \alpha_i \,\frac{\sigma_i^2 - 2\sigma_i^1}{2} + \frac{\sigma_0^2 - 2\sigma_0^1 + \sigma_0^0}{2}\right) t^2$$
$$+ \left(\sum_{i\in K^*} \alpha_i^2 \,\frac{4\sigma_{ii}^1 - \sigma_{ii}^2}{2} + \sum_{i,j\in K^*, i<j} \alpha_i\alpha_j \,\frac{4\sigma_{ij}^1 - \sigma_{ij}^2}{2} + \sum_{i\in K^*} \alpha_i \,\frac{4\sigma_i^1 - \sigma_i^2}{2} + \frac{4\sigma_0^1 - \sigma_0^2 - 3\sigma_0^0}{2}\right) t + \sigma_0.$$

Letting $\alpha_i t = u_i, \forall i \in K^*, (1 - \sum_{i\in K^*} \alpha_i)t = u_k$, we have $\alpha_i = u_i / \sum_{i\in K} u_i$ and $t = \sum_{i\in K} u_i$, and the previous equation becomes, $g(u_1, \ldots, u_k) =$

$$\sum_{i\in K^*} u_i^2 \,\frac{\sigma_{ii}^2 - 2\sigma_{ii}^1}{2} + \sum_{i,j\in K^*, i<j} u_i u_j \,\frac{\sigma_{ij}^2 - 2\sigma_{ij}^1}{2} + \sum_{i\in K^*} u_i \sum_{j\in K} u_j \,\frac{\sigma_i^2 - 2\sigma_i^1}{2}$$
$$+ \left(\sum_{i\in K} u_i\right)^2 \,\frac{\sigma_0^2 - 2\sigma_0^1 + \sigma_0^0}{2} + \sum_{i\in K^*} \frac{u_i^2}{\sum_{j\in K} u_j} \,\frac{4\sigma_{ii}^1 - \sigma_{ii}^2}{2}$$
$$+ \sum_{i,j\in K^*, i<j} \frac{u_i u_j}{\sum_{j\in K} u_j} \,\frac{4\sigma_{ij}^1 - \sigma_{ij}^2}{2} + \sum_{i\in K^*} u_i \,\frac{4\sigma_i^1 - \sigma_i^2}{2}$$
$$+ \sum_{i\in K} u_i \,\frac{4\sigma_0^1 - \sigma_0^2 - 3\sigma_0^0}{2} + \sigma_0.$$

For any $j \in K^*$, the graph of $g(u_1, \ldots, u_k)$ must be a parabola on the line segment corresponding to $u_i = 0, \forall i \in K^* \setminus \{j\}, u_k = u_j + 1$. That is, $g(0, \ldots, 0, u_j, 0, \ldots, 0, u_j + 1) =$

$$u_j^2 \,\frac{\sigma_{jj}^2 - 2\sigma_{jj}^1}{2} + u_j(2u_j + 1) \,\frac{\sigma_j^2 - 2\sigma_j^1}{2} + (2u_j + 1)^2 \,\frac{\sigma_0^2 - 2\sigma_0^1 + \sigma_0^0}{2}$$
$$+ \frac{u_j^2}{2u_j + 1} \,\frac{4\sigma_{jj}^1 - \sigma_{jj}^2}{2} + u_j \,\frac{4\sigma_j^1 - \sigma_j^2}{2} + (2u_j + 1) \,\frac{4\sigma_0^1 - \sigma_0^2 - 3\sigma_0^0}{2} + \sigma_0$$

must be a parabola in u_j. This is possible only if $4\sigma_{jj}^1 - \sigma_{jj}^2 = 0$ for all $j \in K^*$.

Similarly, for any $i, j \in K^*$ with $i < j$, the graph of $g(u_1, \ldots, u_k)$ must be a parabola on the line segment corresponding to $u_i = u_j$, $u_l = 0, \forall l \in$

$K^* \setminus \{i,j\}, u_k = u_i + 1$. That is, $g(0,\dots,0,u_i,0,\dots,0,u_i,0,\dots,0,u_i+1) =$

$$u_i^2 \frac{\sigma_{ii}^2 - 2\sigma_{ii}^1}{2} + u_i^2 \frac{\sigma_{jj}^2 - 2\sigma_{jj}^1}{2} + u_i^2 \frac{\sigma_{ij}^2 - 2\sigma_{ij}^1}{2} + u_i(3u_i+1) \frac{\sigma_i^2 - 2\sigma_i^1}{2}$$
$$+ u_i(3u_i+1) \frac{\sigma_j^2 - 2\sigma_j^1}{2} + (3u_i+1)^2 \frac{\sigma_0^2 - 2\sigma_0^1 + \sigma_0^0}{2} + \frac{u_i^2}{3u_i+1} \frac{4\sigma_{ij}^1 - \sigma_{ij}^2}{2}$$
$$+ u_i \frac{4\sigma_i^1 - \sigma_i^2}{2} + u_i \frac{4\sigma_j^1 - \sigma_j^2}{2} + (3u_i+1) \frac{4\sigma_0^1 - \sigma_0^2 - 3\sigma_0^0}{2} + \sigma_0$$

must be a parabola in u_i. This is possible only if $4\sigma_{ij}^1 - \sigma_{ij}^2 = 0$ for all $i,j \in K^*$ with $i \neq j$. We have therefore reached the conclusion that g has the desired form. □

Lemma 5. Let S be a rational convex subset of $\mathbb{Q}^k$ such that $(0,0,\dots,0) \in \operatorname{int} S$. Let $g : S \to \mathbb{R}_+$ be a mapping such that $g(u_1,\dots,u_k) = 0$ if and only $(u_1,\dots,u_k) = (0,\dots,0)$ and the graph of g is a parabola on any line segment $r \subset S$. Then $g(u_1,\dots,u_k) = \sum_{i\in K} \rho_{ii} u_i^2 + \sum_{i,j\in K, i<j} \rho_{ij} u_i u_j$ for some real ρ_{ii}, ρ_{ij}.

Proof. By induction and Lemmas 4 and 3 , $g(u_1,\dots,u_k) = \sum_{i\in K} \rho_{ii} u_i^2 + \sum_{i,j\in K,i<j} \rho_{ij} u_i u_j + \sum_{i\in K} \rho_i u_i + \rho_0$ for some real $\rho_{ii}, \rho_{ij}, \rho_i, \rho_0$. On the line $u_2 = u_3 = \ldots = u_k = 0$, $g(u_1, 0,\dots,0) = \rho_{11} u^2 + \rho_1 u + \rho_0$. The graph of this function of u must be a parabola with vertex in 0. Hence $\rho_1 = \rho_0 = 0$. Similarly, for any $i \in K$, considering the line defined by $u_j = 0, \forall j \neq i$ entails $\rho_i = 0$. In conclusion, $g(u_1,\dots,u_k) = \sum_{i\in K} \rho_{ii} u_i^2 + \sum_{i,j\in K,i<j} \rho_{ij} u_i u_j$. □

Lemma 6. Suppose π is fixed and the dissimilarity measure f satisfies Homogeneity of degree 0 and Deviations Balancedness. Then, for all $p,q \in \Pi$ such that $p+q-\pi \in \Pi$ and $p-q+\pi \in \Pi$, we have

$$F(p+q-\pi) + F(p-q+\pi) = 2F(p) + 2F(q). \tag{5.8}$$

Proof. Let p,q be as in the statement of the lemma. There are two distributions $x,y \in X$ such that $x/s(x) = p$ and $y/s(y) = q$. Hence $F(p) = f(x,\pi)$ and $F(q) = f(y,\pi)$. By Homogeneity of degree 0, $F(p) = f(x,\pi) = f(s(y)x,\pi)$ and $F(q) = f(y,\pi) = f(s(x)y,\pi)$. The two distributions $s(y)x$ and $s(x)y$ have the same size, i.e., $s(x)s(y)$. Hence we can apply Deviations Balancedness and we find

$$\begin{aligned} 2F(p) + 2F(q) &= 2f(s(y)x,\pi) + 2f(s(x)y,\pi) \\ &= f(s(y)x + s(x)y - s(x)s(y)\pi, \pi) \\ &\quad + f(s(y)x - s(x)y + s(x)s(y)\pi, \pi). \end{aligned}$$

By the definition of F,

$$f(s(y)x + s(x)y - s(x)s(y)\pi, \pi) = F(p+q-\pi)$$

and

$$f(s(y)x - s(x)y + s(x)s(y)\pi, \pi) = F(p-q+\pi).$$

In conclusion,

$$F(p+q-\pi)+F(p-q+\pi)=2F(p)+2F(q). \qquad \square$$

For every $l \in N$, define $N_l = N \setminus \{l\}$ and $N_{lm} = N \setminus \{l,m\}$.

Lemma 7. Suppose π is fixed and the dissimilarity measure f satisfies Homogeneity of degree 0 and Deviations Balancedness. Then, for every $l \in N$ and $p \in \Pi$,

$$F(p)=\sum_{i\in N_l}\rho^l_{ii}(p_i-\pi_i)^2+\sum_{i,j\in N_l:i<j}\rho^l_{ij}(p_i-\pi_i)(p_j-\pi_j)$$

for some real ρ^l_{ii}, ρ^l_{ij}.

Proof. Let r be a line segment with extremities $s,t \in \Pi$, with $s \neq t$. Every point of $r \cap \Pi$ can be written as $\alpha t+(1-\alpha)s$ with $\alpha \in [0,1]$ and α rational. Consider any two points $p,q \in r\cap\Pi$, the position of which on r is characterized by α and β respectively. Then $p+q-\pi$ lies on the line segment between $2t-\pi$ and $2s-\pi$ and it can be written as

$$\left(\frac{\alpha+\beta}{2}\right)(2t-\pi)+\left(1-\frac{\alpha+\beta}{2}\right)(2s-\pi).$$

Similarly, $p-q+\pi$ lies on the line segment between $t-s+\pi$ and $s-t+\pi$ and it can be written as

$$\left(\frac{\alpha-\beta+1}{2}\right)(t-s+\pi)+\left(1-\frac{\alpha-\beta+1}{2}\right)(s-t+\pi).$$

Notice that $(\alpha+\beta)/2 \in [0,1]$ and $(\alpha-\beta+1)/2 \in [0,1]$ for any $\alpha,\beta \in [0,1]$. Define three mappings as follows:

- $L : [0,1]\cap\mathbb{Q} \to \mathbb{R}_+$ by $L(\alpha)=F(p)$ if $p=\alpha t+(1-\alpha)s$;
- $G : [0,1]\cap\mathbb{Q} \to \mathbb{R}_+$ by $G(\alpha)=F(p)$ if $p=\alpha(2t-\pi)+(1-\alpha)(2s-\pi)$;
- $H : [0,1]\cap\mathbb{Q} \to \mathbb{R}_+$ by $H(\alpha)=F(p)$ if $p=\alpha(t-s+\pi)+(1-\alpha)(s-t+\pi)$.

Then (5.8) can be rewritten as

$$G\left(\frac{\alpha+\beta}{2}\right)+H\left(\frac{\alpha-\beta+1}{2}\right)=2L(\alpha)+2L(\beta) \qquad (5.9)$$

and it holds for all rational $\alpha,\beta \in [0,1]$. This functional equation is a generalization of Equation (18) discussed in [1, p.82].

If $\alpha=\beta$, then $G(\alpha)+H(1/2)=4L(\alpha)$. In other words, $G(\alpha)=4L(\alpha)+\delta'$ for some real number δ'. If $\alpha=1-\beta$, then $G(1/2)+H(1-\beta)=2L(1-\beta)+2L(\beta)$. So, $H(1-\beta)=2L(\beta)+2L(1-\beta)+\delta''$ for some real number δ''. We can now rewrite (5.9) as

$$4L\left(\frac{\alpha+\beta}{2}\right)+2L\left(\frac{\alpha-\beta+1}{2}\right)+2L\left(\frac{\beta-\alpha+1}{2}\right)=2L(\alpha)+2L(\beta)+\delta'''$$

with $\delta'''=-\delta'-\delta''$. If we now let $\alpha=(m-2)c$ and $\beta=mc$ with m a positive integer ($m \geq 2$) and c a positive rational number such that $mc \in [0,1]$, then

$$4L\left((m-1)c\right)+2L\left((1/2)-c\right)+2L\left((1/2)+c\right)=2L((m-2)c)+2L(mc)+\delta'''.$$

If we divide this equation by 2 and reorder the terms (with $\delta = \delta'''/2$), we obtain that $L(mc) - L((m-1)c)$ is equal to

$$\begin{aligned} & L((m-1)c) - L((m-2)c) + L((1/2)-c) + L((1/2)+c) + \delta \\ = \;& L((m-2)c) - L((m-3)c) + 2\big(L((1/2)-c) + L((1/2)+c) + \delta\big) \\ = \;& L((m-3)c) - L((m-4)c) + 3\big(L((1/2)-c) + L((1/2)+c) + \delta\big) \\ = \;& \dots \\ = \;& L(c) - L(0) + (m-1)\big(L((1/2)-c) + L((1/2)+c) + \delta\big). \end{aligned}$$

Notice that, for all $m \in \mathbb{N} (m \geq 2)$ and $c \in \mathbb{Q}_{++}$ such that $mc \in [0,1]$,

$$\begin{aligned} L(mc) &= \sum_{i=1}^{m} \big(L(ic) - L((i-1)c)\big) + L(0) \\ &= \sum_{i=1}^{m} \Big(L(c) - L(0) + (i-1)\big(L((1/2)-c) + L((1/2)+c) + \delta\big)\Big) \\ &\qquad + L(0) \\ &= mL(c) + (1-m)L(0) \\ &\qquad + \frac{m(m-1)}{2}\,\big(L((1/2)-c) + L\big((1/2)+c\big) + \delta\big). \end{aligned} \tag{5.10}$$

If $m = 0$, it is easy to verify that (5.10) still holds. Indeed,

$$L(0c) = 0L(c) + (1-0)L(0) + \frac{0(0-1)}{2}\,\big(L((1/2)-c) + L\big((1/2)+c\big) + \delta\big).$$

A similar reasoning holds when $m = 1$. Hence, for any fixed value of c (a positive rational number), $L(\beta)$ is a polynomial of degree 2 in β for every $\beta \in [0,1]$ that can be written as mc for some integer m, hence we can also write (5.10) as

$$L(\beta) = a\beta^2 + b\beta + d,$$

for some real numbers a, b, d. Let $c = 1/2$; β can take the values $0, 1/2, 1$. We have

$$\begin{cases} L(0) &= d \\ L(1/2) &= 1/4\, a + 1/2\, b + d \\ L(1) &= a + b + d \end{cases}$$

This is a non-singular (determinant $= -1/4$) system of linear equations which determines unique values for a, b, d. These are a linear combination of the values of $L(\beta)$ for $\beta = 0, 1/2, 1$. There is a single parabola that passes through the three points $(\beta, L(\beta))$ for $\beta = 0, 1/2, 1$.

Consider now any rational number $\beta = \frac{w}{w'}$. If w' is odd, we also have that $\beta = \frac{2w}{2w'}$ so that we can assume that w' is even. Using (5.10), we have that $L(\beta) = a'\beta^2 + b'\beta + d'$ for some a', b', d' and for any integer w in the interval $[0, w']$. We have to show that these constants are a, b, d. Indeed, for $w = 0, w'/2, w'$, we have that $\beta = 0, 1/2, 1$ respectively. Hence, the points $(\beta, L(\beta))$ for $\beta = 0, 1/2, 1$ are also on the parabola with coefficients a', b', c' and these points are the same as in the case $w' = 2$ since L is a single function.

Since there is only one parabola through these points and the parabola with coefficients a, b, c passes through these points, we conclude that $a = a'$, $b = b'$ and $d = d'$.

This being true for any (even) value of w', we have that $L(\beta)$ is a quadratic function of β independently of the denominator of the rational number β.

Of course, the coefficients of the polynomial a, b and d depend on the line segment r joining t and s. So, we had better write $L_r(\beta) = a_r\beta^2 + b_r\beta + d_r$. We now go back to F. Since L_r is a polynomial of degree 2 for any r, we find that the graph of F is a parabola on any line segment r. Define $G : \Pi \to \mathbb{R}_+$ by $G(p - \pi) = F(p)$. Then G is like g in the statement of Lemma 5, with $k = n - 1$ (because Π has only $n - 1$ dimensions). Then, $G(p - \pi) = F(p) = \sum_{i\in N_l} \rho^l_{ii}(p_i - \pi_i)^2 + \sum_{i,j\in N_l,i<j} \rho^l_{ij}(p_i - \pi_i)(p_j - \pi_j)$ for some real ρ^l_{ii}, ρ^l_{ij}. Notice that, for each $l \in N$, we have such an expression. □

Proof of Theorem 2.1.

Case $n = 2$.

f is homogeneous of degree 0. Since $n = 2$, we have $(p_1 - \pi_1)^2 = (p_2 - \pi_2)^2$. Then Lemma 7 yields $F(p) = \rho^2_{11}(p_1 - \pi_1)^2 = \frac{\rho^2_{11}}{2}\left((p_1 - \pi_1)^2 + (p_2 - \pi_2)^2\right)$ and

$$f(x, \pi) = \frac{\rho^2_{11}}{2}\left(\left(\frac{x_1}{s(x)} - \pi_1\right)^2 + \left(\frac{x_2}{s(x)} - \pi_2\right)^2\right). \tag{5.11}$$

f is homogeneous of degree 1. By Lemma 1, $f/s(x)$ satisfies Homogeneity of degree 0 and Deviations Balancedness. By (5.11), $f/s(x)$ is proportional to χ^2_0 and f to χ^2_1.

Case $n \geq 3$.

f is homogeneous of degree 0. We know from the previous lemma that F can be expressed as

$$F(p) = \sum_{i\in N_n} \rho^n_{ii}(p_i - \pi_i)^2 + \sum_{i,j\in N_n:i<j} \rho^n_{ij}(p_i - \pi_i)(p_j - \pi_j) \tag{5.12}$$

for some real ρ^n_{ii}, ρ^n_{ij}. Recall that $N_n = N \setminus \{n\}$. So, for all $j, l \in N_n$, with $j \neq l$,

$$f(k\pi + \mathbf{1}^j - \mathbf{1}^l, \pi) = \frac{\rho^n_{jj} + \rho^n_{ll} - \rho^n_{jl}}{k^2},$$

and

$$f(k\pi + \mathbf{1}^n - \mathbf{1}^j, \pi) = \frac{\rho^n_{jj}}{k^2}.$$

For the sake of simplicity, for all $j, l \in N$, let A_{jl} denote $\frac{\pi_j+\pi_l}{\pi_j\pi_l}$. Then, thanks to Restricted Inverse Effects, we can write, for all $j, l \in N_n$ with $j \neq l$.

$$\frac{f(k\pi + \mathbf{1}^j - \mathbf{1}^l, \pi)}{f(k\pi + \mathbf{1}^1 - \mathbf{1}^2, \pi)} = \frac{\rho^n_{jj} + \rho^n_{ll} - \rho^n_{jl}}{\rho^n_{11} + \rho^n_{22} - \rho^n_{12}} = \frac{A_{jl}}{A_{12}},$$

$$\frac{f(k\pi + \mathbf{1}^j - \mathbf{1}^n, \pi)}{f(k\pi + \mathbf{1}^1 - \mathbf{1}^n, \pi)} = \frac{\rho^n_{jj}}{\rho^n_{11}} = \frac{A_{jn}}{A_{1n}}$$

and

$$\frac{f(k\pi + \mathbf{1}^1 - \mathbf{1}^n, \pi)}{f(k\pi + \mathbf{1}^1 - \mathbf{1}^2, \pi)} = \frac{\rho^n_{11}}{\rho^n_{11} + \rho^n_{22} - \rho^n_{12}} = \frac{A_{1n}}{A_{12}}.$$

Using each of these three equations separately, we find

$$\rho^n_{jl} = \rho^n_{jj} + \rho^n_{ll} + \frac{A_{jl}}{A_{12}}(\rho^n_{12} - \rho^n_{11} - \rho^n_{22}),\ \forall j, l \in N_n : j \neq l, \tag{5.13}$$

$$\rho^n_{jj} = \rho^n_{11}\frac{A_{jn}}{A_{1n}},\ \forall j \in N_n \tag{5.14}$$

and

$$\frac{\rho^n_{11} + \rho^n_{22} - \rho^n_{12}}{A_{12}} = \frac{\rho^n_{11}}{A_{1n}}. \tag{5.15}$$

If we substitute (5.14) and (5.15) in (5.13), we obtain

$$\begin{aligned} \rho^n_{jl} &= \rho^n_{11}\left(\frac{A_{jn}}{A_{1n}} + \frac{A_{ln}}{A_{1n}} - \frac{A_{jl}}{A_{1n}}\right),\ \forall j, l \in N_n : j \neq l, \\ &= 2\rho^n_{11}\ \frac{\pi_1}{\pi_1 + \pi_n}. \end{aligned}$$

From (5.14), we find

$$\rho^n_{jj} = \rho^n_{11}\frac{\pi_j + \pi_n}{\pi_j}\ \frac{\pi_1}{\pi_1 + \pi_n},\ \forall j \in N_n.$$

Let us define

$$\gamma = \rho^n_{11}\ \frac{\pi_1\pi_n}{\pi_1 + \pi_n}.$$

Then,

$$\rho^n_{jl} = 2\gamma\ \frac{\pi_1 + \pi_n}{\pi_1\pi_n}\ \frac{\pi_1}{\pi_1 + \pi_n} = 2\ \frac{\gamma}{\pi_n},\ \forall j, l \in N_n : j \neq l,$$

and

$$\rho^n_{jj} = \gamma\ \frac{\pi_1 + \pi_n}{\pi_1\pi_n}\ \frac{\pi_j + \pi_n}{\pi_j}\ \frac{\pi_1}{\pi_1 + \pi_n} = \gamma\ \frac{\pi_j + \pi_n}{\pi_j\pi_n},\ \forall j \in N_n.$$

If we now substitute the expressions of ρ^n_{jl} and ρ^n_{jj} into (5.12), we obtain $f(x,\pi)$

$$\begin{aligned}
&= \sum_{i\in N_n}\left(\frac{\gamma}{\pi_i}+\frac{\gamma}{\pi_n}\right)\left(\pi_i-\frac{x_i}{s(x)}\right)^2 \\
&\quad + \sum_{i,j\in N_n: i<j}\frac{2\gamma}{\pi_n}\left(\pi_i-\frac{x_i}{s(x)}\right)\left(\pi_j-\frac{x_j}{s(x)}\right) \\
&= \sum_{i\in N_n}\frac{\gamma}{\pi_i}\left(\pi_i-\frac{x_i}{s(x)}\right)^2+\sum_{i\in N_n}\frac{\gamma}{\pi_n}\left(\pi_i-\frac{x_i}{s(x)}\right)^2 \\
&\quad + \sum_{i,j\in N_n: i<j}\frac{2\gamma}{\pi_n}\left(\pi_i-\frac{x_i}{s(x)}\right)\left(\pi_j-\frac{x_j}{s(x)}\right) \\
&= \sum_{i\in N_n}\frac{\gamma}{\pi_i}\left(\pi_i-\frac{x_i}{s(x)}\right)^2+\frac{\gamma}{\pi_n}\left(\sum_{i\in N_n}\left(\pi_i-\frac{x_i}{s(x)}\right)\right)^2 \\
&= \sum_{i\in N_n}\frac{\gamma}{\pi_i}\left(\pi_i-\frac{x_i}{s(x)}\right)^2+\frac{\gamma}{\pi_n}\left(1-\sum_{i\in N_n}\pi_i-1+\sum_{i\in N_n}\frac{x_i}{s(x)}\right)^2 \\
&= \sum_{i\in N_n}\frac{\gamma}{\pi_i}\left(\pi_i-\frac{x_i}{s(x)}\right)^2+\frac{\gamma}{\pi_n}\left(\pi_n-\frac{x_n}{s(x)}\right)^2 \\
&= \sum_{i\in N}\frac{\gamma}{\pi_i}\left(\pi_i-\frac{x_i}{s(x)}\right)^2=\gamma\chi^2_0(x,\pi). \qquad (5.16)
\end{aligned}$$

f is homogeneous of degree 1. By Lemma 1, $f/s(x)$ satisfies Homogeneity of degree 0, Deviations Balancedness and Inverse Effects. By (5.16), $f/s(x)$ is proportional to χ^2_0 and f to χ^2_1. □

Proof of Theorem 2.2. We prove the result only for χ^2_1 (the case of χ^2_0 being similar). We want to prove that there exists $\gamma>0$ such that $f(x,\pi)=\gamma\chi^2_1(x,\pi)$ for all $x\in X$ and all $\pi\in\Pi$. From Theorem 2.1, we have $f(x,\pi)=\gamma_\pi\chi^2_1(x,\pi)$, where we now use the notation γ_π to make clear that γ_π can depend on π. Thanks to Inverse Effects, for all $\pi,\pi'\in\Pi$, we have

$$\frac{f(k\pi+\mathbf{1}^j-\mathbf{1}^l,\pi)}{f(k\pi'+\mathbf{1}^j-\mathbf{1}^l,\pi')}=\frac{\gamma_\pi\left(\frac{1}{\pi_j}+\frac{1}{\pi_l}\right)}{\gamma_{\pi'}\left(\frac{1}{\pi'_j}+\frac{1}{\pi'_l}\right)}=\frac{\frac{1}{\pi_j}+\frac{1}{\pi_l}}{\frac{1}{\pi'_j}+\frac{1}{\pi'_l}}.$$

This is possible only if $\gamma_\pi=\gamma_{\pi'}$ for all $\pi,\pi'\in\Pi$. □

References

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

[2] C. Bertoluzza, M. Di Bacco, and V. Doldi. An axiomatic characterization of the measures of similarity. *Sankhya: The Indian Journal of Statistics*, 66(3): 474–486, 2004.

[3] D. Bouyssou, T. Marchant, and M. Pirlot. An axiomatic characterization of two disproportionality and malapportionment indices: the Duncan and Duncan index and the Lijphart index. *Annals of Operations Research*, 284(1):147–163, 2020.

[4] F. A. Cowell, R. Davidson, and E. Flachaire. Goodness of fit: An axiomatic approach. *Journal of Business and Economic Statistics*, 33(1):54–67, 2015.

[5] M. D'Agostino and V. Dardanoni. What's so special about Euclidean distance? *Social Choice and Welfare*, 33:211–233, 2009.

[6] M. Greenacre. Ordination with any dissimilarity measure: a weighted Euclidean solution. *Ecology*, 98(9):2293–2300, 2017.

[7] H. Kaufman, A.M. Mathai, and P.N. Rathie. A mathematical foundation for Pearson's chisquare goodness-of-fit statistic. *Sankhya Series A*, 34:441–442, 1972.

[8] S. F. Reardon and G. Firebaugh. Measures of multigroup segregation. *Sociological Methodology*, 32(1):33–67, 2002.

[9] G. Yalonetzky. A dissimilarity index of multidimensional inequality of opportunity. *Journal of Economic Inequality*, 10(3):343–373, 2012.

Denis Bouyssou
CNRS and Université Paris Dauphine
e-mail: `bouyssou@lamsade.dauphine.fr`

Thierry Marchant
Ghent University
e-mail: `thierry.marchant@ugent.be`

Marc Pirlot
UMons
e-mail: `marc.pirlot@umons.ac.be`