Austrian Journal of Statistics
2024, Volume 53, 85–98.
http://www.ajs.or.at/
doi:10.17713/ajs.v53i4.1845

Symmetry of Square Contingency Tables Using Simplicial Geometry

Keita Nakamura
Tokyo University of Science
Japan

Tomoyuki Nakagawa
Meisei University
Japan

Kouji Tahata
Tokyo University of Science
Japan

Abstract

Two-way contingency tables illustrate the relationship between two discrete variables. Their corresponding probability tables can be regarded as an element in a simplex. Herein we discuss the symmetry of a square contingency table with the same row and column classifications. Specifically, we identify symmetric probability tables as a linear subspace using the Aitchison geometry of the simplex. Then given a probability table, an orthogonal projection onto the symmetric subspace yields the nearest symmetric table. The (i, j) cell of the nearest symmetric table is characterized as the geometric mean of symmetric cells. This characterization does not agree with the standard maximum likelihood estimators, except in the symmetric case. The original probability table is subsequently decomposed into symmetric and skew-symmetric tables, which are orthogonal to each other. Finally, we develop a method to test the symmetry of a contingency table based on a parametric bootstrap and provide an example.

Keywords: contingency table, Aitchison geometry, Bowker test, orthogonal decomposition.

1. Introduction

The variables in a square two-way contingency table often depend on each other. Consider an $I \times I$ square contingency table with the same row and column classifications, where n individuals are arranged in I^2 cells with n_{ij} individuals assigned to the (i, j) cell. Assuming the data is obtained from a multinomial sampling, individuals fall into each of the (i, j) cells with a certain probability, $p_{ij} > 0$, $\sum_{i=1}^{I}\sum_{j=1}^{I} p_{ij} = 1$. We will refer to a table with p_{ij} in the ith row and jth column as a probability table (PT). Our research is interested in the symmetric structure of p_{ij}, which means the probability that an individual falls into the (i, j) cell is equal to the probability that the individual falls into the (j, i) cell. Bowker (1948) proposed a method to test the hypothesis, $H_0 : p_{ij} = p_{ji}$, for $i, j = 1, 2, \ldots, I$, using χ^2 type statistic,

$$\chi^2 = n \sum_{i=1}^{I} \sum_{j=1}^{I} \frac{(\hat{p}_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}, \tag{1}$$

where $\hat{E}_{ij} = (\hat{p}_{ij} + \hat{p}_{ji})/2$ and $\hat{p}_{ij} = n_{ij}/n$. It has an approximately χ^2 distribution with

Table 1: Data of unaided distance vision of 7477 women aged 30 − 39 (Stuart 1955)

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade	Total
Highest grade	1520	266	124	66	1976
Second grade	234	1512	432	78	2256
Third grade	117	362	1772	205	2456
Lowest grade	36	82	179	492	7890
Total	1907	2222	2507	841	7477

Table 2: Estimated PT(%) corresponding to Table 1

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade	Total
Highest grade	20.33	3.56	1.66	0.88	26.43
Second grade	3.13	20.22	5.78	1.04	30.17
Third grade	1.56	4.84	23.70	2.74	32.85
Lowest grade	0.48	1.10	2.39	6.58	10.55
Total	25.50	29.72	33.53	11.25	100.00

$I(I-1)/2$ degrees of freedom as $n \to +\infty$. We can also use likelihood ratio statistics,

$$G^2 = 2n \sum_{i=1}^{I} \sum_{j=1}^{I} \hat{p}_{ij} \log \left(\frac{\hat{p}_{ij}}{\hat{E}_{ij}} \right), \tag{2}$$

which has the same limiting distribution as (1). These statistics are interpreted as divergences from the maximum likelihood estimator, $\hat{p}_{ij}$, to the one under H_0, $\hat{E}_{ij}$.

Throughout this paper, we use the contingency table given by Stuart (1955) as an example. This table is a 4×4 contingency table, which classifies the eyesight of 7477 women (Highest, Second, Third, and Lowest grades). Table 1 displays the contingency table, while Table 2 shows the corresponding estimated PT by percentage. For this table, Bowker's test on the symmetry of the square contingency table using (1) produces $\chi^2 = 19.1$ with a p-value $\simeq 0.004$. Consequently, the null hypothesis is rejected at the 0.05 significance level.

When the symmetry model, H_0, does not hold for the given data, we are interested in applying the asymmetric or more relaxed models. For example, Kateri and Papaioannou (1997), Kateri and Agresti (2007), and Tahata (2020) proposed the asymmetry models based on the f-divergence for analyzing square contingency tables. Moreover, since we are interested in measuring the degree of departure from symmetry, for example, Tomizawa (1994), Tomizawa, Seo, and Yamamoto (1998), and Momozaki, Cho, Nakagawa, and Tomizawa (2023) proposed the measures that represent the degree of departure from symmetry using various approaches. Tahata (2022) gives a review that focuses on modeling based on f-divergence and the measure of asymmetry. However, there has yet to be a detailed investigation of symmetry analysis using *Aitchison geometry*. Therefore, this paper considers the method based on Aitchison geometry for square contingency table analysis.

We can view $I \times J$ PTs associated with contingency tables as an element of the standard $(IJ-1)$-dimensional simplex, where the subset of $\mathbb{R}^{IJ}$ is given by

$$\mathcal{S}^{IJ} = \left\{ (t_{11}, t_{12}, \ldots, t_{IJ}) \in \mathbb{R}^{IJ} \mid t_{ij} > 0, \sum_{i=1}^{I} \sum_{j=1}^{J} t_{ij} = 1 \right\}$$

In compositional data analysis, a simplex equipped with an algebraic-geometric structure, called the Aitchison geometry, has been used because it becomes a Euclidean space (Billheimer, Guttorp, and Fagan 2001; Pawlowsky-Glahn and Egozcue 2001). For two-way probability tables, Egozcue, Pawlowsky-Glahn, Templ, and Hron (2015) introduced the decomposition of a PT into two geometric marginal tables and an interaction table using Aitchison geometry and proposed an analysis method similar to the ordinary log-linear model. Herein we define the symmetry of PTs and propose a method to analyze the symmetry of square contingency tables using the Aitchison geometry.

In Section 2, we recall simplicial operations and the Aitchison geometry for compositional data: perturbation, powering, centered log-ratio (clr) transformation, and the Aitchison metrics. Section 3 introduces the transposition of a PT and the symmetric PTs, which create the subspace and the decomposition obtained by the orthogonal projection of a given PT. Moreover, we present the properties of a skew-symmetric PT obtained by subtracting the symmetric PT from the original one and proposes measures of departure from symmetry. Section 4 provides a method to test the symmetry based on the distance between the original PT and the nearest symmetric one. Section 5 illustrates the proposed decomposition of PTs and the testing of symmetry through an example 1. Finally, Section 6 concludes this paper. All proofs are deferred to the Appendix.

2. Simplicial operations and the Aitchison geometry

Consider an $I \times J$ contingency table obtained from multinomial sampling. Let $\mathbf{T} = (t_{ij})$ denote the vector of its associated PTs, whose probabilities are assumed to be strictly positive and add up to 1. Therefore, the table is regarded as an element in a simplex, $\mathcal{S}^{IJ}$.

Aitchison (1982) introduced an operation called *perturbation* between elements of a simplex. In addition, Aitchison (1986) defined an operation called *powering* as a repeated perturbation. For two PTs, $\mathbf{R}$ and $\mathbf{T}$, and a real number, α, the perturbation, $\oplus$, and powering, $\odot$, are respectively given as

$$\mathbf{R} \oplus \mathbf{T} = \left(\frac{r_{11}t_{11}}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij}t_{ij}}, \ldots, \frac{r_{IJ}t_{IJ}}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij}t_{ij}} \right),$$

$$\alpha \odot \mathbf{R} = \left(\frac{r_{11}^{\alpha}}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij}^{\alpha}}, \ldots, \frac{r_{IJ}^{\alpha}}{\sum_{i=1}^{I}\sum_{j=1}^{J} r_{ij}^{\alpha}} \right).$$

Simplex $\mathcal{S}^{IJ}$ equipped with $\oplus$ and $\odot$ is a $(IJ-1)$-dimensional vector space. (Aitchison, Barceló-Vidal, Martín-Fernández, and Pawlowsky-Glahn 2001; Pawlowsky-Glahn and Egozcue 2001)

Because analyzing a simplex with the constraint that its entries add to 1 is difficult, Aitchison (1982, 1986) achieved a flexible analysis by mapping the simplex onto a space of real numbers with two different isomorphisms: the additive log-ratio (alr) transformation and the clr transformation. The clr transformation is suited to analyze PTs since each part of a PT is represented by a contrast with the geometric mean of all the parts of the PT, which is easier to interpret.

Let $g(\mathbf{T})$ be the geometric mean of the entries of $\mathbf{T}$, that is,

$$g(\mathbf{T}) = \left(\prod_{i=1}^{I}\prod_{j=1}^{J} t_{ij} \right)^{1/(IJ)}.$$

The clr transformation of $\mathbf{T} \in \mathcal{S}^{IJ}$, $\mathrm{clr}(\mathbf{T})$, is an $I \times J$ table whose entries are expressed as

$$\mathrm{clr}_{ij}(\mathbf{T}) = \log \frac{t_{ij}}{g(\mathbf{T})}, \quad i = 1, \ldots, I, \; j = 1, \ldots, J,$$

where log refers to the natural logarithm. The overall sum of entries $\mathrm{clr}(\mathbf{T})$ is 0. Therefore, it is not full rank and spans a $(IJ-1)$-dimensional subspace of $\mathbb{R}^{IJ}$. The inverse of this transformation is given as

$$\mathbf{T} = \mathcal{C} \exp[\mathrm{clr}(\mathbf{T})],$$

where exp operates component-wise. $\mathcal{C}$ is the closure operator, which is defined as

$$\mathcal{C}\mathbf{T} = \left(\frac{t_{11}}{\sum_{i=1}^{I}\sum_{j=1}^{J} t_{ij}}, \ldots, \frac{t_{IJ}}{\sum_{i=1}^{I}\sum_{j=1}^{J} t_{ij}} \right),$$

where $\mathbf{T} \in \left\{ \mathbf{t} \in \mathbb{R}^{IJ} \mid t_{ij} > 0 \right\}$. Therefore, the clr transformation is a one-to-one mapping from $\mathcal{S}^{IJ}$ to $(IJ-1)$-dimensional subspace of $\mathbb{R}^{IJ}$.

The Aitchison inner product, $\langle \cdot, \cdot \rangle_a$, the norm, $\| \cdot \|_a$, and the distance, $d_a(\cdot, \cdot)$, are defined as

$$\langle \cdot, \cdot \rangle_a = \langle \mathrm{clr}(\cdot), \mathrm{clr}(\cdot) \rangle, \ \ \| \cdot \|_a = \| \mathrm{clr}(\cdot) \|, \ \ d_a(\cdot, \cdot) = d(\mathrm{clr}(\cdot), \mathrm{clr}(\cdot)),$$

where $\langle \cdot, \cdot \rangle$, $\| \cdot \|$, and $d(\cdot, \cdot)$ denote the ordinary Euclidean inner product, norm, and distance in $\mathbb{R}^{IJ}$, respectively. The simplex equipped with this metric in addition to perturbation and powering is a $(IJ-1)$-dimensional Euclidean space (Billheimer *et al.* 2001; Pawlowsky-Glahn and Egozcue 2001). Therefore, one table can be added to another one, the distance between two PTs and the norm of a PT can be measured, and a PT can be orthogonally projected onto some linear subspace.

Egozcue *et al.* (2015) proposed a method to study the independence between two categorical variables using the Aitchison geometry. They showed that independent PTs constitute an $(I-1)(J-1)$-dimensional linear subspace of $\mathcal{S}^{IJ}$, which is denoted by $\mathcal{S}^{IJ}_{\mathrm{ind}}$, and any PT can be orthogonally projected onto this subspace. As a result, any PT, $\mathbf{T}$, can be uniquely decomposed as

$$\mathbf{T} = \mathbf{T}_{\mathrm{ind}} \oplus \mathbf{T}_{\mathrm{int}},$$

where $\mathbf{T}_{\mathrm{ind}}$ is the closest independent PT in the sense of the Aitchison distance and $\mathbf{T}_{\mathrm{int}}$ is the interaction PT orthogonal to $\mathbf{T}_{\mathrm{ind}}$. The square Aitchison norm of $\mathbf{T}_{\mathrm{int}}$ is called *simplicial deviance*, $\Delta^2(\mathbf{T})$. Its relative version to falls within $[0, 1]$, which is called the *relative simplicial deviance*, $R^2_{\Delta}(\mathbf{T})$. $\Delta^2(\mathbf{T})$ and $R^2_{\Delta}(\mathbf{T})$ can be used to measure the departure from independence. They also developed another method to test independence using multinomial simulations.

3. Symmetric PT

In this section, we consider the symmetric structure of a PT in the similar manner as Egozcue *et al.* (2015) The probability that an individual will fall in the (i, j) cell is equal to the probability that an individual will fall in the (j, i) cell because the symmetry of a two-way contingency table is normally defined as corresponding to a PT, $\mathbf{T}$, whose transpose, $T(\mathbf{T})$, is the same. Therefore, a symmetric PT is defined as follows:

Definition 1 (Symmetric PT)**.** Let $\mathbf{T} \in \mathcal{S}^{I^2}$ be an $I \times I$ table. $\mathbf{T}$ is a symmetric PT, if

$$\mathbf{T} = T(\mathbf{T}).$$

This leads to Lemma 1 and Lemma 1 states that symmetric PTs form a subspace of $\mathcal{S}^{I^2}$.

Lemma 1 (Subspace of symmetric PTs)**.** *Let $\mathcal{S}^{I^2}_{\mathrm{sym}}$ be the set of $I \times I$ symmetric tables in $\mathcal{S}^{I^2}$. Then $\mathcal{S}^{I^2}_{\mathrm{sym}}$ is an $(I-1)(I+2)/2$-dimentional subspace of $\mathcal{S}^{I^2}$.*

Egozcue and Maldonado (2021) state the same for elements of $R^{+}_{>0} = \{(x_{11}, x_{12}, \ldots, x_{II}) \in \mathbb{R}^{I^2} \mid x_{ij} > 0\}$ with log-geometry (non-normalized Aitchison geometry); Birtea and Gavra (2024) show the relation between them, that is, $R^{+}_{>0}$ with log-geometry is a Lie group and a simplex with Aitchison geometry is associated to its quotient Lie group with respect to the equivalence relation.

Table 3: MLE(%) of Table 1 under the symmetry hypothesis

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade	Total
Highest grade	20.33	3.34	1.61	0.68	25.97
Second grade	3.34	20.22	5.31	1.07	29.95
Third grade	1.61	5.31	23.70	2.57	33.19
Lowest grade	0.68	1.07	2.57	6.58	10.90
Total	25.97	29.95	33.19	10.90	100.00

Table 4: Nearest symmetric PT(%) corresponding to Table 1

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade	Total
Highest grade	20.36	3.34	1.61	0.65	25.96
Second grade	3.34	20.25	5.30	1.07	29.96
Third grade	1.61	5.30	23.70	2.57	33.20
Lowest grade	0.65	1.07	2.57	6.59	10.88
Total	25.96	29.96	33.20	10.88	100.00

According to Lemma 1 and Hilbert projection theorem, for every PT $\mathbf{T} \in \mathcal{S}^{I^2}$, there is a unique PT, $\mathbf{T}_{\text{sym}} \in \mathcal{S}^{I^2}_{\text{sym}}$, for which $\|\mathbf{R} \ominus \mathbf{T}\|_a$ is minimized over the PTs, $\mathbf{R} \in \mathcal{S}^{I^2}_{\text{sym}}$. Moreover, $\mathbf{T}$ is uniquely decomposed into $\mathbf{T}_{\text{sym}}$ and

$$\mathbf{T}_{\text{skew}} = \mathbf{T} \oplus ((-1) \odot \mathbf{T}_{\text{sym}}) = \mathbf{T} \ominus \mathbf{T}_{\text{sym}},$$

which are orthogonal to each other ($\langle \mathbf{T}_{\text{sym}}, \mathbf{T}_{\text{skew}} \rangle_a = 0$).

Theorem 1 describes orthogonal projections onto the subspace of symmetric PTs.

Theorem 1. *Let $\mathbf{T} \in \mathcal{S}^{I^2}$ be an $I \times I$ table. The nearest symmetric PT to $\mathbf{T}$ is $\mathbf{T}_{\text{sym}} = 0.5 \odot (\mathbf{T} \oplus T(\mathbf{T}))$ in the sense of the Aitchison distance in $\mathcal{S}^{I^2}$.*

Table 3 shows MLE under the symmetry hypothesis, while Table 4 is the symmetric PT, which is the orthogonal projection of Stuart's example onto the symmetric subspace. The latter coincides with the minimum discrimination information (MDI) estimator satisfying symmetry in Ireland, Ku, and Kullback (1969).

As a result of the decomposition, the *departure of symmetry* can be measured using the square-Aitchison norm.

Corollary 1. *Let $\mathbf{T} \in \mathcal{S}^{I^2}$ be an $I \times I$ table. Consider the decomposition, $\mathbf{T} = \mathbf{T}_{\text{sym}} \oplus \mathbf{T}_{\text{skew}}$, defined in Theorem 1. Then,*

$$\|\mathbf{T}\|_a^2 = \|\mathbf{T}_{\text{sym}}\|_a^2 + \|\mathbf{T}_{\text{skew}}\|_a^2.$$

Next, we describe the properties of the skew-symmetric PT, $\mathbf{T}_{\text{skew}}$, and the decomposition of the square-Aitchison norm into symmetric and skew-symmetric parts.

Theorem 2. *Let $\mathbf{T} \in \mathcal{S}^{I^2}$ be an $I \times I$ table. The nearest symmetric PT of $\mathbf{T}_{\text{skew}}$ is the neutral element of $\mathcal{S}^{I^2}$, $\mathbf{N}$, whose elements are all equal. In addition, the (i,j) and (j,i) cells of $\text{clr}(\mathbf{T}_{\text{skew}})$ sum to 0.*

From the properties shown in Theorem 2, $\mathbf{T}_{\text{skew}}$ is called a skew-symmetric PT. The nearest symmetric PT of $\mathbf{T}_{\text{skew}}$ is the neutral element with $\|\mathbf{N}\|_a = 0$. Therefore, $\|\mathbf{T}_{\text{skew}}\|_a^2$ is a proper measure of asymmetry. The following theorem states this important property for $\mathbf{T}_{\text{skew}}$.

Theorem 3. *Let* $\mathbf{T} \in \mathcal{S}^{I^2}$ *be an* $I \times I$ *table and* $\mathbf{R} \in \mathcal{S}^{I^2}_{\text{sym}}$ *be any symmetric* $I \times I$ *PT. Then the perturbed PT,* $\mathbf{V} = \mathbf{R} \oplus \mathbf{T}$*, is orthogonally decomposed into* $\mathbf{V} = \mathbf{V}_{\text{sym}} \oplus \mathbf{T}_{\text{skew}}$*, where* $\mathbf{V}_{\text{sym}} = \mathbf{R} \oplus \mathbf{T}_{\text{sym}}$*.*

The following two definitions adopt $\|\mathbf{T}_{\text{skew}}\|_a^2$ and its relative version as a measure of asymmetry.

Definition 2 (Simplicial skewness)**.** Let $\mathbf{T} \in \mathcal{S}^{I^2}$ be an $I \times I$ table with an orthogonal decomposition $\mathbf{T} = \mathbf{T}_{\text{sym}} \oplus \mathbf{T}_{\text{skew}}$ (Corollary 1). The simplicial skewness, which measures the degree of symmetry, is defined as $E^2(\mathbf{T}) = \|\mathbf{T}_{\text{skew}}\|_a^2$ with $0 \leq E^2(\mathbf{T}) < \infty$.

Definition 3 (Relative simplicial skewness)**.** Using the notation in Definition 2, the relative simplicial skewness is

$$R_E^2(\mathbf{T}) = \frac{E^2(\mathbf{T})}{\|\mathbf{T}\|_a^2} = \frac{\|\mathbf{T}_{\text{skew}}\|_a^2}{\|\mathbf{T}\|_a^2} = \frac{\|\mathbf{T}\|_a^2 - \|\mathbf{T}_{\text{sym}}\|_a^2}{\|\mathbf{T}\|_a^2},\ 0 \leq R_E^2(\mathbf{T}) \leq 1.$$

$E^2(\mathbf{T})$ has a desirable property. It offers a measure for the departure of symmetry, which we call the *symmetric invariant.* That is, when each symmetric cell of $\mathbf{T}$ is multiplied by arbitrary strictly positive constants, $\mathbf{T}_{\text{skew}}$ remains the same. When PT, $\mathbf{T}$, is described by

$$p_{ij} = \begin{cases} \delta_{ij}\phi_{ij} & (i < j), \\ \phi_{ij} & (i \geq j), \end{cases}$$

where $\phi_{ij} = \phi_{ji}$, the simplicial skewness is expressed as

$$E^2(\mathbf{T}) = \frac{1}{2}\sum_{i<j}(\log \delta_{ij})^2.$$

Consequently, it measures only the asymmetric components under an asymmetric model such as the conditional symmetry (CS) model or the diagonals-parameter symmetry (DPS) model. (McCullagh 1978; Goodman 1979)

The relative simplicial skewness is not symmetric invariant because its denominator changes when another symmetric PT is perturbed to $\mathbf{T}$. However, the relative simplicial skewness indicates whether the skew-symmetric PT is in the overall norm. Therefore, a low value of $R_E^2(\mathbf{T})$ may be due to either a small simplicial skewness or a large norm of symmetric PT.

When the simplicial skewness is large, it is useful to determine which cell is responsible for the departure of symmetry. $\mathbf{T}_{\text{skew}}$ itself provides a detailed description of each cell. However, reading the description directly from the cells is incomprehensible because it must be read in contrast to $1/I^2$. Therefore, $\mathbf{T}_{\text{skew}}$ is transformed using the centered log-ratio transformation,and $\text{clr}(\mathbf{T}_{\text{skew}})$ is analyzed like a log-linear model with its constraint summing to 0. This detailed description of skewness can be obtained using the following definition of cell skewness:

Definition 4 (Cell skewness)**.** Let $\mathbf{T}_{\text{skew}} \in \mathcal{S}^{I^2}$ be a skew-symmetric $I \times I$ table. The coefficient $\text{clr}_{ij}(\mathbf{T}_{\text{skew}})$ is called the (i, j) cell skewness.

A direct result is that the square-simplicial skewness is the sum of square-cell skewnesses:

$$E^2(\mathbf{T}) = \|\mathbf{T}_{\text{skew}}\|_a^2 = \sum_{i=1}^{I}\sum_{j=1}^{I}(\text{clr}_{ij}(\mathbf{T}_{\text{skew}}))^2$$

Because cell skewnesses are functions of the symmetric invariant, $\mathbf{T}_{\text{skew}}$, cell skewnesses are symmetric invariants.

For simplicity, the cell-skewness analysis is presented as an $I \times I$ table, which is called a *skewness array.* In a skewness array, the entries are the signed proportions or percents of the simplicial skewness, $\mathrm{sgn}(\mathrm{clr}_{ij}(\mathbf{T}_{\mathrm{skew}}))(\mathrm{clr}_{ij}(\mathbf{T}_{\mathrm{skew}}))^2/E^2(\mathbf{T})$. The strength of the skewness is quantified, and the direction of the skewness is illustrated by the sign if its probability is a surplus or deficit compared with the corresponding symmetry PT.

Theorem 4 shows how $\mathbf{T}_{\mathrm{skew}}$ and the simplicial skewness change under weighting operations.

Theorem 4 (Weighting properties)**.** *Let* $\mathbf{T} \in \mathcal{S}^{I^2}$ *be an* $I \times I$ *table. Consider a symmetric PT denoted as* $\mathbf{W}$ *and a real number,* w*. Then,*

$$\begin{aligned} \|\mathbf{T} \oplus \mathbf{W}\|_a^2 &= \|\mathbf{T}_{\mathrm{sym}} \oplus \mathbf{W}\|_a^2 + \|\mathbf{T}_{\mathrm{skew}}\|_a^2, \\ \|w \odot \mathbf{T}\|_a^2 &= w^2 \cdot \|\mathbf{T}_{\mathrm{sym}}\|_a^2 + w^2 \cdot \|\mathbf{T}_{\mathrm{skew}}\|_a^2. \end{aligned} \tag{3}$$

These properties in Theorem 4 have practical consequences. Sampling of the population may affect the symmetric PTs of $\mathbf{T}$, but does not affect the analysis. Perturbation by $\mathbf{W}$ works as a weighting of the symmetric sampling without affecting the simplicial skewness. Weighting by powering can appear in the case of repeated perturbation, which is the likelihood of an independent sample being produced.

4. Symmetry test

Here, we present a method to test the symmetry of contingency tables derived on the theory developed in the previous sections. We use the sample simplicial skewness, $E^2(\hat{\mathbf{T}})$ (Definition 2), and the sample relative simplicial skewness, $R_E^2(\hat{\mathbf{T}})$, as the test statistic.

Consider a sample $I \times I$ contingency table obtained by multinomial sampling, and the corresponding estimated PT, which is denoted $\hat{\mathbf{T}}$. To use the Aitchson geometry, the entries of PT must be strictly positive and not 0. Therefore, $\hat{\mathbf{T}}$ must be an estimator that avoids zero cells. Perks (1947) proposed a Bayesian estimator using the uniform Dirichlet prior, which is given as

$$\hat{t}_{ij} = \frac{n_{ij} + 1/I^2}{n+1}, \quad i, j = 1, 2, \ldots, I, \tag{4}$$

Thus, (4) is biased toward $\|\mathbf{N}\|_a = 0$. If we interpret the Aitchison norm as information like Egozcue and Vera (2018), it is natural to set it to 0 before acquiring data. Let $\mathbf{T}$ be the true unkown multinomial probabilities and $\mathcal{S}^{I^2}_{\mathrm{sym}}$ be the subspace of symmetric tables, as in Section 3, Definition 1. Then the hypotheses,

$$H_0: \quad \mathbf{T} = \mathbf{T}_{\mathrm{sym}} \in \mathcal{S}^{I^2}_{\mathrm{sym}} \ vs. \ H_1; \quad \mathbf{T} \notin \mathcal{S}^{I^2}_{\mathrm{sym}},$$

are tested. Under H_0, we must estimate which symmetric table corresponds to $\mathbf{T} = \mathbf{T}_{\mathrm{sym}}$. For the Bowker test statistic (1), the arithmetic mean of the PT and the transpose of the PT are specified. According to the theory developed in the previous sections, $\mathbf{T}_{\mathrm{sym}}$ is set to $\hat{\mathbf{T}}_{\mathrm{sym}}$. That is, the symmetric PT obtained from the orthogonal projection onto a symmetric subspace is the same form of estimator obtained in Ireland *et al.* (1969). Therefore, it has the same asymptotic properties as the BAN estimator.

Using a parametric bootstrap, we investigate the sample distribution of two statistics. The test procedure is as follows:

1. Estimate the sample PT, $\hat{\mathbf{T}}$, using the estimator that avoids zero cells;
2. Compute $\hat{\mathbf{T}}_{\mathrm{sym}}$ from $\hat{\mathbf{T}}$;
3. For $\hat{\mathbf{T}}_{\mathrm{sym}}$, compute $\hat{\mathbf{T}}_{\mathrm{skew}} = \hat{\mathbf{T}} \ominus \hat{\mathbf{T}}_{\mathrm{sym}}$, $E^2(\hat{\mathbf{T}})$, and $R_E^2(\hat{\mathbf{T}})$;
4. Simulate 10^4 multinomial samples with the probabilities specified in $H_0 : \mathbf{T} = \hat{\mathbf{T}}_{\mathrm{sym}}$. For each simulated contingency table, repeat steps 1 to 3;

Table 5: Sizes and powers of six test statistics for 4×4 contingency tables with $\hat{\mathbf{T}}_{\text{sym}} \oplus (r \odot \hat{\mathbf{T}}_{\text{skew}})$: simplicial skewness ($E^2$), relative simplicial skewness (R^2_E), Pearson statistic (χ^2), log-likelihood ratio statistic (L), and their bootstrap versions (χ^2_B) and (L_B)

		r						
Method	n	-1.5	-1.0	-0.5	0	0.5	1.0	1.5
E^2	50	0.1542	0.1405	0.1391	0.1379	0.1451	0.1432	0.1574
	100	0.1017	0.0941	0.0881	0.0852	0.0871	0.0914	0.0996
	1000	0.1873	0.0772	0.0267	0.0155	0.0275	0.0754	0.1878
	5000	0.9776	0.7060	0.2029	0.0437	0.2086	0.7064	0.9793
	7477	0.9989	0.8860	0.3061	0.0492	0.3067	0.8878	0.9988
R^2_E	50	0.2630	0.2493	0.2440	0.2414	0.2502	0.2528	0.2706
	100	0.2448	0.2252	0.2120	0.2074	0.2143	0.2235	0.2433
	1000	0.2622	0.1161	0.0460	0.0268	0.0474	0.1160	0.2627
	5000	0.9791	0.7153	0.2132	0.0476	0.2167	0.7178	0.9805
	7477	0.9990	0.8903	0.3139	0.0512	0.3144	0.8919	0.9989
χ^2_B	50	0.1420	0.1365	0.1289	0.1254	0.1329	0.1300	0.1451
	100	0.1110	0.1054	0.0966	0.0937	0.0949	0.1009	0.1138
	1000	0.3476	0.1565	0.0623	0.0391	0.0627	0.1576	0.3512
	5000	0.9911	0.7729	0.2093	0.0493	0.2245	0.7738	0.9907
	7477	0.9998	0.9317	0.3302	0.0518	0.3270	0.9312	0.9997
L_B	50	0.1533	0.1450	0.1377	0.1339	0.1411	0.1408	0.1545
	100	0.1187	0.1091	0.0984	0.0952	0.0979	0.1061	0.1206
	1000	0.3701	0.1674	0.0683	0.0444	0.0698	0.1724	0.3695
	5000	0.9912	0.7717	0.2084	0.0495	0.2240	0.7718	0.9907
	7477	0.9998	0.9313	0.3297	0.0519	0.3266	0.9308	0.9997
χ^2	50	0.0050	0.0046	0.0036	0.0029	0.0049	0.0044	0.0056
	100	0.0282	0.0218	0.0171	0.0138	0.0153	0.0203	0.0255
	1000	0.3720	0.1684	0.0725	0.0471	0.0729	0.1771	0.3715
	5000	0.9912	0.7689	0.2051	0.0493	0.2209	0.7706	0.9907
	7477	0.9999	0.9307	0.3281	0.0520	0.3253	0.9300	0.9997
L	50	0.0511	0.0444	0.0390	0.0386	0.0460	0.0444	0.0489
	100	0.0913	0.0776	0.0691	0.0671	0.0690	0.0800	0.0940
	1000	0.3978	0.1879	0.0828	0.0543	0.0834	0.1953	0.3989
	5000	0.9913	0.7739	0.2083	0.0505	0.2250	0.7739	0.9911
	7477	0.9999	0.9315	0.3306	0.0530	0.3281	0.9313	0.9997

5. Compare the values of the test statistic and significant critical points.

To verify the size and power of the test, we conducted the Monte Carlo simulations. We used the data set in Table 1 and its corresponding PT estimated by $\hat{\mathbf{T}}$, which was proposed in Perks (1947). Then 10^4 contingency tables with $n = 50, 100, 1000, 5000$, and 7477 trials under $\hat{\mathbf{T}}_{\text{sym}} \oplus (r \odot \hat{\mathbf{T}}_{\text{skew}})$ were generated. The probability of rejection was estimated as the proportion of times that the test statistic was larger than the 95% quantile of the distribution generated under the null hypothesis, H_0. The probability of rejection should approximately coincide with a significance level $\alpha = 0.05$. As the departure from symmetry value increase, the probability of rejection should increase up to 1.

The sizes and powers of six test statistics were compared: simplicial skewness (E^2), relative simplicial skewness (R^2_E), Pearson statistic (χ^2), loglikelihood ratio statistic (L), and their bootstrap versions (χ^2_B) and (L_B). χ^2 and L were computed as (1) and (2) and, if $n_{ij}+n_{ji} = 0$, then the corresponding summand was set to 0. χ^2_B and L_B were computed by a parametric bootstrap using the following test statistic:

$$\chi_B^2 = n \sum_{i=1}^{I} \sum_{j=1}^{I} \frac{(\hat{t}_{ij} - \hat{E}_{ij})^2}{\hat{E}_{ij}}, \ L_B = 2n \sum_{i=1}^{I} \sum_{j=1}^{I} \hat{t}_{ij} \log \frac{\hat{t}_{ij}}{\hat{E}_{ij}},$$

where t_{ij} is (i,j) cell of $\hat{\mathbf{T}}$ and $\hat{E}_{ij}$ is (i,j) element of $\hat{\mathbf{T}}_{\text{sym}}$. Table 5 shows the simulation results. In all test statistics, the probability of rejection with a large sample ($n = 5000, 7477$), converges to 0.05 under the null hypothesis ($r = 0$) and increases to 1 as the value of r increases. Due to the unstable estimation of zero cells in contingency tables with small samples, we cannot control the type I error of all test statistics. The controls of the type I error for χ^2 and L begin to stabilize from $n = 100$, while those for other statistics begin to stabilize from $n = 1000$.

5. Illustrative example

Table 6: Statistics

Statistics	Sample value	0.05 crit-v	p-value
$\Vert\hat{\mathbf{T}}\Vert_a^2$	20.560		
$\Vert\hat{\mathbf{T}}_{\text{sym}}\Vert_a^2$	20.341		
$E^2(\hat{\mathbf{T}})$	0.219	0.131	49×10^{-4}
$R_E^2(\hat{\mathbf{T}})$	0.011	0.006	46×10^{-4}

Table 1 shows the contingency table reported by Stuart (1955). Their table classified 7477 women by left and right eyesight (Highest, Second, Third, and Lowest grades). Table 4 shows the consequence of orthogonal projection onto a symmetric subspace. The simplicial skewness and other summary statistic are described in Table 6. Table 7 and 8 show the skewness table and array, respectively.

The two symmetry tests show p-values below the 0.05 significant level. Therefore, the symmetry hypothesis is rejected. This agrees with the method proposed by Bowker (1948). Once the hypothesis of symmetry is rejected, the next question is how the PT deviates from symmetry. The skewness array in Table 8 shows that most departures from symmetry are due to positive skewness in the Highest-Lowest grade and a negative skewness in the Lowest-Highest grade. The bold skewnesses suggest that the Highest-Lowest and Lowest-Highest grades are skewed in the female employees at Royal Ordnance factories.

Table 7: $\text{clr}(\hat{\mathbf{T}}_{\text{skew}})$ for Table 1

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade
Highest grade	0.000	0.064	0.029	0.363
Second grade	−0.064	0.000	0.088	−0.025
Third grade	−0.029	−0.088	0.000	0.068
Lowest grade	−0.363	0.025	−0.068	0.000

6. Conclusions

PTs are commonly obtained from a multinomial sampling. PTs with strictly positive entries are elements of a simplex. Mathematical tools can be applied because a simplex with the structure called the Aitchison geometry is a Euclidean space. Tools can identify the set of symmetric PTs as a linear vector subspace. Given a PT, we obtained the nearest symmetric

Table 8: Skewness array for Table 1

Right eye \ Left eye	Highest grade	Second grade	Third grade	Lowest grade
Highest grade	0.00	1.87	0.38	**41.80**
Second grade	−1.87	0.00	3.56	−0.28
Third grade	−0.38	−3.56	0.00	2.10
Lowest grade	−**41.80**	0.28	−2.10	0.00

PT and the orthogonal decomposition into the symmetric PT and the skew-symmetric one by projecting the original PT onto a symmetric subspace. Hence, the nearest symmetric PT is a perturbation of the original PT and its transposition.

The square-distance to the subspace of symmetric tables, which is called the *simplicial skewness*, is the Aitchison square-norm of the skew-symmetric table. The newly developed method using simplicial skewness and its relative value as a test statistic is an understandable analytical tool, which should provide useful and practical results.

Acknowledgements

This work was supported by JSPS KAKENHI (Grant Number 20K03756). We would like to thank the editor for their helpful support and two referees for their valuable comments and suggestions.

References

Aitchison J (1982). “The Statistical Analysis of Compositional Data.” *Journal of the Royal Statistical Society: Series B (Methodological)*, **44**(2), 139–160. ISSN 0035-9246. doi:10.1111/j.2517-6161.1982.tb01195.x.

Aitchison J (1986). *The Statistical Analysis of Compositional Data*. Monographs on Statistics and Applied Probability (Series). Chapman and Hall, London. ISBN 0-412-28060-4 978-0-412-28060-3.

Aitchison J, Barceló-Vidal C, Martín-Fernández JA, Pawlowsky-Glahn V (2001). “Reply to Letter to the Editor by S. Rehder and U. Zier.” *Mathematical Geology*, **33**(7), 849–860. ISSN 1573-8868. doi:10.1023/A:1010954915624.

Billheimer D, Guttorp P, Fagan WF (2001). “Statistical Interpretation of Species Composition.” *Journal of the American Statistical Association*, **96**(456), 1205–1214. ISSN 0162-1459. doi:10.1198/016214501753381850.

Birtea P, Gavra I (2024). “Parametric Lie Group Structures on the Probabilistic Simplex and Generalized Compositional Data.” *Journal of Geometry and Physics*, **196**, 105061. ISSN 0393-0440. doi:10.1016/j.geomphys.2023.105061.

Bowker AH (1948). “A Test for Symmetry in Contingency Tables.” *Journal of the American Statistical Association*, **43**(244), 572–574. ISSN 0162-1459. doi:10.1080/01621459.1948.10483284.

Egozcue JJ, Maldonado WL (2021). “An Interpretable Orthogonal Decomposition of Positive Square Matrices.” In P Filzmoser, K Hron, JA Martín-Fernández, J Palarea-Albaladejo (eds.), *Advances in Compositional Data Analysis: Festschrift in Honour of Vera Pawlowsky-Glahn*, pp. 1–18. Springer International Publishing, Cham. ISBN 978-3-030-71175-7. doi:10.1007/978-3-030-71175-7_1.

Egozcue JJ, Pawlowsky-Glahn V, Templ M, Hron K (2015). "Independence in Contingency Tables Using Simplicial Geometry." *Communications in Statistics - Theory and Methods*, **44**(18), 3978–3996. ISSN 0361-0926. doi:10.1080/03610926.2013.824980.

Egozcue JJ, Vera PG (2018). "Evidence Functions: A Compositional Approach to Information." *SORT-Statistics and Operations Research Transactions*, **42**(2), 101–124.

Goodman LeoA (1979). "Multiplicative Models for Square Contingency Tables with Ordered Categories." *Biometrika*, **66**(3), 413–418. ISSN 0006-3444. doi:10.1093/biomet/66.3.413.

Ireland CT, Ku HH, Kullback S (1969). "Symmetry and Marginal Homogeneity of an r × r Contingency Table." *Journal of the American Statistical Association*, **64**(328), 1323–1341. ISSN 01621459. doi:10.2307/2286071. 2286071.

Kateri M, Agresti A (2007). "A Class of Ordinal Quasi-Symmetry Models for Square Contingency Tables." *Statistics & Probability Letters*, **77**(6), 598–603. ISSN 0167-7152. doi:10.1016/j.spl.2006.09.015.

Kateri M, Papaioannou T (1997). "Asymmetry Models for Contingency Tables." *Journal of the American Statistical Association*, **92**(439), 1124–1131. ISSN 0162-1459. doi:10.1080/01621459.1997.10474068.

McCullagh P (1978). "A Class of Parametric Models for the Analysis of Square Contingency Tables with Ordered Categories." *Biometrika*, **65**(2), 413–418. ISSN 0006-3444. doi:10.1093/biomet/65.2.413.

Momozaki T, Cho K, Nakagawa T, Tomizawa S (2023). "Improving the Accuracy of Estimating Indexes in Contingency Tables Using Bayesian Estimators." *Journal of Statistical Theory and Practice*, **18**(1), 1. ISSN 1559-8616. doi:10.1007/s42519-023-00353-4.

Pawlowsky-Glahn V, Egozcue JJ (2001). "Geometric Approach to Statistical Analysis on the Simplex." *Stochastic Environmental Research and Risk Assessment*, **15**(5), 384–398. ISSN 1436-3259. doi:10.1007/s004770100077.

Perks W (1947). "Some Observations on Inverse Probability Including a New Indifference Rule." *Journal of the Institute of Actuaries*, **73**(2), 285–334. ISSN 0020-2681. doi:10.1017/S0020268100012270.

Stuart A (1955). "A Test for Homogeneity of the Marginal Distributions in A Two-Way Classification." *Biometrika*, **42**(3-4), 412–416. ISSN 0006-3444. doi:10.1093/biomet/42.3-4.412.

Tahata K (2020). "Separation of Symmetry for Square Tables with Ordinal Categorical Data." *Japanese Journal of Statistics and Data Science*, **3**(2), 469–484. doi:10.1007/s42081-019-00066-8.

Tahata K (2022). "Advances in Quasi-Symmetry for Square Contingency Tables." *Symmetry*, **14**(5). ISSN 2073-8994. doi:10.3390/sym14051051.

Tomizawa S (1994). "Two Kinds of Measures of Departure from Symmetry in Square Contingency Tables Having Nominal Categories." *Statistica Sinica*, **4**(1), 325–334. ISSN 10170405, 19968507. 24305290.

Tomizawa S, Seo T, Yamamoto H (1998). "Power-Divergence-Type Measure of Departure from Symmetry for Square Contingency Tables That Have Nominal Categories." *Journal of Applied Statistics*, **25**(3), 387–398. ISSN 0266-4763. doi:10.1080/02664769823115.

A. Proof of Lemma 1

Lemma 1 (Subspace of symmetric PTs)**.** *Let* $\mathcal{S}_{\text{sym}}^{I^2}$ *be the set of* $I \times I$ *symmetric tables in* $\mathcal{S}^{I^2}$*. Then* $\mathcal{S}_{\text{sym}}^{I^2}$ *is an* $(I-1)(I+2)/2$*-dimentional subspace of* $\mathcal{S}^{I^2}$*.*

Proof. For any $\mathbf{R} = (r_{ij}^{\text{sym}}), \mathbf{T} = (t_{ij}^{\text{sym}}) \in \mathcal{S}_{\text{sym}}^{I^2}$, and $\alpha \in \mathbb{R}$, we have that the (i,j) cell of $\alpha \odot (\mathbf{R} \oplus \mathbf{T})$ is

$$\frac{(r_{ij}^{\text{sym}})^\alpha (t_{ij}^{\text{sym}})^\alpha}{\sum_{i=1}^{I}\sum_{j=1}^{I}(r_{ij}^{\text{sym}})^\alpha (t_{ij}^{\text{sym}})^\alpha} = \frac{(r_{ji}^{\text{sym}})^\alpha (t_{ji}^{\text{sym}})^\alpha}{\sum_{j=1}^{I}\sum_{i=1}^{I}(r_{ji}^{\text{sym}})^\alpha (t_{ji}^{\text{sym}})^\alpha}.$$

The right-hand side of this equation is the (i,j) cell of $T(\alpha \odot (\mathbf{R} \oplus \mathbf{T}))$. Therefore, $\alpha \odot (\mathbf{R} \oplus \mathbf{T}) = T(\alpha \odot (\mathbf{R} \oplus \mathbf{T}))$ and $\mathcal{S}_{\text{sym}}^{I^2}$ is closed under perturbation and powering. Hence, it is a subspace of $\mathcal{S}^{I^2}$. The dimension of $\mathcal{S}_{\text{sym}}^{I^2}$ is $(I-1)(I+2)/2$ because the upper triangle of the square PT contains $I(I-1)/2$ cells and the dimension is reduced by one due to the constraint that the cells add up to a constant. □

B. Proof of Theorem 1

Theorem 1. *Let* $\mathbf{T} \in \mathcal{S}^{I^2}$ *be an* $I \times I$ *table. The nearest symmetric PT to* $\mathbf{T}$ *is* $\mathbf{T}_{\text{sym}} = 0.5 \odot (\mathbf{T} \oplus T(\mathbf{T}))$ *in the sense of the Aitchison distance in* $\mathcal{S}^{I^2}$*.*

Proof. Let $\mathbf{T}_{\text{sym}}^*$ be $0.5 \odot (\mathbf{T} \oplus T(\mathbf{T}))$. Then, the (i,j) elements of $\mathbf{T}_{\text{sym}}^*$ is

$$t_{ij}^{\text{sym}} = \frac{\sqrt{t_{ij}t_{ji}}}{\sum_{k=1}^{I}\sum_{l=1}^{I}\sqrt{t_{kl}t_{lk}}}.$$

To prove this theorem, for any $\mathbf{R} \in \mathcal{S}_{\text{sym}}^{I^2}$, whose (i,j) cell is $r_{ij}^{\text{sym}} = r_{ji}^{\text{sym}}$, we must show $\langle \mathbf{T} \ominus \mathbf{T}_{\text{sym}}^*, \mathbf{T}_{\text{sym}}^* \ominus \mathbf{R} \rangle_a = 0$. The (i,j) entry of $\text{clr}(\mathbf{T} \ominus \mathbf{T}_{\text{sym}}^*)$ is given as

$$\begin{aligned}
\text{clr}_{ij}(\mathbf{T} \ominus \mathbf{T}_{\text{sym}}^*) &= \log \frac{\frac{t_{ij}}{\sqrt{t_{ij}t_{ji}}}}{\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{t_{kl}}{\sqrt{t_{kl}t_{lk}}}} - \log\left(\prod_{i=1}^{I}\prod_{j=1}^{I}\frac{\frac{t_{ij}}{\sqrt{t_{ij}t_{ji}}}}{\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{t_{kl}}{\sqrt{t_{kl}t_{lk}}}}\right)^{1/I^2} \\
&= \log \frac{t_{ij}}{\sqrt{t_{ij}t_{ji}}} - \log \sum_{k=1}^{I}\sum_{l=1}^{I}\frac{t_{kl}}{\sqrt{t_{kl}t_{lk}}} - \frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\log\frac{t_{ij}}{\sqrt{t_{ij}t_{ji}}} + \log\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{t_{kl}}{\sqrt{t_{kl}t_{lk}}} \\
&= \log\sqrt{\frac{t_{ij}}{t_{ji}}} - \frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\log\sqrt{\frac{t_{ij}}{t_{ji}}} \\
&= \log\sqrt{\frac{t_{ij}}{t_{ji}}} - \left(\frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\log\sqrt{t_{ij}} - \frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\log\sqrt{t_{ji}}\right) \\
&= \log\sqrt{\frac{t_{ij}}{t_{ji}}}
\end{aligned}$$

and the (i,j) entry of $\mathrm{clr}(\mathbf{T}^*_{\mathrm{sym}} \ominus \mathbf{R})$ is given as

$$\begin{aligned}
\mathrm{clr}_{ij}(\mathbf{T}^*_{\mathrm{sym}} \ominus \mathbf{R}) &= \log \frac{\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}}}{\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{\sqrt{t_{kl}t_{lk}}}{r^{\mathrm{sym}}_{kl}}} - \log\left(\prod_{i=1}^{I}\prod_{j=1}^{I}\frac{\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}}}{\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{\sqrt{t_{kl}t_{lk}}}{r^{\mathrm{sym}}_{kl}}}\right)^{1/I^2} \\
&= \log\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}} - \log\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{\sqrt{t_{kl}t_{lk}}}{r^{\mathrm{sym}}_{kl}} - \frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\log\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}} + \log\sum_{k=1}^{I}\sum_{l=1}^{I}\frac{\sqrt{t_{kl}t_{lk}}}{r^{\mathrm{sym}}_{kl}} \\
&= \log\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}} - \frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\log\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}}.
\end{aligned}$$

Then we have the Aitchison inner product of $\mathbf{T} \ominus \mathbf{T}^*_{\mathrm{sym}}$ and $\mathbf{T}^*_{\mathrm{sym}} \ominus \mathbf{R}$ as follows.

$$\begin{aligned}
\langle \mathbf{T} \ominus \mathbf{T}^*_{\mathrm{sym}}, \mathbf{T}^*_{\mathrm{sym}} \ominus \mathbf{R}\rangle_a &= \sum_{i=1}^{I}\sum_{j=1}^{I}\mathrm{clr}_{ij}(\mathbf{T} \ominus \mathbf{T}^*_{\mathrm{sym}}) \cdot \mathrm{clr}_{ij}(\mathbf{T}^*_{\mathrm{sym}} \ominus \mathbf{R}) \\
&= \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{\frac{t_{ij}}{t_{ji}}}\right)\left(\log\frac{\sqrt{t_{ij}t_{ji}}}{r^{\mathrm{sym}}_{ij}}\right) \\
&\quad - \left(\frac{1}{I^2}\sum_{k=1}^{I}\sum_{l=1}^{I}\log\frac{\sqrt{t_{kl}t_{lk}}}{r^{\mathrm{sym}}_{kl}}\right)\sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{\frac{t_{ij}}{t_{ji}}}\right) \\
&= \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ij}} - \log\sqrt{t_{ji}}\right)\left(\log\sqrt{t_{ij}} + \log\sqrt{t_{ji}} - \log r^{\mathrm{sym}}_{ij}\right) \\
&\quad - \left(\frac{1}{I^2}\sum_{k=1}^{I}\sum_{l=1}^{I}\log\frac{\sqrt{t_{kl}t_{lk}}}{r^{\mathrm{sym}}_{kl}}\right)\left(\sum_{i=1}^{I}\sum_{j=1}^{I}\log\sqrt{t_{ij}} - \sum_{i=1}^{I}\sum_{j=1}^{I}\log\sqrt{t_{ji}}\right) \\
&= \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ij}}\right)^2 + \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ij}}\right)\left(\log\sqrt{t_{ji}}\right) \\
&\quad - \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ij}}\right)\left(\log r^{\mathrm{sym}}_{ij}\right) - \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ji}}\right)^2 \\
&\quad - \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ij}}\right)\left(\log\sqrt{t_{ji}}\right) + \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ji}}\right)\left(\log r^{\mathrm{sym}}_{ij}\right) \\
&= -\sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ij}}\right)\left(\log r^{\mathrm{sym}}_{ij}\right) + \sum_{i=1}^{I}\sum_{j=1}^{I}\left(\log\sqrt{t_{ji}}\right)\left(\log r^{\mathrm{sym}}_{ji}\right) \\
&= 0.
\end{aligned}$$

□

C. Proof of Theorem 2

Theorem 2. *Let* $\mathbf{T} \in \mathcal{S}^{I^2}$ *be an* $I \times I$ *table. The nearest symmetric PT of* $\mathbf{T}_{\mathrm{skew}}$ *is the neutral element of* $\mathcal{S}^{I^2}$, $\mathbf{N}$, *whose elements are all equal. In addition, the* (i,j) *and* (j,i) *cells of* $\mathrm{clr}(\mathbf{T}_{\mathrm{skew}})$ *sum to* 0.

Proof. Let t_{ij} be the (i,j) cell of $\mathbf{T}$ and $t^{\mathrm{sym}}_{ij}(= \sqrt{t_{ij}t_{ji}}/\sum_{k=1}^{I}\sum_{l=1}^{I}\sqrt{t_{kl}t_{lk}})$ be the (i,j) cell of the nearest symmetric PT, $\mathbf{T}_{\mathrm{sym}}$. The (i,j) cell of $\mathbf{T}_{\mathrm{skew}}$, t^{skew}_{ij}, is

$$t^{\mathrm{skew}}_{ij} = \frac{t_{ij}/t^{\mathrm{sym}}_{ij}}{\sum_{k=1}^{I}\sum_{l=1}^{I} t_{kl}/t^{\mathrm{sym}}_{lk}} = \frac{\sqrt{t_{ij}}/\sqrt{t_{ji}}}{\sum_{k=1}^{I}\sum_{l=1}^{I}\sqrt{t_{kl}}/\sqrt{t_{lk}}}.$$

Therefore, all elements of $\mathbf{T}_{\text{skew}} \oplus T(\mathbf{T}_{\text{skew}})$ are equal. The (i, j) and (j, i) cells of $\text{clr}(\mathbf{T}_{\text{skew}})$ sum to 0 because the geometric means of symmetric cells in $\mathbf{T}_{\text{skew}}$ are equal to the overall geometric mean. □

D. Proof of Theorem 3

Theorem 3. *Let* $\mathbf{T} \in \mathcal{S}^{I^2}$ *be an* $I \times I$ *table and* $\mathbf{R} \in \mathcal{S}^{I^2}_{\text{sym}}$ *be any symmetric* $I \times I$ *PT. Then the perturbed PT,* $\mathbf{V} = \mathbf{R} \oplus \mathbf{T}$*, is orthogonally decomposed into* $\mathbf{V} = \mathbf{V}_{\text{sym}} \oplus \mathbf{T}_{\text{skew}}$*, where* $\mathbf{V}_{\text{sym}} = \mathbf{R} \oplus \mathbf{T}_{\text{sym}}$*.*

Proof. The projection of $\mathbf{V}$ onto the symmetric subspace, $\mathbf{V}_{\text{sym}} = \mathbf{R} \oplus \mathbf{T}_{\text{sym}}$, is a symmetric PT because $\mathbf{R}$ is symmetric PT and $\mathcal{S}^{I^2}_{\text{sym}}$ is a linear subspace. Therefore,

$$\mathbf{V}_{\text{skew}} = \mathbf{V} \ominus \mathbf{V}_{\text{sym}} = \mathbf{T} \ominus \mathbf{T}_{\text{sym}} = \mathbf{T}_{\text{skew}}.$$

□

E. Proof of Theorem 4

Theorem 4 (Weighting properties)**.** *Let* $\mathbf{T} \in \mathcal{S}^{I^2}$ *be an* $I \times I$ *table. Consider a symmetric PT denoted as* $\mathbf{W}$ *and a real number,* w*. Then,*

$$\begin{aligned} \|\mathbf{T} \oplus \mathbf{W}\|_a^2 &= \|\mathbf{T}_{\text{sym}} \oplus \mathbf{W}\|_a^2 + \|\mathbf{T}_{\text{skew}}\|_a^2, \\ \|w \odot \mathbf{T}\|_a^2 &= w^2 \cdot \|\mathbf{T}_{\text{sym}}\|_a^2 + w^2 \cdot \|\mathbf{T}_{\text{skew}}\|_a^2. \end{aligned} \tag{3}$$

Proof. The first equation in (3) is the same result in Theorem 3. Due to the distributive property of powering by w, $w \odot \mathbf{T} = w \odot \mathbf{T}_{\text{sym}} \oplus w \odot \mathbf{T}_{\text{skew}}$. The second equation in (3) is due to the property of the Aitchison norm and the Pythagorean theorem. □

Affiliation:

Keita Nakamura
Department of Information Sciences, Tokyo University of Science
2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
E-mail: 6324703@ed.tus.ac.jp

Tomoyuki Nakagawa
School of Data Science, Meisei University
2-1-1 Hodokubo, Hino City, Tokyo 191-8506, Japan
E-mail: tomoyuki.nakagawa@meisei-u.ac.jp

Kouji Tahata
Department of Information Sciences, Tokyo University of Science
2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan
E-mail: kouji_tahata@rs.tus.ac.jp

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume 53 *Submitted:* 2023-12-08
2024 *Accepted:* 2024-04-15