Send Orders for Reprints to reprints@benthamscience.ae

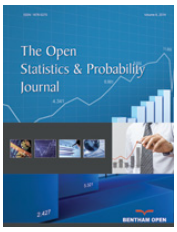
BENTHAM OPEN

CrossMark

The Open Statistics & Probability Journal

Content list available at: www.benthamopen.com/TOSPJ/

DOI: 10.2174/1876527001607010020



RESEARCH ARTICLE

Comparing Measures of Association in 2×2 Probability Tables

Dirk Hasenclever[1,*] and Markus Scholz[1,2]

[1]*Institute for Medical Informatics, Statistics and Epidemiology, Leipzig, Germany*
[2]*LIFE Research Center University of Leipzig, Leipzig, Germany*

Received: March 23, 2015 Revised: April 25, 2016 Accepted: May 02, 2016

Abstract: Measures of association play a role in selecting 2×2 tables exhibiting strong dependence in high-dimensional binary data. Several measures are in use differing on specific tables and in their dependence on the margins. We study a 2-dimensional group of margin transformations on the 3-dimensional manifold $\mathbb{T}$ of all 2×2 probability tables. The margin transformations allow introducing natural coordinates that identify $\mathbb{T}$ with the real 3-space such that the x-axis corresponds to $log\sqrt{odds-ratio}$ and margins vary on planes x =const. We use these coordinates to visualise and compare measures of association with respect to their dependence on the margins given the odds-ratio, their limit behaviour when cells approach zero and their weighting properties. We propose a novel measure of association in which tables with single small entries are up-weighted but those with skewed margins are down-weighted according to the relative entropy among the tables of the same odds-ratio.

Keywords: Entropy, Margin, Measures of association, Odds-ratio, Statistical dependence, Two by two probability tables.

INTRODUCTION

2×2 tables of binary markers with random margins are intriguing in several respects: First, there is a confusing plethora of measures of association in 2×2 tables with random margins that are used in statistical practice. Their relative merit is unclear. Some of them were developed for 2×2 tables with fixed margins and then extended to the case considered here. Measures typically agree in the ordering by strength of association on 2×2 tables that have diagonal symmetry and in case of independence. But they differ markedly in asymmetric tables or in tables which are ”far from independence”. We develop a unified framework to analyse, visualise and compare measures of association in 2×2 probability tables especially with respect to their dependence on the margins.

Second, 2×2 tables ”far from independence” may approximate logical forms like logical equivalence (one diagonal is zero) or implication (one entry zero). The task of selecting particularly interesting and informative tables among a large number of tables is often encountered in the analysis of data consisting of high dimensional binary patterns (*e.g.* linkage disequilibrium of SNPs, patterns of aberration at various DNA loci, patterns of protein expression *etc*.). We suggest a principled approach for picking tables which approximate logical relations. This approach relies on an entropy-based weighting of tables and aims to improve existing measures often used in Genetical Statistics.

Defining and justifying measures and estimating them from empirical data are radically different tasks. We have investigated methods of estimating measures of association in a separate paper [1]. Here we deal exclusively with abstract 2×2 probability models and their mathematical structure.

* Address correspondence to this author at the Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Haertelstrasse 16-18 04107, Leipzig, Germany; Tel: +49 341 97 16121; Fax: +49 341 97 16109; E-mail: dirk.hasenclever@imise.uni-leipzig.de

RESULTS

Mathematical Structure of 2×2 Probability Models

2×2 tables of binary markers with random margins can be considered as tetranomial distributions with a symmetry structure. Symmetry of 2×2 tables can be described by the dihedral group D_4 generated by the transposition of the binary markers (matrix transposition) and transposition of their values (transposition of columns or rows).

We consider the manifold $\mathbb{T}$ of all non-degenerate tetranominal probability models which we write in two by two lay-out: $\mathbb{T}$ consists of all two by two matrices t with entries $p_{ij} \in \mathbb{R}$, $(i, j \in \{0, 1\})$ subject to the constraints $p_{ij} > 0$, $\Sigma_{i,j} p_{ij} = 1$. The p_{ij} denote the probabilities of the corresponding combination of the states of two binary markers i and j. In the following, we abbreviate $\Sigma^1_{i=0}\Sigma^1_{j=0} = \Sigma_{i,j}$, $p_{i\cdot} = p_{io} + p_{i1}$ and $p_{\cdot j} = p_{oj} + p_{1j}$. The margins $p_{i\cdot}$ and $p_{\cdot j}$ give the marginal distributions of the marker i and j respectively.

In $\mathbb{T}$ we have several relevant submanifolds. There is a marked point m, namely the midpoint $\begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{pmatrix}$. There is the 1-dimensional submanifold $\mathbb{DS}$ of all tables with diagonal symmetry of the form $\begin{pmatrix} a & b \\ b & a \end{pmatrix}$. And there is the 2-dimensional submanifold $\mathbb{IND}$ of independent tables with $p_{ij} = p_{i\cdot} \bullet p_{\cdot j} \ \forall\, i, j$.

By $\bar{\mathbb{T}}$ we denote the closure of $\mathbb{T}$. The border $\partial\bar{\mathbb{T}} = \bar{\mathbb{T}} - \mathbb{T}$ consists of tables with at least one zero: four two dimensional sides $\{p_{ij} = 0\}$ for any i, j, six one dimensional edges of vanishing rows $\{p_{\cdot j} = 0\}$, vanishing columns $\{p_{i\cdot} = 0\}$ and two vanishing diagonals $\{p_{00} = p_{11} = 0\}$, $\{p_{01} = p_{10} = 0\}$ as well as four triple zero vertices $\{p_{ij} = 1\}$.

Manipulating the margins defines an additional structure on $\mathbb{T}$. We can multiply rows or columns with positive numbers and renormalise: Formally, consider the group $G = (\mathbb{R}^+ \times \mathbb{R}^+, \bullet)$ with component-wise multiplication.

For every $(\mu, \nu) \in \mathbb{R}^+ \times \mathbb{R}^+$ we define a map: $g(\mu, \nu) : \mathbb{T} \to \mathbb{T}$

$$t = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} \longmapsto g(\mu,\nu)(t) \quad = \quad \frac{1}{\mu\nu p_{00} + \mu p_{01} + \nu p_{10} + p_{11}} \begin{pmatrix} \mu\nu p_{00} & \mu p_{01} \\ \nu p_{10} & p_{11} \end{pmatrix} \tag{1}$$

Since $g(\mu, \nu) \circ g(\mu', \nu') = g(\mu \cdot \mu', \nu \cdot \nu')$ and $g(1, 1) = \mathrm{Id}_{\mathbb{T}}$ this defines a G-group action on $\mathbb{T}$.

Lying in the same group orbit defines an equivalence relation on $\mathbb{T}$: We say two elements $t_1, t_2 \in \mathbb{T}$ are equivalent $t_1 \sim t_2$ if and only if there are $(\mu, \nu) \in \mathbb{R}^+ \times \mathbb{R}^+$ with $g(\mu, \nu)(t_1) = t_2$. G-Orbits are diffeomorph to $\mathbb{R}^+ \times \mathbb{R}^+$.

A real function $\eta : \mathbb{T} \to \mathbb{R}$ is G-invariant if $\eta(t) = \eta(g(\mu, \nu)(t))$ for all $(\mu, \nu) \in \mathbb{R}^+ \times \mathbb{R}^+$.

Proposition 1 (odds-ratio):

a) The odds-ratio $\lambda : \mathbb{T} \to \mathbb{R}$; $t = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} \mapsto \lambda(t) = \frac{p_{00}p_{11}}{p_{01}p_{10}}$ is G-invariant.

b) The odds-ratio classifies the G-orbits. Let $\tilde{\mathbb{T}}$ be the quotient space of $\mathbb{T}$ by the equivalence relation induced by G. λ induces a bijective map $\tilde{\lambda} : \tilde{\mathbb{T}} \to \mathbb{R}^+$.

c) The inverse mapping $\tilde{\lambda}^{-1} : \mathbb{R}^+ \to \tilde{\mathbb{T}}$ can be described by $l \longmapsto \left[\begin{pmatrix} \frac{\sqrt{l}}{2\cdot(1+\sqrt{l})} & \frac{1}{2\cdot(1+\sqrt{l})} \\ \frac{1}{2\cdot(1+\sqrt{l})} & \frac{\sqrt{l}}{2\cdot(1+\sqrt{l})} \end{pmatrix} \right]$

d) Every G-invariant function $\eta : \mathbb{T} \to \mathbb{R}$ can be written as a function of λ, namely $\eta = (\tilde{\eta} \circ \tilde{\lambda}^{-1}) \circ \lambda$.

Proof: a) is easily verified. b) Every equivalence class [t] in $\tilde{\mathbb{T}}$ has a representant with margins ½, namely $[g(\sqrt{\frac{p_{10}p_{11}}{p_{00}p_{01}}}, \sqrt{\frac{p_{01}p_{11}}{p_{00}p_{10}}})(t)]$ which has the form given in c). d) is trivial.

We next define new coordinates on $\mathbb{T}$ to make use of this insight.

Proposition 2 (Margin transformation coordinates on T highlighting the G-action and its invariant):

The map $\Theta : \mathbb{T} \to \mathbb{R}^3$

$$t = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} \longmapsto \Theta(t) = (\ln \sqrt{\frac{p_{00}p_{11}}{p_{01}p_{10}}}, \ln \sqrt{\frac{p_{00}p_{01}}{p_{10}p_{11}}}, \ln \sqrt{\frac{p_{00}p_{10}}{p_{01}p_{11}}}) \tag{2}$$

is a diffeomorphism.

The inverse $\Psi = \Theta^{-1} : \mathbb{R}^3 \to \mathbb{T}$ is given by

$$\begin{aligned} \Psi(x,y,z) &= g\,(e^y, e^z)) \left(\begin{pmatrix} \frac{e^x}{2(1+e^x)} & \frac{1}{2(1+e^x)} \\ \frac{1}{2(1+e^x)} & \frac{e^x}{2(1+e^x)} \end{pmatrix} \right) \\ &= \frac{1}{e^{x+y+z} + e^x + e^y + e^z} \begin{pmatrix} e^{x+y+z} & e^y \\ e^z & e^x \end{pmatrix} \end{aligned}$$

In these new coordinates, x corresponds to the logarithmized odds-ratio [2], while y and z determine the G-transformation that maps the table to diagonal symmetry. In addition, the midpoint m_0 corresponds to the origin (0, 0, 0). G-orbits (odds-ratio = constant) correspond to planes $\{a\} \times \mathbb{R}^2$. In particular, the submanifold of independent tables $\mathbb{IND}$ maps to $\{0\} \times \mathbb{R}^2$. The tables with diagonal symmetry $\mathbb{DS}$ form the line $\mathrm{R}\times\{0\}\times\{0\}$. Transposing rows and columns of a table is equivalent to transformations $y \to -y$ and $z \to -z$, while matrix transposition is equivalent to the transformation $y \leftrightarrow z$.

Let $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ be the two point compactification of $\mathbb{R}$. $\bar{\mathbb{R}}^3$ is a compactification of $\mathbb{R}^3$ as a cube. We use a short hand notation to describe the boundaries abbreviating $+\infty$ as "+", $-\infty$ as "–" and any finite real number as "*". The eight vertices $\mathrm{V} = \{(\pm\pm\pm)\}$ split into two sets of four: $V_g = \{(+++), (+--), (-+-), (--+)\}$ and $V_b = \{(---), (-++), (+-+), (++-)\}$.

Proposition 3 (Extension to the borders): Ψ and Θ considered as set valued functions can be extended to $\bar{\mathbb{R}}^3$ respectively $\bar{\mathbb{T}}$. They remain inverse to each other. The mappings of the borders can be characterized as follows:

- The vertices V_g together with their respective adjacent edges map to the vertices in $\bar{\mathbb{T}}$.
- The faces of $\bar{\mathbb{T}}$ correspond to the vertices V_b.
- The faces $(\pm * *)$ of the cube map to the diagonal edges $p_{00} = p_{11} = 0$ and $p_{01} = p_{10} = 0$ in $\bar{\mathbb{T}}$.
- The faces $(* \pm *)$, $(* * \pm)$ correspond to tables with vanishing rows $\{p_{\cdot j} = 0\}$ or vanishing columns $\{p_{i\cdot} = 0\}$ in $\bar{\mathbb{T}}$ respectively.

This behaviour is illustrated in Fig. (**1**). These different compactifications will later be used to characterise the limit behaviour of association measures. It will turn out that the limit behaviour can be easier described using the margin transformation coordinates.

Measures of Association

We will now investigate various measures of associations between two binary markers. First we define the objects of interest.

Definition (Measures of association): A measure of association between binary markers is a continuous function $\eta : \mathbb{T} \to \mathbb{R}$ with the following properties:

a) η is zero on independent tables.

b) η is a strictly increasing function of the odds-ratio when restricted to tables with fixed margins.

c) η respects the symmetry group D_4, namely:

c1) η is symmetric in the markers, *i.e.* invariant to matrix transposition.

c2) η changes sign when states of a marker are transposed (row or column transposition).

d) The range of the function is restricted to (−1, 1).

The first two conditions are equivalent to basic properties proposed by Piatetsky-Shapiro [3]. Condition c) is added to acknowledge that associations between interchangeable markers are of interest here. Finally, condition d) is added to define a unique scale for all measures of association which is often referred as *standardization*.

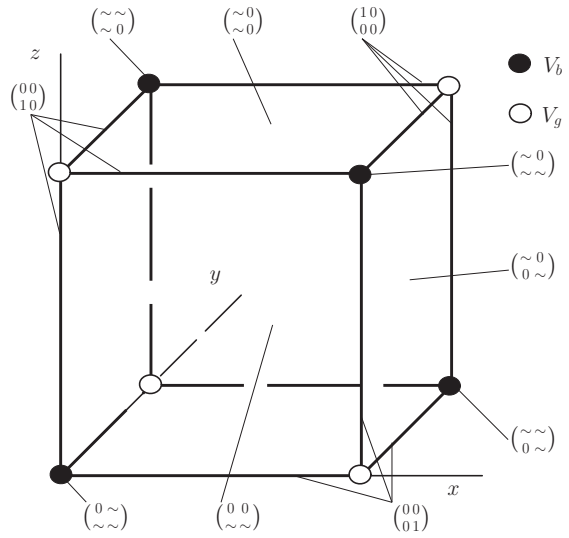


Fig. (1). Illustration of the maps θ and Ψ on the boundaries of $\bar{\mathbb{R}}^3$ and $\bar{\mathbb{T}}$: "∼" represents positive numbers adding up to 1.

Measures Based on the Odds-Ratio

The odds-ratio *Odds-ratio* λ:

$$\lambda = \frac{p_{00}p_{11}}{p_{01}p_{10}}$$

can be used to define measures of association. As λ is G-invariant, monotone transformations automatically fulfill condition b) of the definition.

Measures of association derived from the odds-ratio include *Yule's Q* [4]:

$$Q = \frac{\lambda - 1}{\lambda + 1}$$

and *Yule's Y* [4]:

$$Y = \frac{\sqrt{\lambda} - 1}{\sqrt{\lambda} + 1}$$

Obviously, both *Q* and *Y* are measures of association in our sense. Similar to the odds-ratio, both are extremal if one of the p_{ij} tends to zero.

Measures Based on Additive Deviations from Independence Given the Margins

Fixing margins results is a one dimensional submanifold of tables that can be additively parametrised by a parameter *D*.

All such tables have the form:

$$\begin{pmatrix} p_{0.} \cdot p_{.0} + D & p_{0.} \cdot p_{.1} - D \\ p_{1.} \cdot p_{.0} - D & p_{1.} \cdot p_{.1} + D \end{pmatrix}$$

$D = p_{00}p_{11} - p_{01}p_{10} = p_{00} - p_{0.}p_{.0}$ describes the additive deviation from the independent table with the given margins. This measure is zero in case of independence of the markers but extremal values depend on the margins.

Lewontin's D' [5]: The measure D' is a standardisation of the original measure D:

$$D' = \frac{D}{D_{max}} \quad \text{where} \quad D_{max} = \begin{cases} \min\{p_{0.}p_{.1}, p_{.0}p_{1.}\} & \text{if} \quad D \geq 0 \\ \min\{p_{0.}p_{.0}, p_{1.}p_{.1}\} & \text{if} \quad D < 0 \end{cases}$$

Lewontin's D' ranges from −1 to 1 and tends to these values if at least one of the p_{ij} tends to zero.

D' is widely used in genetics to measure linkage disequilibrium. When a new SNP emerges in a population by a single mutation event, the new allele is exclusively found in conjunction with only one of the two alleles of already existing SNPs. As long as no recombination events occurs, the new SNP remains in complete linkage disequilibrium with the other SNPs. The corresponding 2×2 tables feature a single zero cell. Thus in this context a measure is needed that is extremal whenever a single entry tends to zero.

Since D_{max} is constant for tables with fixed margins and D increases with increasing odds-ratio, D' is a monotone function of the odds-ratio for constant margins. Symmetry is obvious. Hence, D' is a measure of association in our sense.

Correlation coefficient r [6]: The correlation coefficient applied to binary data has similar popularity in genetics as D'. It ranges also from −1 to 1, but, in contrast to D', the absolute value 1 is obtained when a diagonal of t tends to zero:

$$r = \frac{D}{\sqrt{p_{0.}p_{.0}p_{1.}p_{.1}}} = \frac{p_{00}p_{11} - p_{01}p_{10}}{\sqrt{p_{0.}p_{.0}p_{1.}p_{.1}}}$$

With reasoning similar as for D', r is a measure of association.

Proposition 4 (Equality of *r*, *D'* and *Y* on diagonal tables): The measures r, D' and Y coincide on the set of diagonal tables, *i.e.* tables with pair-wise equal diagonal elements.

Proof: This follows directly after calculating these measures for the tables $t = \frac{1}{2a+2b}\begin{pmatrix} a & b \\ b & a \end{pmatrix},\ a, b > 0.$

Measures Based on Information Theory

The *mutual information* [7] is defined as the difference between the information of the given table and the independent table with the same margins.

$$\text{MutInf} = \sum_{i,j} p_{ij} \cdot \log_2(p_{ij}) - \sum_{i} p_{i.} \cdot \log_2(p_{i.}) - \sum_{j} p_{.j} \cdot \log_2(p_{.j})$$

MutInf takes values only in (0, 1). In order to make it a measure of association according to our definition, we define a signed version:

$$\text{sMutInf} = \text{sign}(D) \cdot \text{MutInf}$$

Proposition 5: sMutInf is a measure of association.

Proof: The symmetry of this measure is clear. To show that sMutInf is a monotone function of the odds-ratio, we consider the tables $t_\varepsilon = \begin{pmatrix} p_{00}+\varepsilon & p_{01}-\varepsilon \\ p_{10}-\varepsilon & p_{11}+\varepsilon \end{pmatrix}$ for a sufficiently small ε > 0. These tables have the same margins as the

table $t = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}$ but higher odds-ratios. Assume that $\lambda > 1$, we see that $\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}$ sMutInf $(t_\varepsilon) = \log_2 \lambda > 0$. Hence sMutInf is monotone, and thus, a measure of association.

MutInf approaches 1 only if tables approach $\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$; while *r* approaches 1 if tables approach the form $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, $a, b > 0$.

Counter Example

Kappa coefficient [8]: The Kappa coefficient which is useful in quantifying the agreement between two raters is defined as:

$$\kappa = \frac{p_{00} + p_{11} - p_{0.}p_{.0} - p_{1.}p_{.1}}{1 - p_{0.}p_{.0} - p_{1.}p_{.1}}$$

Kappa is **not** a measure of association in our sense. Although it fulfils the condition of monotonicity, it is not symmetric.

Comparing Measures of Association

We use the coordinates introduced in Proposition 2 in order to describe and visualise how measures of association depend on the margins. In particular we study measures of association η restricted to *x*=const *i.e.* for fixed odds-ratios. The restricted functions will be denoted η_x and called margin weighting functions. We characterise the shape of the margin weighting functions and study their limiting behaviours and extensibility to the compactification $\bar{\mathbb{R}}^3$ in comparison to $\bar{\mathbb{T}}$.

The association measure *r* expressed in margin transformation coordinates reads:

$$r(x,y,z) = \frac{\left(e^{2x}-1\right)e^{y+z}}{\sqrt{\left(e^{x+y+z}+e^{y}\right)\left(e^{x+y+z}+e^{z}\right)\left(e^{x}+e^{y}\right)\left(e^{x}+e^{z}\right)}} \quad \textbf{(3)}$$

The margin weighting function of *r* for odds-ratio $\lambda = 40$ is shown in Fig. (**2**).

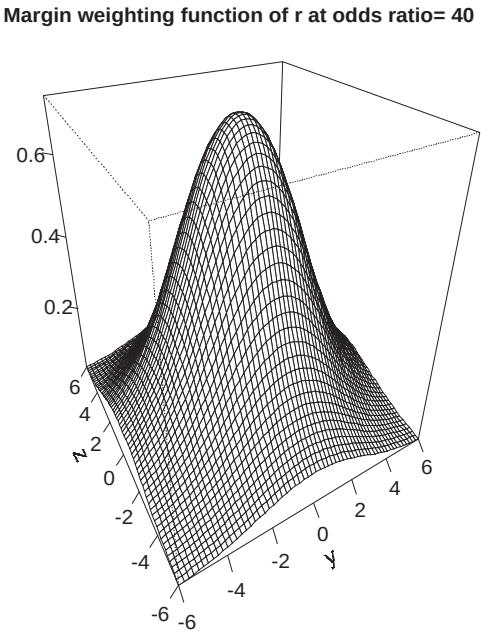


Fig. (2). Margin-weighting function of *r*.

Proposition 6 (Margin weighting function for *r*): For all $x \in \mathbb{R} \setminus \{0\}$:

a) r_x has exactly one extremum at the origin $(y, z) = (0, 0)$, corresponding to the diagonal symmetric table with the fixed odds-ratio.

b) $\lim_{\|(y,z)\|\to\infty} r_x = 0$.

c) $\lim_{x\to\pm\infty} r_x = \pm 1$

d) r can be extended to $\bar{\mathbb{R}}^3$ except for the lines (±,±, *) and (±, *,±) and the vertices V.

e) r can be extended to $\bar{\mathbb{T}}$ except for the vertices.

Proof: see **Supplement Material**.

The measure r down-weights tables with skewed margins.

The association measure D' expressed in margin transformation coordinates reads:

$$
\begin{aligned}
D'(x,y,z) &= \frac{\left(e^{2x}-1\right)e^{y+z}}{D_{max}} \qquad \text{where} \\
D_{max} &= \begin{cases}
\left(e^{x+y+z}+e^{y}\right)\left(e^{x}+e^{y}\right) & : \quad x>0, \quad y<z \\
\left(e^{x+y+z}+e^{z}\right)\left(e^{x}+e^{z}\right) & : \quad x>0, \quad y\geq z \\
\left(e^{x+y+z}+e^{y}\right)\left(e^{x+y+z}+e^{z}\right) & : \quad x<0, \quad y<-z \\
\left(e^{x}+e^{y}\right)\left(e^{x}+e^{z}\right) & : \quad x<0, \quad y\geq -z
\end{cases}
\end{aligned}
\tag{4}
$$

The margin weighting function of D' for odds-ratio $\lambda = 40$ is shown in Fig. (**3**).

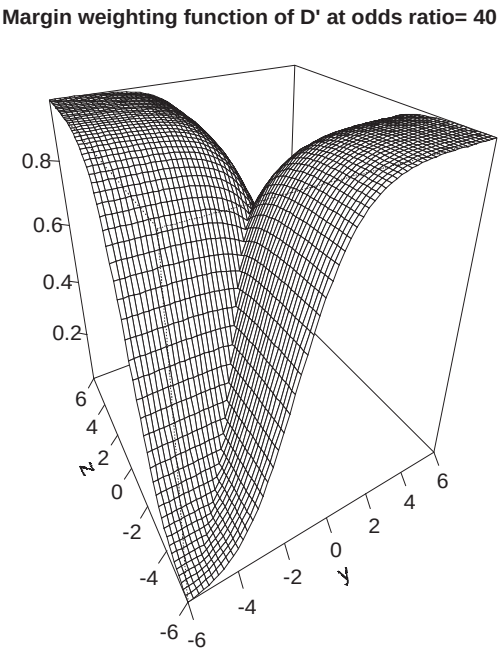


Fig. (3). Margin-weighting function of D'.

Proposition 7 (Margin weighting function for *D'*): For all $x \in \mathbb{R}\ \{0\}$:

a) D_x' has a non-differentiable edge along the diagonal $y = z$ for $D' > 0$ and along the diagonal $y = -z$ for $D' < 0$. There is a non-smooth saddle point in the origin.

b)

$$\lim_{y\to\pm\infty} D'_x = \left(e^{2x}-1\right)\cdot\begin{cases}\left(e^{2x}+e^{x\pm z}\right)^{-1} & : \quad x>0\\ \left(e^{x\mp z}+1\right)^{-1} & : \quad x<0\end{cases}$$

$$\lim_{z\to\pm\infty} D'_x = \left(e^{2x}-1\right)\cdot\begin{cases}\left(e^{2x}+e^{x\pm y}\right)^{-1} & : \quad x>0\\ \left(e^{x\mp y}+1\right)^{-1} & : \quad x<0\end{cases}$$

Thus, limit functions have a range of $(0, 1 - e^{-2x})$ for $x > 0$ and $(e^{2x} - 1, 0)$ for $x < 0$, where 0 is obtained for $y \to \pm\infty$, $z \to \pm\infty$, $x > 0$ and $y \to \mp\infty$, $z \to \pm\infty$, $x < 0$.

c) $\lim_{x\to\pm\infty} D_x' = \pm 1$

d) D' can be extended to $\bar{\mathbb{R}}^3$ except for the vertices V_g.

e) D' can be extended to $\bar{\mathbb{T}}$ except for the edges and vertices.

Proof: see **Supplement Material**.

D' gives higher weights to certain tables without diagonal symmetry. The measure up-weights or down-weights tables with skewed margins depending on the position of zeros which occur in the limiting tables (see Fig. **3**). Comparing d) and e) one recognizes that the introduction of the odds-ratio as coordinate allows extending D' to limit tables with vanishing colums or rows.

The association measure sMutInf can also be written in margin transformation coordinates but this is skipped due to the lengthy formula. The margin weighting function of sMutInf for odds-ratio $\lambda = 40$ is shown in Fig. (**4**).

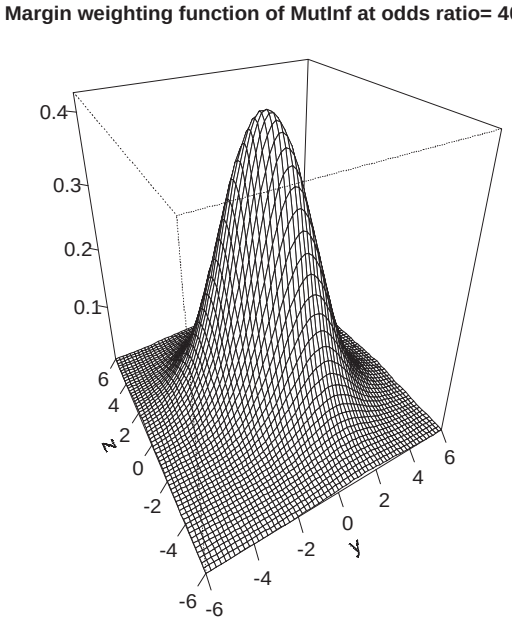


Fig. (4). Margin-weighting function of sMutInf.

Proposition 8 (Margin weighting function for sMutInf): For all $x \in \mathbb{R} \setminus \{0\}$:

a) sMutInf_x has exactly one maximum at the origin $(y, z) = (0, 0)$.

b) $\lim_{|(y,z)|\to\infty} \text{sMutInf}_x = 0$.

c) $\lim_{x\to\pm\infty} \text{sMutInf}_x = \pm\left(\log_2\left(e^{y\pm z}+1\right) - \frac{e^{y\pm z}}{e^{y\pm z}+1}\log_2 e^{y\pm z}\right)$

Thus $\text{sMutInf}_x \to \pm 1$ for $y = \mp z$ and $x \to \pm\infty$ respectively.

d) sMutInf can be extended to $\bar{\mathbb{R}}^3$ except for the vertices V_b.

e) sMutInf can be extended completely to $\bar{\mathbb{T}}$.

Proof: see **Supplement Material**.

Thus, similarly to *r*, sMutInf down-weights tables with skewed margins (see Fig. **4**).

The association measure *Y* in margin transformation coordinates can be simply written as:

$$Y(x,y,z) \quad = \quad \tanh\frac{x}{2} \tag{5}$$

Proposition 9 (Margin weighting function for *Y*): For all $x \in \mathbb{R}$:

a) Y_x is constant.

b) $\lim_{\|(y,z)\|\to\infty} Y_x = \tanh\frac{x}{2}$

c) $\lim_{x\to\pm\infty} Y_x = \pm 1$

d) *Y* can be extended completely to $\bar{\mathbb{R}}^3$.

e) *Y* can be extended to $\bar{\mathbb{T}}$ except for edges and vertices corresponding to vanishing rows or columns.

Proof: is trivial.

Entropy

Among tables of a fixed odds-ratio, we look for a principled approach to prefer interesting tables and down-weight obscure "junk" tables. As a candidate we study the table entropy on $\mathbb{T}$. The entropy function $H : \mathbb{T} \to \mathbb{R}$ is defined as the negative expectation of the loglikelihood of the tables:

$$H\left(\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}\right) \quad := \quad -\left(p_{00}\cdot\log_2(p_{00}) + p_{01}\cdot\log_2(p_{01}) + p_{10}\cdot\log_2(p_{10}) + p_{11}\cdot\log_2(p_{11})\right)$$

Why is entropy a candidate to select among tables? It can be characterised in multiple ways: For general finite discrete distributions the entropy was introduced by Shannon (1948) [9]. Shannon characterised *H* by a set of postulates to measure the uncertainty in a discrete distribution:

Shannon's Characterisation of Entropy: If functions $H_n(p_1, ..., p_n)$ with $p_i \geq 0$, $\sum p_i = 1$, $n \geq 2$ satisfy the conditions

a) $H_2(p, 1-p)$ is a continuous positive function of *p*.

b) $H_n(p_1, ..., p_n)$ is symmetric, *i.e.* invariant under permutations of the $p_1, ..., p_n$ for all *n*.

c) $H_n(p_1, ..., p_n) = H_{n-1}(p_1 + p_2, p_3, ..., p_n) + (p_1 + p_2)\cdot H_2\left(\frac{p_1}{(p_1+p_2)}, \frac{p_2}{(p_1+p_2)}\right)$

then $H_n(p_1, ..., p_n) = -K\cdot\Sigma\, p_i \log_2(p_i)$ for some $K > 0$.

Tables with high entropy are interesting as they have high uncertainty and "surprise value".

Jaynes [10] gives an independent combinatorial characterisation: When we sample sequentially from a table $t \in \mathbb{T}$ we obtain a vector of observations of length N, which we summarise as a frequency table $\hat{t}_N = 1/N \cdot \begin{pmatrix} n_{00} & n_{01} \\ n_{10} & n_{11} \end{pmatrix}$. Each frequency table $\hat{t}_N$ is characterised by the number $W(\hat{t}_N) = \frac{N!}{n_{00}!n_{01}!n_{10}!n_{11}!}$ of sequences which realise $\hat{t}_N$. Intuitively, tables that can be realised in multiple ways

are more plausible than those that can be realised only by few sequences. We can use Stirlings formula for $n!$ to approximate $W(\hat{t}_N)$. In the limit $N \to \infty$, $\hat{t}_N \to t$ in probability and $1/N \cdot \log(W(t^N)) \to H(t)$. Thus the entropy describes the combinatorial plausibility of a table.

Given a set of distributions fulfilling certain constraints, Jaynes [10] proposes to pick the corresponding maximum entropy distribution as the most uncommitted and prototypical distribution. Looking at the margin weighting function of the entropy leads to a surprise (see Fig. **5**):

Recall that Lambert's W-function is defined as the inverse function to $x \exp x$. W is a multi-branch function since $y = x \exp(x)$ has two solutions for $y \in (-1/e, 0)$. We can prove the following:

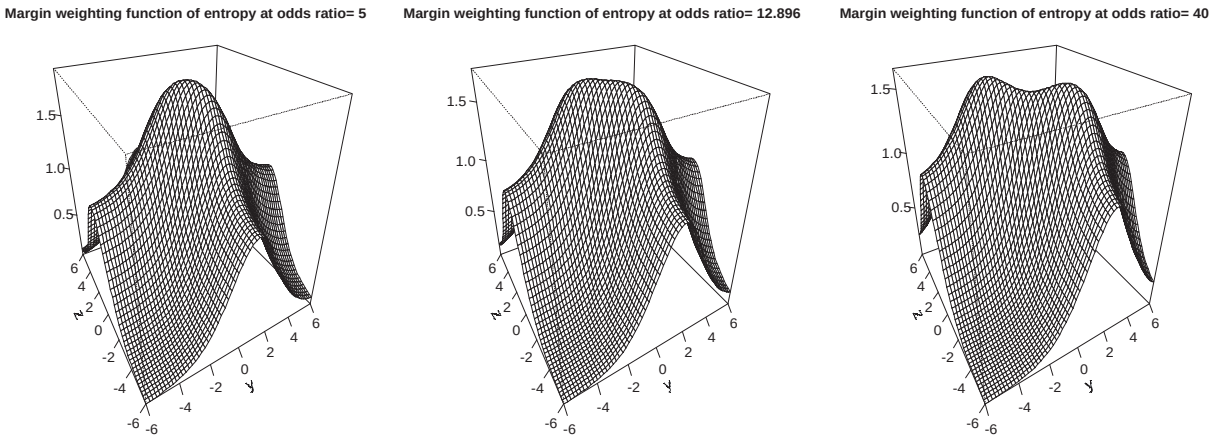


Fig. (5). Margin-weighting function of entropy: The margin-weighting function of the entropy is shown, conditioned to the odds-ratios $\lambda = 5$ which results in a single maximum, $\lambda = 12.896$ at which the maximum splits into two and $\lambda = 40$ with two maxima and a saddle point.

Theorem 1 (magic odds-ratio): Define the "magic odds-ratio" by $L_{magic} = W(1/e)^{-2} \approx 12.89$. Let $L > 1$. The entropy H restricted to the submanifold of constant odds-ratio L in $\mathbb{T}$

- has a single maximum at the diagonal table of odds-ratio L if $1 < L \leq L_{magic}$.
- has a saddle point at the diagonal table of odds-ratio L and two "L-shaped" tables as maxima which transpose with matrix transposition if $L_{magic} < L$.

"L-shaped" means that for $L \to \infty$ one of the maxima approaches the table $\begin{pmatrix} 1/3 & 1/3 \\ 0 & 1/3 \end{pmatrix}$. For the case $L < 1$ a similar result can be derived by transposing principal and secondary diagonals.

Proof: There are two constraints to be considered, one of them not linear in p_{ij} :

$$\ln(p_{00}) - \ln(p_{01}) - \ln(p_{10}) + \ln(p_{11}) \quad = \quad \ln(L) \tag{6}$$

$$p_{00} + p_{01} + p_{10} + p_{11} \quad = \quad 1 \tag{7}$$

Using Langrange multipliers, critical tables of H restricted to odds-ratio equals L can be expressed in terms of Lambert's W function. The bifurcation occurs for $L_{magic} < L$ because Lambert's W is multibranched. See supplement material for details.

This theorem suggests that the "magic odds-ratio" is a natural cutpoint between weak and strong association. For weak association $L < L_{magic}$, interesting tables are those near $\mathbb{DS}$. For strong association $L_{magic} < L$, particularly interesting tables are those that approach "L-shape", *i.e.* those in which one cell differs in magnitude from the three others.

An Entropy-Based Measure of Association

Using these insights on the entropy of a table, in this section we aim to define a measure of association with similar properties to D', Y but better limit behaviour, *i.e.* the measure should down-weight tables with almost vanishing rows or columns or single entries. These tables are denoted as *junk tables* in the following. We have seen in the last sections that D' and Y could be large for these tables.

We also like to recall that both, D' and Y become extremal if the table features a single entry equals zero while r, sMutInf require a vanishing diagonal. We like to retain this property for a new measure to be defined. Another feature to be retained is the agreement of measures for diagonal tables which holds for Y, D' and r.

According to our definition, an important property of a measure of association is that it is a monotone function of the odds-ratio when the margins are kept fixed. For the entropy, one can prove the following lemma:

Lemma 1 (Monotony of the entropy difference): Let H be the entropy of t and H_{diag} be the entropy of the corresponding diagonal table of the same odds-ratio λ. Then, $H_{diag} - H$ is monotonically decreasing for increasing $\lambda > 1$ and constant margins.

Proof: see supplement material.

As a direct consequence of this lemma, it is easy to see that:

Corollary:

$$\mathrm{HS}_n \quad := \quad \operatorname{sign} Y \, |Y|^{\exp n(H_{diag}-H)} \tag{8}$$

is a measure of association for arbitrary $n \geq 0$.

This newly defined measure fulfils all above mentioned properties: It coincides with Y, D', r at diagonal tables, is extremal for tables with a single zero, up-weights L-shaped tables for large odds-ratios in the sense that $\mathrm{HS}_n > Y$ and down-weights junk-tables in the sense that $\mathrm{HS}_n < Y$ at the margins (proof see below). However, the down-weighting is imperfect as $\mathrm{HS}_n > 0$ for junk-tables.

The parameter n can be chosen in order to define the degree of up- and down-weighting. According to our observations, $n = 4$ is a reasonable choice resulting in a satisfactory down-weighting of junk tables (see later).

The measure HS_n can be written in margin transformation coordinates using

$$\begin{aligned} H_{diag}(x,y,z) &= 1 + \log_2(1+e^x) - \frac{x}{\ln 2\,(e^{-x}+1)} \\ H(x,y,z) &= \log_2\left(e^{x+y+z}+e^x+e^y+e^z\right) - \frac{(x+y+z)\,e^{x+y+z} + xe^x + ye^y + ze^z}{\ln 2\,(e^{x+y+z}+e^x+e^y+e^z)} \end{aligned}$$

At Fig. (**6**) we present the margin weighting functions of HS_n for $\lambda = 5$ and $\lambda = 40$. These functions can be easily characterised using the results of the previous section:

Proposition 10 (Margin weighting function for HS_n): For all $x \in \mathbb{R} \setminus \{0\}$:

a) For $x \in (-1 - W(1/e),\ 1 + W(1/e))$, HS_{n_x} has exactly one maximum at the origin $(y,z) = (0,0)$. If $x < -1 - W(1/e)$ or $x > 1 + W(1/e)$, HS_{n_x} has a saddle-point at the origin and two extrema elsewhere. At these extrema, the elements of one diagonal are equal while at the other diagonal there is one (small) element.

b) HS_{n_x} has the following limit functions

$$\lim_{\|(y,z)\|\to\infty} \mathrm{HS}_{n\,x} \quad = \quad \operatorname{sign}\left(\tanh\frac{x}{2}\right)\left|\tanh\frac{x}{2}\right|^{\exp n\left\{1+\log_2(1+e^x)-\frac{x}{\ln 2\left(e^{-x}+1\right)}+p\log_2 p+(1-p)\log_2(1-p)\right\}}$$

where $p = (1 + e^{x\pm z})^{-1}$ for $y \to \pm\infty$ or $p = (1 + e^{x\pm y})^{-1}$ for $z \to \pm\infty$ respectively. Thus, the limit functions have an extremum at $p = 0.5$ that is $z = \mp x$ for $y \to \pm\infty$ and $y = \mp x$ for $z \to \pm\infty$ respectively.

c) $\lim_{x\to\pm\infty} \mathrm{HS}_{n_x} = \pm 1$

d) $\mathrm{HS}_{n_x} < Y_x$ at the margins, *i.e.* HS_n down-weights junk-tables.

e) HS_n can be extended completely to $\bar{\mathbb{R}}^3$.

f) HS_n can be extended to $\bar{\mathbb{T}}$ except for edges and vertices corresponding to vanishing rows or columns.

g) For all $x \in \mathbb{R}$, HS_n coincides with Y, D', r at diagonal tables.

Proof: a) follows from Theorem 1. b) is easy to see taking the limit of the tables first. c) is clear since $\lim_{x\to\pm\infty} \tanh {}^{x}/_{2} = \pm 1$ and the exponent is finite. d) holds since $H_{diag} > 1$ and $H \leq 1$ at the margins of finite x. e) and f) are consequences of b) and c). g) is obvious.

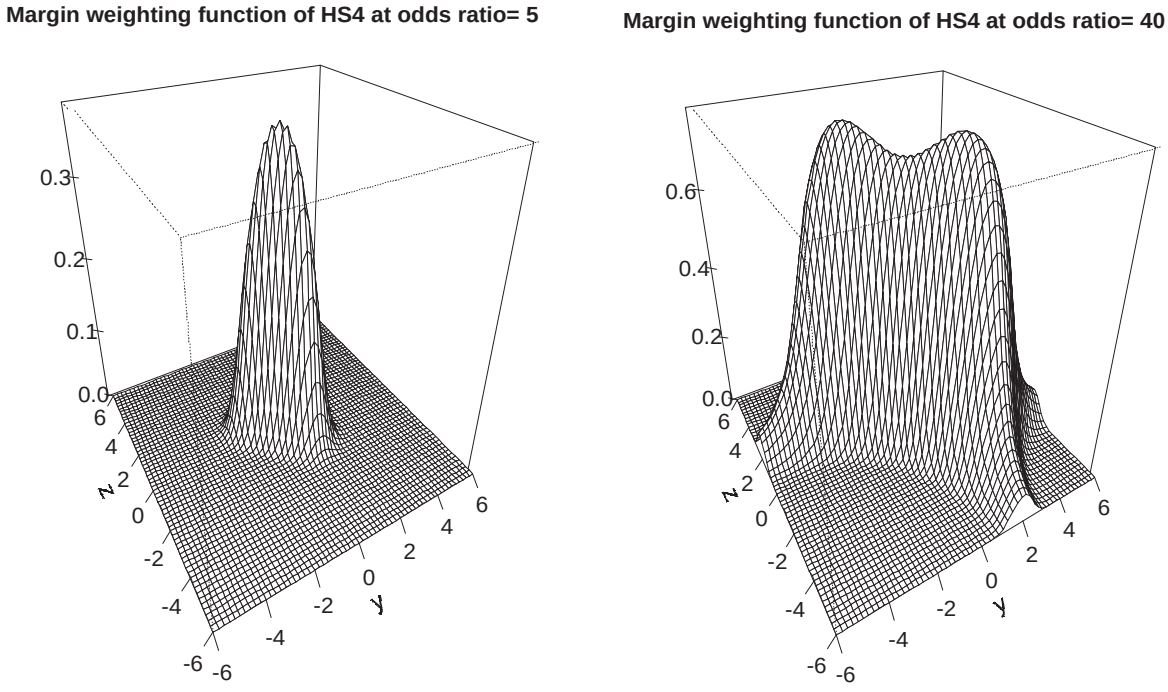


Fig. (6). Margin-weighting function of HS_4.

Examples of Tables and Corresponding Association Measures

We now study the behaviour of the measures Y, r, D' and the newly proposed measure HS_4 for a variety of selected tables (see Table **1**). For this purpose, we study the odds-ratios $\lambda \in \{1, 2, 5, 10, 20, 50, 100\}$ and consider the following tables for $x = \ln \sqrt{\lambda}$:

- The diagonal table ($y = z = 0$).
- An L-shaped table, characterized by $y = x$, $z = -x$.
- A junk table with $y = 10$, $z = -y$ corresponding to $p_{01} \approx 1$.
- A junk table with $y = 10$, $z = -x$ corresponding to $p_{00} \approx p_{01} \approx 0.5$.
- A junk table with $y = 10$, $z = y$ corresponding to $p_{00} \approx 1$.

We also like to remark that the table with three equal entries has maximum entropy if $\lambda \to \infty$.

Per definition of a measure, for $\lambda = 1$ all measures equals zero independent of the concrete realization of the table. Since Y is based on the odds-ratio, Y is constant for all tables of the same odds-ratio. Y, r, D' and HS_4 always coincide at diagonal tables. r is maximal at diagonal tables and becomes small for all kinds of junk tables. D' is always greater for L-shaped tables than for diagonal tables. D' is close to zero in case of $p_{00} \approx 1$ but could become large for $p_{01} \approx 1$ which is highly counter-intuitive. HS_4 also becomes larger for L-shaped tables compared to diagonal tables if λ is large. In contrast to D', HS_4 is close to zero for both junk configurations $p_{00} \approx 1$ and $p_{01} \approx 1$ respectively. The limit tables have a

maximum of the entropy at $p_{00} = p_{01} = 0.5$. This induces a maximum of HS_4 for limit tables which increases with λ (see Table **1**, fourth rows of each odds-ratio).

Table 1. (Measures of association for selected tables): All values are rounded to three decimals. To allow discrimination between hard zeros and small values, all values within (0, 0.0005) are presented as "< 0.001". Entries of 2x2 tables are presented in columns 1 to 4. The fifth columns presents the odds-ratio of the tables. For each odds-ratio, we studied five tables: the diagonal table (first row of the corresponding odds-ratio), a table with three equal entries (second row), a table for which it holds that $p_{01} \approx 1$ or $p_{00} \approx 1$ (third and fifth row respectively) and a table for which $p_{00} \approx p_{01} \approx 0.5$ (fourth row). The last four columns contain the corresponding measures of association rounded to three decimals.

p_{00}	p_{01}	p_{10}	p_{11}	λ	Y	r	D'	HS_4
0.25	0.25	0.25	0.25	1	0	0	0	0
0.25	0.25	0.25	0.25	1	0	0	0	0
< 0.001	1	< 0.001	< 0.001	1	0	0	0	0
0.5	0.5	< 0.001	< 0.001	1	0	0	0	0
1	< 0.001	< 0.001	< 0.001	1	0	0	0	0
0.293	0.207	0.207	0.293	2	0.172	0.172	0.172	0.172
0.286	0.286	0.143	0.286	2	0.172	0.167	0.222	0.139
< 0.001	1	< 0.001	< 0.001	2	0.172	< 0.001	0.5	< 0.001
0.5	0.5	< 0.001	< 0.001	2	0.172	0.002	0.333	< 0.001
1	< 0.001	< 0.001	< 0.001	2	0.172	< 0.001	< 0.001	< 0.001
0.345	0.155	0.155	0.345	5	0.382	0.382	0.382	0.382
0.312	0.312	0.062	0.312	5	0.382	0.333	0.556	0.282
< 0.001	1	< 0.001	< 0.001	5	0.382	< 0.001	0.8	< 0.001
0.5	0.5	< 0.001	< 0.001	5	0.382	0.005	0.667	< 0.001
1	< 0.001	< 0.001	< 0.001	5	0.382	< 0.001	< 0.001	< 0.001
0.38	0.12	0.12	0.38	10	0.519	0.519	0.519	0.519
0.323	0.323	0.032	0.323	10	0.519	0.409	0.744	0.441
< 0.001	1	< 0.001	< 0.001	10	0.519	< 0.001	0.9	< 0.001
0.5	0.5	< 0.001	< 0.001	10	0.519	0.007	0.818	< 0.001
1	< 0.001	< 0.001	< 0.001	10	0.519	< 0.001	< 0.001	< 0.001
0.409	0.091	0.091	0.409	20	0.635	0.635	0.635	0.635
0.328	0.328	0.016	0.328	20	0.635	0.452	0.862	0.627
< 0.001	1	< 0.001	< 0.001	20	0.635	< 0.001	0.95	< 0.001
0.5	0.5	< 0.001	< 0.001	20	0.635	0.009	0.905	0.001
1	< 0.001	< 0.001	< 0.001	20	0.635	< 0.001	< 0.001	< 0.001
0.438	0.062	0.062	0.438	50	0.752	0.752	0.752	0.752
0.331	0.331	0.007	0.331	50	0.752	0.48	0.942	0.821
< 0.001	0.999	< 0.001	< 0.001	50	0.752	< 0.001	0.98	< 0.001
0.5	0.5	< 0.001	< 0.001	50	0.752	0.012	0.961	0.086
1	< 0.001	< 0.001	< 0.001	50	0.752	< 0.001	< 0.001	< 0.001
0.455	0.045	0.045	0.455	100	0.818	0.818	0.818	0.818
0.332	0.332	0.003	0.332	100	0.818	0.49	0.97	0.904
< 0.001	0.999	< 0.001	< 0.001	100	0.818	< 0.001	0.99	< 0.001
0.5	0.5	< 0.001	< 0.001	100	0.818	0.015	0.98	0.316
1	< 0.001	< 0.001	< 0.001	100	0.818	< 0.001	< 0.001	< 0.001

DISCUSSION

In this paper we studied measures of association of 2×2 contingency tables. We defined our measures of interest by four conditions: The first two of them (zero in case of independence, monotonicity with odds-ratio in case of fixed margins) are basic properties according to Piatetsky-Shapiro [3]. There is an ongoing debate regarding desired properties of association measures [11]. Here, we additionally postulate a standardization and symmetry under matrix transposition, *i.e.* interchangeability of markers to be associated. In contrast to traditional independence analysis, we asked for the selection of tables which are far away from independence. This objective was motivated by the analysis of high-dimensional molecular genetic data such as SNP array data in which a high number of 2×2 tables occur from which one would like to select cases of high dependence called *linkage disequilibrium*.

In contrast to detecting a (moderate) deviation from independence, quantifying the strength of association is multiform as pointed out for example by Tan *et al*. [12]. A large number of possible measures were proposed in the literature. Those fulfilling our conditions are shortly reviewed. Most of these measures (r, D', Y) agree at diagonal tables. Some of them become extremal for a vanishing diagonal (r, sMutInf) while for others it suffices that a single cell becomes zero (D', odds-ratio based measures). The measures also markedly differ in cases were one of the rows or columns of the table becomes small. Since in practice, it can hardly be decided for these tables whether the dependence is strong or not, these tables are not really of interest and are considered as *junk tables* here. Nevertheless, the measure D' can become large in these cases which is undesirable to our opinion. D' also varies markedly in a small neighbourhood of the vertices of $\mathbb{T}$.

To study the properties of measures of association, we introduced coordinates on the manifold $\mathbb{T}$ of all tables mapping it to 3-dimensional space such that the x-axis corresponds to the logarithmized square root of the odds-ratio. We study the measures on the hyperplanes of constant odds-ratio, looking at the so called *margin weighting functions*. These functions are constant for all measures based on the odds-ratio which is known to be independent of the margins of the table. For other measures, these functions describe the dependence of the measure on the margins for tables with constant odds-ratio. Hence, our construction acknowledges the fact that the odds-ratio completely captures the information of the joint distribution of the two markers except for those contained in the margins [13]. Our margin weighting functions illustrate major properties of the association measures considered. It also helps designing new measures with desired properties, which we demonstrated in the second part of the paper.

The mathematical properties of the margin weighting functions were derived for three measures of association, namely r, sMutInf and D'. It revealed that r and sMutInf behave very similarly by up-weighting diagonal tables but down-weighting tables with small rows or columns. In contrast, D' is not maximal for diagonal tables. Furthermore, it expresses a strange weighting behaviour for tables with small rows and columns, up-weighting or down-weighting these tables in dependence on the position of the structural zeros. Such tables occur frequently *e.g.* in SNP data. This property also explains, why the estimation problem for D' is not well behaved [1]. On the other hand, D' as well as odds-ratio based measures are constructed to up-weight tables which feature a single small entry. These tables represent a prototype of a logical table for which one can conclude the state of the column for one row but not for the other row. These kinds of tables are interesting in genetical statistics since they correspond to situations at which no recombinations occurred between two SNPs, *i.e*. only three of the four theoretically possible haplotypes are observed. Therefore, we aimed to define an alternative measure also highlighting L-shaped tables but with a better behaviour at the margins than D' or odds-ratio based measures.

For this purpose, the entropy [9] as another canonical structure at 2×2 tables was studied. We proved that the margin weighting function of this quantity is maximal at the diagonal for odds-ratios within a critical range, namely ($W(1/e)^2$, $W(1/e)^{-2}$). Outside this range, there are two maxima at L-shaped tables, *i.e*. tables with a single small cell while the others are (almost) equal. More precisely, the elements of the opposite diagonal are equal for the maxima.

The difference between the entropy of a non-diagonal table and the corresponding diagonal table of the same odds-ratio is a monotone function of the odds-ratio for fixed margins. A new measure of association called HS_n is defined, which is essentially *Yule's Y* weighted by the exponential of this entropy difference.This quantity fulfils all requirements of an association measure, *i.e*. ranges between -1 and 1, is zero in case of independence, is symmetric and a monotone function of the odds-ratio for fixed margins. In addition, it agrees with Y , D' and r at diagonal tables, up-weights tables with an L-shape and large odds-ratio and is extremal in case of a single zero in the table. Hence, the measure has similar properties than D' except for a better limit behaviour. Since the entropy difference of tables with vanishing row or column is smaller than the entropy of the corresponding diagonal table, degenerated tables are markedly down-weighted

relative to the diagonal table. The free constant n allows tuning the degree of this down-weighting. For practical issues we recommend using $n = 4$ which yields satisfactory results to our experiences. However, our procedure of down-weighting junk tables is neither unique nor perfect in the sense that the junk tables are down-weighted to zero. The latter one is not possible within the framework of weighting by entropy without loosing other desired properties of the measure, because the minimum of the absolute differences between the diagonal table and the degenerated tables of the same odds-ratio approaches zero if the odds-ratio tends to 0 or ∞.

We recommend using HS_4 instead of D' when interested in selecting L-shaped tables from a large set of tables mostly far away from independence and when tables with small marginal frequencies are common. When HS_4 is estimated from count data, we recommend using Bayesian plug-in estimators of the frequencies of single cells showing a good compromise between accuracy and computational burden [1].

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers Website along with the published article.

CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This research was funded by the Leipzig Interdisciplinary Research Cluster of Genetic Factors, Clinical Phenotypes, and Environment (LIFE Center, University of Leipzig). LIFE is funded by means of the European Union, by the European Regional Development Fund (ERDF), the European Social Fund (ESF), and by means of the Free State of Saxony within the framework of its excellence initiative.

REFERENCES

[1] M Scholz, and D Hasenclever, "Comparison of estimators for measures of linkage disequilibrium", *The International Journal of Biostatistics,* vol. 6, no. 1, 2010. Article 1.
[http://dx.doi.org/10.2202/1557-4679.1162]

[2] A.W. Edwards, "The measure of association in a 2x2 table", *Journal of the Royal Statistical Society [Series A],* vol. 126, pp. 108-114, 1963.
[http://dx.doi.org/10.2307/2982448]

[3] G. Piatetsky-Shapiro, "Discovery, analysis and presentation of strong rules", In: G. Piatetsky-Shapiro, and W. Frawley, Eds., *Knowledge Discovery in Databases.,* MIT Press: Cambridge, MA, 1991, pp. 229-248.

[4] G.U. Yule, "On the association of attributes in statistics", *Philosophical transactions of the Royal Society of London A,* vol. 194, pp. 269-274, 1900.
[http://dx.doi.org/10.1098/rsta.1900.0019]

[5] R.C. Lewontin, "The interaction of selection and linkage. I. general considerations; heterotic models", *Genetics,* vol. 49, no. 1, pp. 49-67, 1964.
[PMID: 17248194]

[6] W.G. Hill, and A. Robertson, "Linkage disequilibrium in finite populations", *Theoretical and Applied Genetics,* vol. 38, no. 6, pp. 226-231, 1968.
[http://dx.doi.org/10.1007/BF01245622] [PMID: 24442307]

[7] W. Weaver, and C.E. Shannon, *The Mathematical Theory of Communication.,* University of Illinois Press: Urbana, IL, 1963.

[8] J. Cohen, "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement,* vol. 20, pp. 37-46, 1960.
[http://dx.doi.org/10.1177/001316446002000104]

[9] CE Shannon, "A mathematical theory of communication", *Bell System Technical Journal,* vol. 27, pp. 379-423, 1948. 623-656

[10] E.T. Jaynes, *Probability Theory: The Logic of Science.,* Cambridge University Press: Cambridge, 2003.
[http://dx.doi.org/10.1017/CBO9780511790423]

[11] MJ Warrens, "On association coefficients for 2x2 tables and properties that do not depend on the marginal distributions", *Psychometrika,* vol. 73, no. 4, pp. 777-789, 2008.

[12] P.N. Tan, V. Kumar, and J. Srivastava, "Selecting the right objective measure for association analysis", *Information Systems,* vol. 29, pp. 293-313, 2004.
[http://dx.doi.org/10.1016/S0306-4379(03)00072-3]

[13] G Osius, "The association between two random elements: A complete characterization and odds ratio models", *Metrika,* vol. 60, pp. 261-277, 2004.
[http://dx.doi.org/10.1007/s001840300309]