Categorical Data Analysis and Visualisation

Part III: Multi-way Contingency Tables

Rosaria Lombardo
Department of Economics,
University of Campania “Luigi Vanvitelli”, Italy
rosaria.lombardo@unicampania.it

Eric J. Beh
University of Wollongong, Wollongong, Australia,
Stellenbosch University, Stellenbosch, South Africa
ericb@uow.edu.au

Multiple Correspondence Analysis

- Indicator table
- Burt table

Multiple Correspondence Analysis

So far we have confined our attention to the case where we are graphically summarising the association between two categorical variables.

In the case where we have three categorical variables, we focus on understanding the association between the rows, columns and tubes.

For more than three categorical variables, the depiction can become very messy for categorical visualisation.

When we have multiple categorical variables, we can perform a correspondence analysis by coding the variables in a number of ways. We can

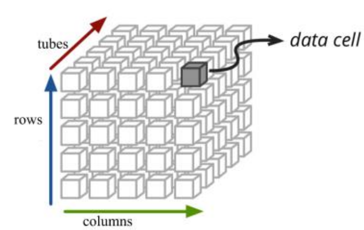
- derive the *indicator table* form of the data
- derive the *Burt table* form of the data
- perform a *joint correspondence analysis, stacking* and *concatenation*

By recoding the contingency table, we perform *multiple correspondence analysis* (MCA)

Notation

For a sample size, n, we denote a three-way contingency table, **N**, so that

- There are I *row* categories, J *column* categories and K *tube* categories
- **N** is of size $I \times J \times K$
- n_{ijk} is the (i, j, k)'th cell frequency
- $n_{i\bullet\bullet}$ is the i'th row marginal frequency
- $n_{\bullet j\bullet}$ is the j'th column marginal frequency
- $n_{\bullet\bullet k}$ as the k'th tube marginal frequency
- the (i, j, k)'th cell proportion by $p_{ijk} = n_{ijk}/n$
- the i'th row marginal proportion by $p_{i\bullet\bullet} = \sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}$
- the j'th column marginal proportion by $p_{\bullet j\bullet} = \sum_{i=1}^{I}\sum_{k=1}^{K} p_{ijk}$
- the k'th tube marginal proportion by $p_{\bullet\bullet k} = \sum_{i=1}^{I}\sum_{j=1}^{J} p_{ijk}$



Indicator Matrix

Multiple Correspondence Analysis

Any sized contingency table can be coded using an *indicator table*, denoted by $\mathbf{Z}$. Suppose we wish to perform a correspondence analysis on our two-way table using its indicator matrix, $\mathbf{Z}$.

$\mathbf{Z}$ is formed by concatenating two sub-matrices (one for each variable) such that

$$\mathbf{Z} = [\mathbf{Z}_1 \ \mathbf{Z}_2]$$

Each row of the indicator matrix represents how each individual is classified into the categories.

For the first variable, $\mathbf{Z}_1$ consists of only the elements 1 and 0; 1 where the individuals has a characteristic, 0 where it doesn't.

$$\mathbf{Z}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Z}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

5

Indicator Matrix

Multiple Correspondence Analysis

Therefore, for an **M – way contingency table** of size $I_1 \times I_2 \times \ldots \times I_M \ldots$

. . . that cross classifies n individuals or units, its indicator table form is of dimension

$$n \times \left(\sum_{m=1}^{M} I_m \right)$$

For example, and $I \times J$ contingency table has an indicator table of dimension

$$n \times (I + J)$$

One way to perform multiple correspondence analysis is to perform simple correspondence analysis on the indicator table.

$$\mathbf{Z}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{Z}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

6

Example: Naples Hospital Data

- In 2003 a group of Italian researchers investigated Patient *Satisfaction* (I) in a Naples hospital (n = 1049). The two most important factors
  - *Cleanliness* of the hospital (J)
  - *Management* of the hospital (K)

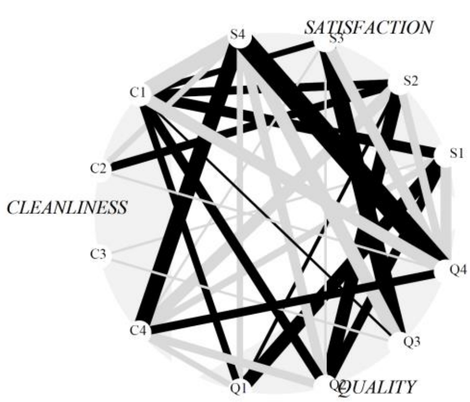
"How are these three variables related to one another?"

Satisfaction	*Cleanliness* 1 (Poor)	2	3	4 (Exce)
	Quality of Management 1 (Low)			
1 (Poor)	16	2	2	1
2 (Fair)	8	1	2	2
3 (Good)	1	4	2	1
4 (Excellent)	1	1	2	3
	Quality of Management 2			
1	14	6	2	1
2	26	2	10	6
3	4	1	6	4
4	2	2	2	2
	Quality of Management 3			
1	6	2	1	1
2	4	14	2	0
3	36	4	14	82
4	4	8	2	14
	Quality of Management 4 (High)			
1	4	2	0	4
2	16	22	10	4
3	44	22	24	50
4	10	24	98	382

Source: D'Ambra, L. *et al.* (2004) In R : `naples.dat`

Example: Naples Hospital Data

Cobweb Diagram



By using standardised residuals

- Black line indicates *positive* association
- Grey line indicates a *negative* association.
- Thick lines indicate a strong association*.
- Thin lines weak association*.

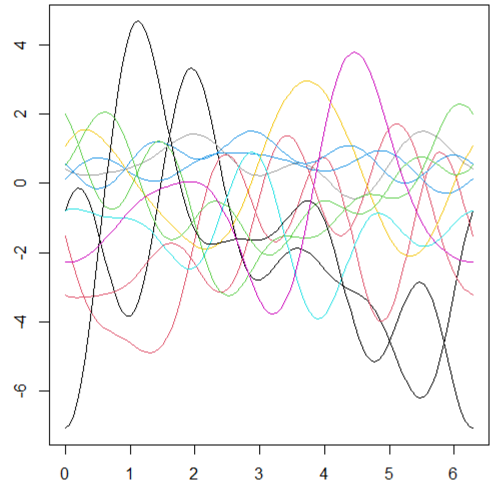
*Positive and negative association

Using the R function `cobweb()` that appears in the Appendix of Upton (2017)

Example: Naples Hospital Data

Multiple Correspondence Analysis

Andrew's Plot



Using the R function `andrewsplot()` in the `pracma` package

Example: Naples Hospital Data

Multiple Correspondence Analysis

The Indicator table

```
> indicator.exe(naples.dat)[c(1,10,20,30,40,50,100,150,300,500),]
```

	S1	S2	S3	S4	C1	C2	C3	C4	Q1	Q2	Q3	Q4
Individual 1	1	0	0	0	1	0	0	0	1	0	0	0
Individual 10	1	0	0	0	1	0	0	0	1	0	0	0
Individual 20	1	0	0	0	1	0	0	0	0	1	0	0
Individual 30	1	0	0	0	1	0	0	0	0	1	0	0
Individual 40	1	0	0	0	1	0	0	0	0	0	0	1
Individual 50	1	0	0	0	0	1	0	0	0	0	1	0
Individual 100	0	1	0	0	1	0	0	0	0	0	1	0
Individual 150	0	1	0	0	0	1	0	0	0	0	0	1
Individual 300	0	0	1	0	0	1	0	0	0	0	0	1
Individual 500	0	0	0	1	1	0	0	0	0	0	0	1

$\mathbf{Z}_1$ — *Satisfaction*; $\mathbf{Z}_2$ — *Cleanliness*; $\mathbf{Z}_3$ — *Quality*

Function `indicator.exe()` appears at the end of these slides (Appendix A)

Example: Naples Hospital Data

Multiple Correspondence Analysis

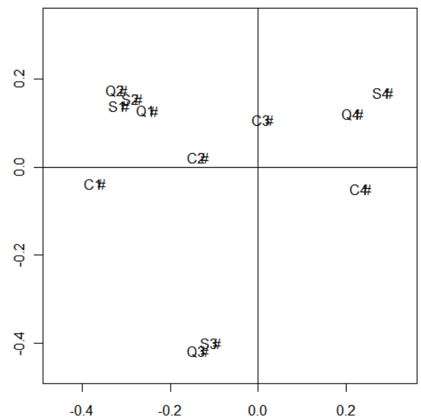
The Indicator table

```
> CAv.naples.indicator <-
+                  CAvariants(indicator.exe(naples.dat))
> CAv.naples.indicator
.
.
Total inertia  3

The inertia values, their percentage contribution to
the total inertia and the cumulative percent inertias

             inertia inertiapc cuminertiapc
value1        0.6800    22.667       22.667
value2        0.4830    16.100       38.767
value3        0.4045    13.483       52.250
value4        0.3478    11.593       63.843
value5        0.3242    10.807       74.650
value6        0.2643     8.811       83.461
value7        0.2082     6.941       90.402
value8        0.1746     5.821       96.223
value9        0.1133     3.777      100.000
```

Total Inertia= $\frac{I+J+K}{3} - 1 = 3$



. . . poor quality (38.8%) 2D visual summary of the association . . .

11

Example: Naples Hospital Data

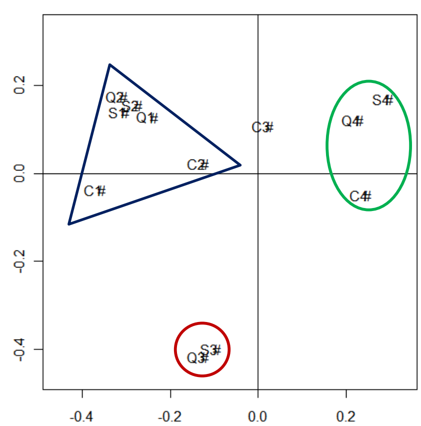
Multiple Correspondence Analysis

The Indicator table

There are three main clusters of points:

1) There exists an association between the two lowest levels of *quality*, *satisfaction* and cleanliness

2) Good quality of management is strongly associated with good satisfaction

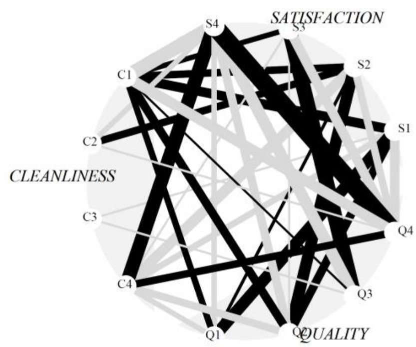


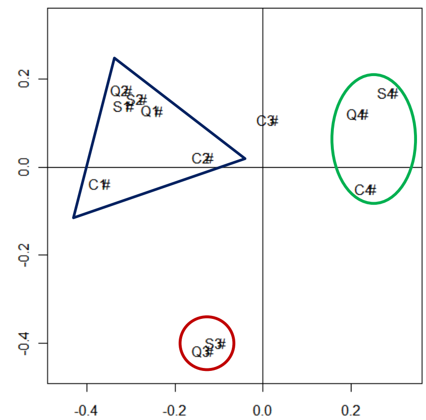
3) There is an association between excellent levels of satisfaction, quality of management and cleanliness

12

Example: Naples Hospital Data

The Indicator table





- Correspondence plot may only reflect 38.8% of the association but it is consistent with the cobweb diagram
- Correspondence plot shows how categories of the same variable compare
- Cobweb diagram looks rather messy

Burt Matrix

For the coding of a two-way contingency table, the *Burt table* consists of the concatenation of the original contingency table with diagonal matrices consisting of the row, column and tube marginal frequencies

The Burt table (Burt, 1950) for a three-way table is

$$\mathbf{B} = \begin{pmatrix} \mathbf{D}_I & \mathbf{N}_{12} & \mathbf{N}_{13} \\ \mathbf{N}_{12}^T & \mathbf{D}_J & \mathbf{N}_{22} \\ \mathbf{N}_{13}^T & \mathbf{N}_{23}^T & \mathbf{D}_K \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_1^T\mathbf{Z}_1 & \mathbf{Z}_1^T\mathbf{Z}_2 & \mathbf{Z}_1^T\mathbf{Z}_3 \\ \mathbf{Z}_2^T\mathbf{Z}_1 & \mathbf{Z}_2^T\mathbf{Z}_2 & \mathbf{Z}_2^T\mathbf{Z}_3 \\ \mathbf{Z}_3^T\mathbf{Z}_1 & \mathbf{Z}_3^T\mathbf{Z}_2 & \mathbf{Z}_3^T\mathbf{Z}_3 \end{pmatrix}$$

Sir Cyril Burt
(1883 – 1971)

where, for example, $\mathbf{N}_{12}$ is the two-way contingency table of the first (row) and second (column) variables formed by aggregating across the third (tube) variable

Example: Naples Hospital Data

Multiple Correspondence Analysis

The Burt table

```
> burt.exe(naples.dat)
```

	S1	S2	S3	S4	C1	C2	C3	C4	Q1	Q2	Q3	Q4
S1	64	0	0	0	40	12	5	7	21	23	10	10
S2	0	12	0	0	54	39	24	12	13	44	20	52
S3	0	$\mathbf{D}_I$	9	0	85	31	46	137	8	15	136	140
S4	0	0	0	557	17	35	104	401	7	8	28	514
C1	40	54	85	17	196	0	0	0	26	46	50	74
C2	12	39	31	35	0	11	0	0	8	11	28	70
C3	5	24	46	104	0	$\mathbf{D}_J$	9	0	8	20	19	132
C4	7	12	137	401	0	0	0	557	7	13	97	440
Q1	21	13	8	7	26	8	8	7	49	0	0	0
Q2	23	44	15	8	46	11	20	13	0	9	$\mathbf{D}_K$	0
Q3	10	20	136	28	50	28	19	97	0	[illegible]	[illegible]	0
Q4	10	52	140	514	74	70	132	440	0	0	0	716

15

Function `burt.exe()` appears at the end of these slides (Appendix B)

Example: Naples Hospital Data

Multiple Correspondence Analysis

The Burt table

```
> burt.exe(naples.dat)
```

	S1	S2	S3	S4	C1	C2	C3	C4	Q1	Q2	Q3	Q4
S1	64	0	0	0	40	12	5	7	21	23	10	10
S2	0	129	0	0	54	[illegible]	$\mathbf{N}_{IJ}$	12	13	[illegible]	$\mathbf{N}_{IK}$	52
S3	0	0	299	0	85	[illegible]	[illegible]	137	8	[illegible]	[illegible]	140
S4	0	0	0	557	17	35	104	401	7	8	28	514
C1	40	54	85	17	196	0	0	0	26	46	50	74
C2	12	[illegible]	$\mathbf{N}_{IJ}^T$	35	0	117	0	0	8	[illegible]	$\mathbf{N}_{JK}$	70
C3	5	[illegible]	[illegible]	104	0	0	179	0	8	[illegible]	[illegible]	132
C4	7	12	137	401	0	0	0	557	7	13	97	440
Q1	21	13	8	7	26	8	8	7	49	0	0	0
Q2	23	[illegible]	$\mathbf{N}_{IK}^T$	8	46	$\mathbf{N}_{JK}^T$	[illegible]	13	0	90	0	0
Q3	10	[illegible]	[illegible]	28	50	[illegible]	[illegible]	97	0	0	194	0
Q4	10	52	140	514	74	70	132	440	0	0	0	716

16

Function `burt.exe()` appears at the end of these slides (Appendix B)

Example: Naples Hospital Data

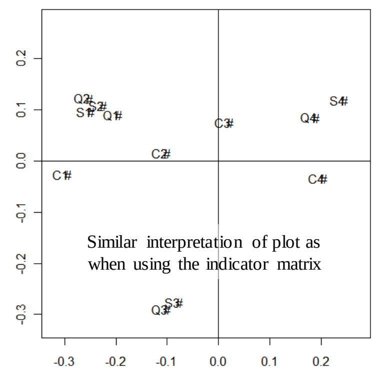
Multiple Correspondence Analysis

The Burt table

```
> CAvariant.naples.burt <- CAvariants(burt.exe(naples.dat))
> CAvariant.naples.burt
.
.
.
Total inertia  1.242

The inertia values, their percentage contribution
to the total inertia and the cumulative percent
inertias

        inertia inertiapc cuminertiapc
value1    0.462    37.234       37.234
value2    0.233    18.787       56.020
value3    0.164    13.171       69.191
value4    0.121     9.742       78.933
value5    0.105     8.464       87.397
value6    0.070     5.625       93.023
value7    0.043     3.489       96.512
value8    0.030     2.454       98.966
value9    0.013     1.034      100.000
value10   0.000     0.000      100.000
value11   0.000     0.000      100.000
```



- OK quality correspondence plot (56.0%)
- Better than when using the indicator matrix (38.8%)

17

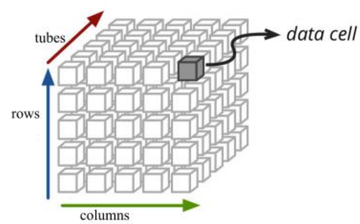
Multi-way Correspondence Analysis & Symmetric Association

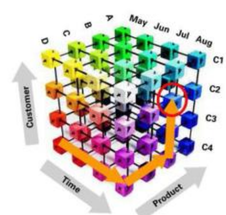
- Partitioning Pearson's Statistic,
- Tucker3 Decomposition,
- Interactive Biplots

Multi-way Correspondence Analysis

An alternative approach to performing correspondence analysis is to consider the following

- Rather than recoding a three- (or multi-) way contingency table into a two way matrix form then performing simple correspondence analysis on that recoded matrix, instead
- Treat the three-way (say) contingency table as a cube (like a Rubic's cube)





This variant is called *multi-way correspondence analysis* (MWCA) . . . and . . .

. . . requires a three- (multi-) way analogue of singular value decomposition.

Multi-Way Correspondence Analysis

19

Pearson's Chi-squared Statistic

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

Then Pearson's three-way chi-squared statistic is

$$X^2 = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{K=1}^{K}\frac{\left(p_{ijk} - p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\right)^2}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{K=1}^{K} p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1\right)^2$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{K=1}^{K} p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\pi_{ijk}^2$$

Multi-Way Correspondence Analysis

20

Decomposing Pearson's Residuals

- To formally test whether there is an (symmetric) association between the row, column, and tubes variables we perform a *chi-squared test of independence*
- Pearson's chi-squared statistic is . . .

$$X^2 = n\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{\left(p_{ijk} - p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\right)^2}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}$$

. . . and has a chi-squared random variable with

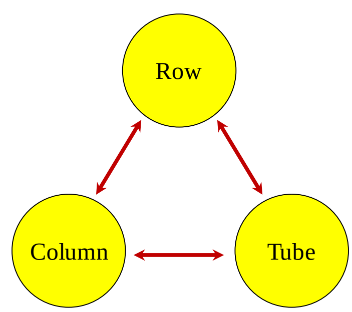
$$(I-1)(J-1) + (I-1)(K-1) + (J-1)(K-1) + (I-1)(J-1)(K-1)$$

degrees of freedom.

Three-way Symmetric Association

Pearson's three-way chi-squared statistic

– Row (I), Column (J) and Tube (K) Variables



$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

$$X^2 = n\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\pi_{ijk}^2$$

Decomposing Pearson's Residuals

Multi-Way Correspondence Analysis

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

CANDECOMP/PARAFAC Decomposition

$$\pi_{ijk} = \sum_{m=1}^{M} a_{im}b_{jm}c_{km}\lambda_m + e_{ijk}$$

- *PARAFAC* (PARallel FACtor analysis)
  - Developed by Harshman (1970)
- *CANDECOMP* (CANonical DECOMPosition)
  - Developed by Carroll & Chang (1970)

} Not easy to define the best rank M!!

Mathematically equivalent (elaborated upon by many)

23

Decomposing Pearson's Residuals

Multi-Way Correspondence Analysis

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

Tucker3 Decomposition

$$\pi_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip}b_{jq}c_{kr}\lambda_{pqr} + e_{ijk}$$

Component matrices: $a_{ip}b_{jq}c_{kr}$

Error: e_{ijk}

- Developed by Tucker (1963) . . . elaborated upon by Tucker (1964, 1966)
- See also Kroonenberg (1983, 2008), Paatero & Andersson (1999), Bro & Kiers (2003), Kiers (2004), Pravdova, Estienne, Walczak & Massart (2001), Beh & Lombardo (2014, 2021) and many others

24

Decomposing Pearson's Residuals

Multi-Way Correspondence Analysis

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

Tucker3 Decomposition

$$\hat{\pi}_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} \lambda_{pqr}$$

Choose a P, Q and R that minimises…not easy the solution are not nested (ALS algorithm)…

$$SSE = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\left(\pi_{ijk} - \hat{\pi}_{ijk}\right)^2$$

25

Decomposing Pearson's Residuals

Multi-Way Correspondence Analysis

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

Tucker3 Decomposition

$$\hat{\pi}_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} \lambda_{pqr}$$

Choose a P, Q and R…

Deviance Plot: A measure of goodness of fit on y-axis and the degree of freedom on x-axis- select a model on or close to an elbow in the higher boundary of the convex hull

(Timmerman, Kiers 2000; Ceulemans, Kiers, 2006; Lombardo, van de Velden, Beh 2022)

26

Decomposing Pearson's Residuals

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

Tucker3 Decomposition

$$\hat{\pi}_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip}b_{jq}c_{kr}\lambda_{pqr}$$

Column-Tube Interactive Biplot (but you can also define principal coords for rows)

$$\tilde{f}_{ip} = a_{ip} \qquad \tilde{g}_{jkp} = \sum_{q=1}^{Q}\sum_{r=1}^{R} b_{jq}c_{kr}\lambda_{pqr}$$

(rows points) ("interactive" column-tube points)

Decomposing Pearson's Residuals

Multi-Way Correspondence Analysis

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

Tucker3 Decomposition

$$\hat{\pi}_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip}b_{jq}c_{kr}\lambda_{pqr}$$

Row-Tube Interactive Biplot (but you can also define principal coords for columns)

$$\tilde{f}_{jq} = b_{jq} \qquad \tilde{g}_{ikq} = \sum_{p=1}^{P}\sum_{r=1}^{R} a_{ip}c_{kr}\lambda_{pqr}$$

(columns points) ("interactive" row-tube points)

Decomposing Pearson's Residuals

Multi-Way Correspondence Analysis

For three *symmetrically* associated categorical variables that are cross-classified to form a three-way contingency table define the *Pearson residuals* by

$$\pi_{ijk} = \frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}} - 1$$

To perform a multi-way correspondence analysis on a three-way contingency table we resort to using two three-way analogues of SVD

Tucker3 Decomposition

$$\hat{\pi}_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip} b_{jq} c_{kr} \lambda_{pqr}$$

Row-Column Interactive Biplot (but you can also define principal coords for tubes)

$$\tilde{f}_{kr} = c_{kr} \qquad \tilde{g}_{ijq} = \sum_{p=1}^{P}\sum_{q=1}^{Q} a_{ip} b_{jq} \lambda_{pqr}$$

(tubes points) ("interactive" row-column points)

29

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

The cross-classification of the three categorical variables (RES, BCA & GEO) that produce the 4 × 4 × 6 contingency table . . .

- RES: indicator of renewable energy share of total final energy consumption
- BCA: an indicator of adjusted emission growth rate for black carbon
- GEO: geographical area

RES		BCA			
		(0, 8.13]	(8.13, 23.1]	(23.1, 49.7]	(49.7, 96.4]
Africa		**BCA1**	**BCA2**	**BCA3**	**BCA4**
(0, 34.5]	RES1	1	0	4	8
(34.5, 53.7]	RES2	0	2	4	19
(53.7, 80.7]	RES3	3	2	1	4
(80.7, 100]	RES4	1	1	0	1
America					
	RES1	1	0	6	2
	RES2	0	2	0	1
	RES3	0	3	1	0
	RES4	1	3	2	1
Asia					
	RES1	5	1	2	0
	RES2	6	2	3	1
	RES3	5	3	3	3
	RES4	4	2	0	0
Australia					
	RES1	1	0	0	0
	RES2	0	1	2	0
	RES3	1	1	5	0
	RES4	1	0	0	0
Carribean					
	RES1	3	2	0	0
	RES2	1	1	0	1
	RES3	0	0	0	0
	RES4	2	3	0	0
Europe					
	RES1	1	0	0	0
	RES2	0	0	1	0
	RES3	0	7	5	0
	RES4	3	11	7	3

30

Example: Sustainable Development Goals Data

The cross-classification of the three categorical variables (RES, BCA & GEO) that produce the 4 × 4 × 6 contingency table . . .

Chi-squared test of independence:

- $X^2 = 290.035$
- A p-value < 0.001
- Thus, a statistically significant association exists between at least two of the variables

. . . but . .

there are other features we'll talk about soon

RES		BCA1 (0, 8.13]	BCA2 (8.13, 23.1]	BCA3 (23.1, 49.7]	BCA4 (49.7, 96.4]
Africa					
(0, 34.5]	RES1	1	0	4	8
(34.5, 53.7]	RES2	0	2	4	19
(53.7, 80.7]	RES3	3	2	1	4
(80.7, 100]	RES4	1	1	0	1
America					
	RES1	1	0	6	2
	RES2	0	2	0	1
	RES3	0	3	1	0
	RES4	1	3	2	1
Asia					
	RES1	5	1	2	0
	RES2	6	2	3	1
	RES3	5	3	3	3
	RES4	4	2	0	0
Australia					
	RES1	1	0	0	0
	RES2	0	1	2	0
	RES3	1	1	5	0
	RES4	1	0	0	0
Carribean					
	RES1	3	2	0	0
	RES2	1	1	0	1
	RES3	0	0	0	0
	RES4	2	3	0	0
Europe					
	RES1	1	0	0	0
	RES2	0	0	1	0
	RES3	0	7	5	0
	RES4	3	11	7	3

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

The cross-classification of the three categorical variables (RES, BCA & GEO) that produce the 4 × 4 × 6 contingency table . . .

Chi-squared test of independence:

- $X^2 = 290.035$

$$X^2 = X^2_{IJ} + X^2_{IK} + X^2_{JK} + X^2_{IJK}$$

Two-way interactions

Three-way interactions

following partition of Lancaster (1951)

RES		BCA1 (0, 8.13]	BCA2 (8.13, 23.1]	BCA3 (23.1, 49.7]	BCA4 (49.7, 96.4]
Africa					
(0, 34.5]	RES1	1	0	4	8
(34.5, 53.7]	RES2	0	2	4	19
(53.7, 80.7]	RES3	3	2	1	4
(80.7, 100]	RES4	1	1	0	1
America					
	RES1	1	0	6	2
	RES2	0	2	0	1
	RES3	0	3	1	0
	RES4	1	3	2	1
Asia					
	RES1	5	1	2	0
	RES2	6	2	3	1
	RES3	5	3	3	3
	RES4	4	2	0	0
Australia					
	RES1	1	0	0	0
	RES2	0	1	2	0
	RES3	1	1	5	0
	RES4	1	0	0	0
Carribean					
	RES1	3	2	0	0
	RES2	1	1	0	1
	RES3	0	0	0	0
	RES4	2	3	0	0
Europe					
	RES1	1	0	0	0
	RES2	0	0	1	0
	RES3	0	7	5	0
	RES4	3	11	7	3

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

Partitioning Pearson's Three-way Chi-squared Statistic

Component	X^2_{IJ}	X^2_{IK}	X^2_{JK}	X^2_{IJK}	X^2
Term	34.970	82.816	95.677	76.573	290.035
P-value	<0.001	<0.001	<0.001	0.002	<0.001
% Contribution	12%	29%	33%	26%	100%

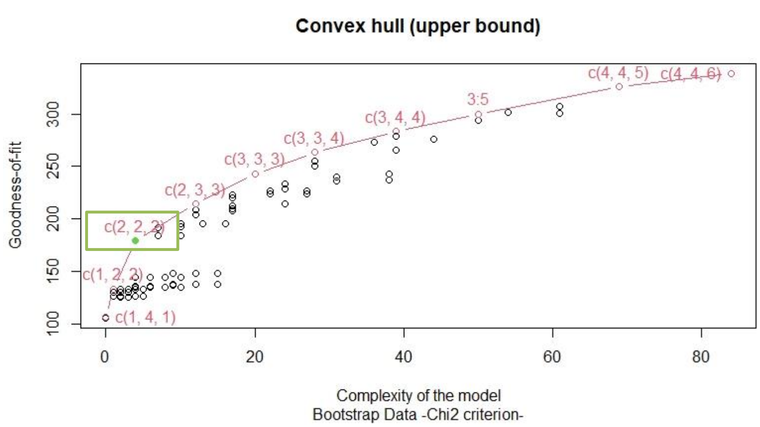
- There is a statistically significant association between at least two of the variables
- There is a statistically significant association between each **pair** of the three variables
- There is a statistically significant association between **all** three variables . . . but . . .
- The association is dominated by the association between BCA & GEO (33%) followed by RES & BCA (29%)

33

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

```
tunelocal(sustainable.dat, boots = T, boottype = "bootpsimple", nboots= 1000)
```



Deviance Plot Goodness of Fit = Chi-squared (vertical axis), df (horizontal axis).

Model D222 is on the hull: $\mathrm{st(i)} = \left(\frac{\mathrm{dev(i+1)} - \mathrm{dev(i)}}{\mathrm{df(i+1)} - \mathrm{df(i)}}\right) \Big/ \left(\frac{\mathrm{dev(i)} - \mathrm{dev(i-1)}}{\mathrm{df(i)} - \mathrm{df(i-1)}}\right)$

34

Example: Sustainable Development Goals Data

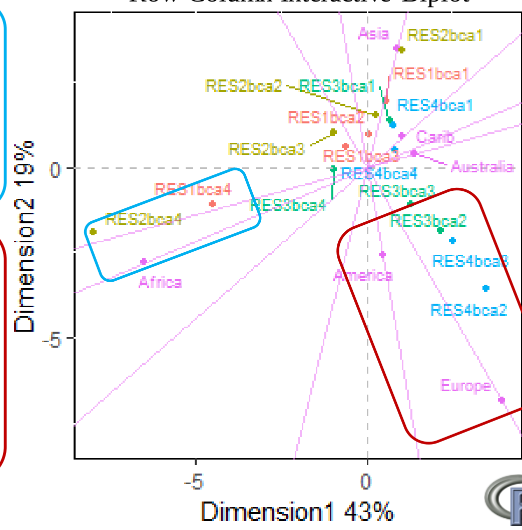
Multi-Way Correspondence Analysis

```
> res.ca3 <- CA3variants(sustainable.dat, dims = c(2,2,2), ca3type = "CA3")
> plot(res.ca3, biptype = "row-column", scaleplot = 1.5)
```

- Association between the lowest levels of *RES* and the highest level of black carbon (RES2bca4 – RES1bca4) in Africa

- The highest levels of RES are strongly associated with middle-low levels of black-carbon (RES4bca2-RES4bca3) in Europe and America

Row-Column Interactive Biplot



35

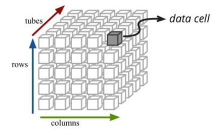
Multi-way Correspondence Analysis & Asymmetric Association

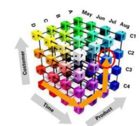
- Marcotorchino’s Index
- Gray-Williams Index
- Lombardo/Lambda Index

Three-way Analogs of $\tau_{A|B}$

Marcotorchnio Index

For three asymmetrically associated variables: dependent on the structure of the asymmetry





Marcotorchino Index

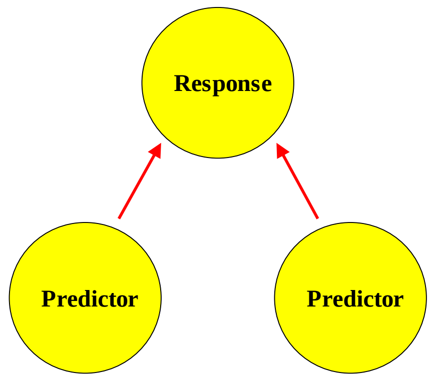
$$\tau_M = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet j \bullet} p_{\bullet\bullet k} \pi_{ijk}^2}{1 - \sum_{i=1}^{I} p_{i\bullet\bullet}^2}$$

where

$$\pi_{ijk} = \frac{p_{ijk}}{p_{\bullet j \bullet} p_{\bullet\bullet k}} - p_{i\bullet\bullet}$$

- Two predictor variables (assumed/treated **independent**)
- One response variable

(Marcotorchino, F., 1984)

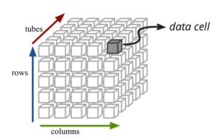


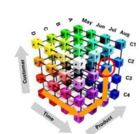
37

Three-way Analogs of $\tau_{A|B}$

Gray-Williams Index

For three asymmetrically associated variables: dependent on the structure of the asymmetry





Gray – Williams Index

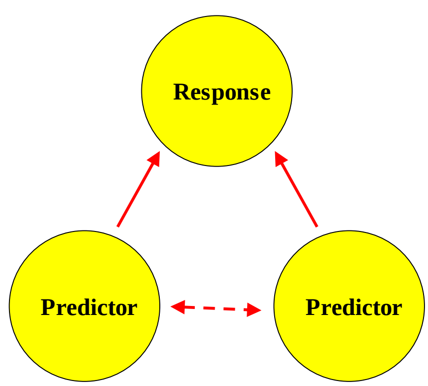
$$\tau_{GW} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet jk} \pi_{ijk}^2}{1 - \sum_{i=1}^{I} p_{i\bullet\bullet}^2}$$

where

$$\pi_{ijk} = \frac{p_{ijk}}{p_{\bullet jk}} - p_{i\bullet\bullet}$$

- Two predictor variables (assumed/treated symmetrically **associated**)
- One response variable

(Gray & Williams, 1981)

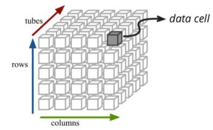


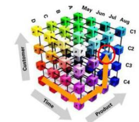
38

Three-way Analogs of $\tau_{A|B}$

Lombardo/Delta Index

For three asymmetrically associated variables: dependent on the structure of the asymmetry





Lombardo/Delta Index

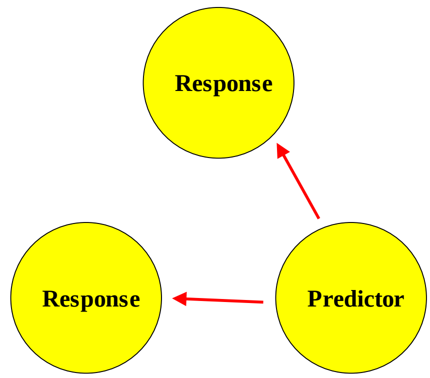
$$\tau_D = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet jk}\pi_{ijk}^2}{1-\sum_{i=1}^{I} p_{i\bullet\bullet}^2}$$

where

$$\pi_{ijk} = \frac{p_{ijk}}{p_{\bullet\bullet k}} - p_{i\bullet\bullet}p_{\bullet j\bullet}$$

- One predictor variable
- Two response variables (assumed/treated **independent**)

(Lombardo, 2011)

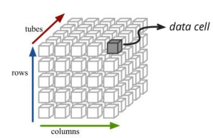


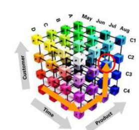
39

Three-way Analogs of $\tau_{A|B}$

Asymmetric Association

For three asymmetrically associated variables: dependent on the structure of the asymmetry





Regardless of the π_{ijk} being considered, there are two ways in which it can be decomposed that are treated as three-way analogues of SVD:

PARAFAC/CANDECOMP Models

$$\pi_{ijk} = \sum_{m=1}^{M} a_{im}b_{jm}c_{km}\lambda_m + e_{ijk}$$

Tucker3 Decomposition

$$\pi_{ijk} = \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} a_{ip}b_{jq}c_{kr}\lambda_{pqr} + e_{ijk}$$

40

Example: Naples Hospital Data

Multi-way Correspondence Analysis

Marcotorchino Index

$$\tau_M = \tau_{IJ} + \tau_{IK} + \tau_{JK} + \tau_{IJK}$$

$$C_{IJ} = 445.614$$ *(Satisfaction vs Cleanliness)*

$$C_{IK} = 712.186$$ *(Satisfaction vs Management)*

$$C_{JK} = 212.348$$ *(Cleanliness vs Management)*

$$C_{IJK} = 299.434$$ *(Three-way Association)*

$$C_M = 1669.581$$ *(Total Association)*

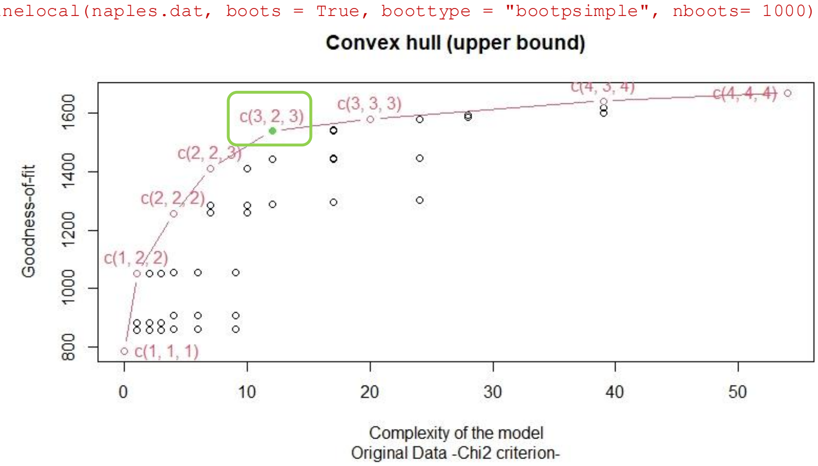
Satisfaction (response)	*Cleanliness* 1 (Poor)	2	3	4 (Exce)
	Quality of Management 1 (Low)			
1 (Poor)	16	2	2	1
2 (Fair)	8	1	2	2
3 (Good)	1	4	2	1
4 (Excellent)	1	1	2	3
	Quality of Management 2			
1	14	6	2	1
2	26	2	10	6
3	4	1	6	4
4	2	2	2	2
	Quality of Management 3			
1	6	2	1	1
2	4	14	2	0
3	36	4	14	82
4	4	8	2	14
	Quality of Management 4 (High)			
1	4	2	0	4
2	16	22	10	4
3	44	22	24	50
4	10	24	98	382

41

Example: Naples Hospital Data

Multi-way Correspondence Analysis

```
tunelocal(naples.dat, boots = True, boottype = "bootpsimple", nboots= 1000)
```



Deviance Plot Goodness of Fit = Marcotorchino (vertical axis), df (horizontal axis).

Model D323 is on the hull: $$st(i) = \left(\frac{dev(i+1) - dev(i)}{df(i+1) - df(i)}\right) \Big/ \left(\frac{dev(i) - dev(i-1)}{df(i) - df(i-1)}\right)$$

42

Example: Naples Hospital Data

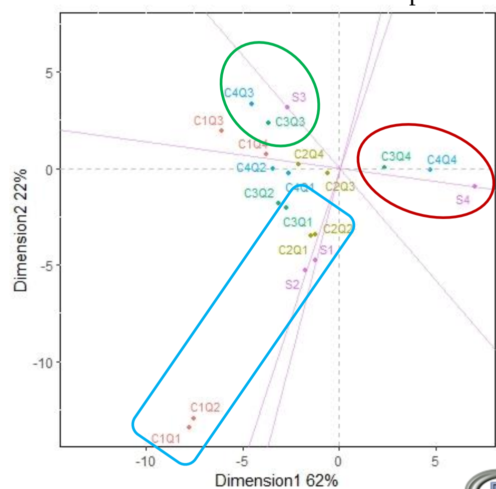
Multi-way Correspondence Analysis

43

```
> res.ca3 <- CA3variants(naples.dat, dims=c(3,2,3),ca3type ="NSCA3")
> plot(res.ca3, biptype = "pred", scaleplot = 1.5)
```

- Predictability of the two lowest levels of *satisfaction* given quality & cleanliness
- Good satisfaction is well predicted by the good quality of management & cleanliness
- Association between excellent levels of satisfaction given excellent levels of cleanliness & quality of management

Column-Tube Interactive Biplot



Multi-way Correspondence Analysis & Symmetric Association

- Cressie-Read family of Divergence Statistics

Three-way Family of Divergence Statistics

Multi-Way Correspondence Analysis

For some δ, the Cressie-Read family of divergence statistics for a three-way contingency table is

$$CR(\delta) = \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} p_{ijk} \left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet} p_{\bullet j \bullet} p_{\bullet\bullet k}} \right)^{\delta} - 1 \right]$$

where $\delta \in (-\infty, \infty)$. (Pardo, 1996; Lombardo & Beh, 2022)

This is a chi-squared random variable with

$$(I-1)(J-1) + (I-1)(K-1) + (J-1)(K-1) + (I-1)(J-1)(K-1)$$

degrees of freedom.

45

Three-way Family of Divergence Statistics

Multi-Way Correspondence Analysis

For some δ, the Cressie-Read family of divergence statistics for a three-way contingency table is

$$CR(\delta) = \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} p_{ijk} \left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet} p_{\bullet j \bullet} p_{\bullet\bullet k}} \right)^{\delta} - 1 \right]$$

where $\delta \in (-\infty, \infty)$. (Pardo, 1996; Lombardo & Beh, 2022)

Pearson's chi-squared statistic

$$CR(\delta = 1) = X^2 = n \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{\left(p_{ijk} - p_{i\bullet\bullet} p_{\bullet j \bullet} p_{\bullet\bullet k}\right)^2}{p_{i\bullet\bullet} p_{\bullet j \bullet} p_{\bullet\bullet k}}$$

Likelihood Ratio Statistic

$$CR(0) = G^2 = 2n \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} p_{ijk} \ln\left(\frac{p_{ijk}}{p_{i\bullet\bullet} p_{\bullet j \bullet} p_{\bullet\bullet k}} \right)$$

46

Three-way Family of Divergence Statistics

For some δ, the Cressie-Read family of divergence statistics for a three-way contingency table is

$$CR(\delta) = \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

where $\delta \in (-\infty, \infty)$. (Pardo, 1996; Lombardo & Beh, 2022)

Freeman-Tukey statistic

$$CR\left(-\frac{1}{2}\right) = T^2 = 4n \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\sqrt{p_{ijk}} - \sqrt{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^2$$

Cressie-Read Statistic

$$CR\left(\frac{2}{3}\right) = CR^2 = \frac{9n}{5}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}\left(\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{2/3} - 1\right)^2$$

Partitioning the Family of Divergence Statistics

Lombardo & Beh (2022) showed that the three-way Cressie-Read divergence statistic can be partitioned as follows

$$CR(\delta) = \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

$$= \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{J}\sum_{j=1}^{J} p_{ij\bullet}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}}\right)^{\delta} - 1\right] + \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{k=1}^{K} p_{i\bullet k}\left[\left(\frac{p_{i\bullet k}}{p_{i\bullet\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

$$= \frac{2n}{\delta(\delta+1)}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet jk}\left[\left(\frac{p_{\bullet jk}}{p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right] + \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{\alpha p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

$$= CR_{IJ}(\delta) + CR_{IK}(\delta) + CR_{JK}(\delta) + CR_{IJK}(\delta)$$

Cressie-Read family of divergence statistics for only the **row** and **column** variables (aggregating across the **tube** variable)

Partitioning the Family of Divergence Statistics

Lombardo & Beh (2022) showed that the three-way Cressie-Read divergence statistic can be partitioned as follows

$$
\begin{aligned}
CR(\delta) &= \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] \\
&= \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{J}\sum_{j=1}^{J} p_{ij\bullet}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}}\right)^{\delta}-1\right] + \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{k=1}^{K} p_{i\bullet k}\left[\left(\frac{p_{i\bullet k}}{p_{i\bullet\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] \\
&= \frac{2n}{\delta(\delta+1)}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet jk}\left[\left(\frac{p_{\bullet jk}}{p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] + \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{\alpha p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] \\
&= CR_{IJ}(\delta) + CR_{IK}(\delta) + CR_{JK}(\delta) + CR_{IJK}(\delta)
\end{aligned}
$$

Cressie-Read family of divergence statistics for only the **row** and **tube** variables (aggregating across the **column** variable)

Multi-Way Correspondence Analysis

49

Partitioning the Family of Divergence Statistics

Lombardo & Beh (2022) showed that the three-way Cressie-Read divergence statistic can be partitioned as follows

$$
\begin{aligned}
CR(\delta) &= \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] \\
&= \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{J}\sum_{j=1}^{J} p_{ij\bullet}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}}\right)^{\delta}-1\right] + \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{k=1}^{K} p_{i\bullet k}\left[\left(\frac{p_{i\bullet k}}{p_{i\bullet\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] \\
&= \frac{2n}{\delta(\delta+1)}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet jk}\left[\left(\frac{p_{\bullet jk}}{p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] + \frac{2n}{\delta(\delta+1)}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{\alpha p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta}-1\right] \\
&= CR_{IJ}(\delta) + CR_{IK}(\delta) + CR_{JK}(\delta) + CR_{IJK}(\delta)
\end{aligned}
$$

Cressie-Read family of divergence statistics for only the **column** and **tube** variables (aggregating across the **row** variable)

Multi-Way Correspondence Analysis

50

Partitioning the Family of Divergence Statistics

Multi-Way Correspondence Analysis

Lombardo & Beh (2022) showed that the three-way Cressie-Read divergence statistic can be partitioned as follows

$$CR(\delta) = \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

$$= \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{J}\sum_{j=1}^{J} p_{ij\bullet}\left[\left(\frac{p_{ijk}}{p_{i\bullet\bullet}p_{\bullet j\bullet}}\right)^{\delta} - 1\right] + \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I}\sum_{k=1}^{K} p_{i\bullet k}\left[\left(\frac{p_{i\bullet k}}{p_{i\bullet\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

$$= \frac{2n}{\delta(\delta+1)} \sum_{j=1}^{J}\sum_{k=1}^{K} p_{\bullet jk}\left[\left(\frac{p_{\bullet jk}}{p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right] + \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} p_{ijk}\left[\left(\frac{p_{ijk}}{\alpha p_{i\bullet\bullet}p_{\bullet j\bullet}p_{\bullet\bullet k}}\right)^{\delta} - 1\right]$$

$$= CR_{IJ}(\delta) + CR_{IK}(\delta) + CR_{JK}(\delta) + CR_{IJK}(\delta)$$

Cressie-Read family of divergence statistics for all **three** the variables (jointly)

51

Partitioning the Family of Divergence Statistics

Multi-Way Correspondence Analysis

Example 1:

When $\delta = 1$

$$CR(1) = CR_{IJ}(1) + CR_{IK}(1) + CR_{JK}(1) + CR_{IJK}(1)$$

is the same as

$$X^2 = X_{IJ}^2 + X_{IK}^2 + X_{JK}^2 + X_{IJK}^2$$

(see Part III slide 29)

Example 2:

When $\delta = 0$

$$CR(0) = CR_{IJ}(0) + CR_{IK}(0) + CR_{JK}(0) + CR_{IJK}(0)$$

is the partition of the three-way *likelihood ratio* statistic such that

$$G^2 = G_{IJ}^2 + G_{IK}^2 + G_{JK}^2 + G_{IJK}^2$$

52

Partitioning the Family of Divergence Statistics

Multi-Way Correspondence Analysis

Example 1:

When $\delta = 1$

$$CR(1) = CR_{IJ}(1) + CR_{IK}(1) + CR_{JK}(1) + CR_{IJK}(1)$$

is the same as

$$X^2 = X^2_{IJ} + X^2_{IK} + X^2_{JK} + X^2_{IJK}$$ (see Part III slide 29)

Example 3:

When $\delta = 1/2$

$$CR\left(\frac{1}{2}\right) = CR_{IJ}\left(\frac{1}{2}\right) + CR_{IK}\left(\frac{1}{2}\right) + CR_{JK}\left(\frac{1}{2}\right) + CR_{IJK}\left(\frac{1}{2}\right)$$

is the partition of the three-way *Freeman-Tukey* statistic such that

$$T^2 = T^2_{IJ} + T^2_{IK} + T^2_{JK} + T^2_{IJK}$$

53

Partitioning the Family of Divergence Statistics

Multi-Way Correspondence Analysis

Example 1:

When $\delta = 1$

$$CR(1) = CR_{IJ}(1) + CR_{IK}(1) + CR_{JK}(1) + CR_{IJK}(1)$$

is the same as

$$X^2 = X^2_{IJ} + X^2_{IK} + X^2_{JK} + X^2_{IJK}$$ (see Part III slide 29)

Example 4:

When $\delta = 2/3$

$$CR\left(\frac{2}{3}\right) = CR_{IJ}\left(\frac{2}{3}\right) + CR_{IK}\left(\frac{2}{3}\right) + CR_{JK}\left(\frac{2}{3}\right) + CR_{IJK}\left(\frac{2}{3}\right)$$

is the partition of the three-way *Cressie-Read* statistic such that

$$CR = CR_{IJ} + CR_{IK} + CR_{JK} + CR_{IJK}$$

54

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

The cross-classification of the three categorical variables (RES, BCA & GEO) that produce the 4 × 4 × 6 contingency table . . .

Chi-squared test of independence:

- $X^2 = 290.035$
- A p-value < 0.001
- Thus, a statistically significant association exists between at least two of the variables

RES		BCA (0, 8.13] BCA1	(8.13, 23.1] BCA2	(23.1, 49.7] BCA3	(49.7, 96.4] BCA4
Africa					
(0, 34.5]	RES1	1	0	4	8
(34.5, 53.7]	RES2	0	2	4	19
(53.7, 80.7]	RES3	3	2	1	4
(80.7, 100]	RES4	1	1	0	1
America					
	RES1	1	0	6	2
	RES2	0	2	0	1
	RES3	0	3	1	0
	RES4	1	3	2	1
Asia					
	RES1	5	1	2	0
	RES2	6	2	3	1
	RES3	5	3	3	3
	RES4	4	2	0	0
Australia					
	RES1	1	0	0	0
	RES2	0	1	2	0
	RES3	1	1	5	0
	RES4	1	0	0	0
Carribean					
	RES1	3	2	0	0
	RES2	1	1	0	1
	RES3	0	0	0	0
	RES4	2	3	0	0
Europe					
	RES1	1	0	0	0
	RES2	0	0	1	0
	RES3	0	7	5	0
	RES4	3	11	7	3

55

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

Partitioning Pearson's Three-way Statistic

Component	X^2_{IJ}	X^2_{IK}	X^2_{JK}	X^2_{IJK}	X^2
Term	34.970	82.816	95.677	76.573	290.035
P-value	<0.001	<0.001	<0.001	0.002	<0.001
% Contribution	12%	29%	33%	26%	100%

- There is a statistically significant association between at least two of the variables
- There is a statistically significant association between each **pair** of the three variables
- There is a statistically significant association between **all** three variables . . . but . . .
- The association is dominated by the association between BCA & GEO (33%) followed by RES & BCA (29%)

56

Example: Sustainable Development Goals Data

The cross-classification of the three categorical variables (RES, BCA & GEO) that produce the 4 × 4 × 6 contingency table . . .

Chi-squared test of independence:

- $X^2 = 290.035$
- A p-value < 0.001
- Thus, a statistically significant association exists between at least two of the variables

. . . but . .

- **Sparse contingency table**
- **Multiple zero cell frequencies**

RES		BCA (0, 8.13] BCA1	(8.13, 23.1] BCA2	(23.1, 49.7] BCA3	(49.7, 96.4] BCA4
Africa					
(0, 34.5]	RES1	1	0	4	8
(34.5, 53.7]	RES2	0	2	4	19
(53.7, 80.7]	RES3	3	2	1	4
(80.7, 100]	RES4	1	1	0	1
America					
	RES1	1	0	6	2
	RES2	0	2	0	1
	RES3	0	3	1	0
	RES4	1	3	2	1
Asia					
	RES1	5	1	2	0
	RES2	6	2	3	1
	RES3	5	3	3	3
	RES4	4	2	0	0
Australia					
	RES1	1	0	0	0
	RES2	0	1	2	0
	RES3	1	1	5	0
	RES4	1	0	0	0
Carribean					
	RES1	3	2	0	0
	RES2	1	1	0	1
	RES3	0	0	0	0
	RES4	2	3	0	0
Europe					
	RES1	1	0	0	0
	RES2	0	0	1	0
	RES3	0	7	5	0
	RES4	3	11	7	3

Example: Sustainable Development Goals Data

Partitioning the Three-way Freeman-Tukey Statistic

Component	T^2_{IJ}	T^2_{IK}	T^2_{JK}	T^2_{IJK}	T^2
Term	35.112	83.340	92.421	39.619	250.491
P-value	<0.001	<0.001	<0.001	0.699	<0.001
% Contribution	14%	33%	37%	16%	100%

- There is a statistically significant association between at least two of the variables
- There is a statistically significant association between each **pair** of the three variables
- . . . but . . .the three-way association is **not** statistically significant (p-value = 0.699)
- The association is still dominated by the association between BCA & GEO (37%) followed by RES & BCA (33%)

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

Partitioning the Three-way Cressie-Read Statistic

Component	CR_{IJ}	CR_{IK}	CR_{JK}	CR_{IJK}	CR
Term	34.998	82.653	93.105	48.559	259.035
P-value	<0.001	<0.001	<0.001	0.332	<0.001
% Contribution	13%	32%	36%	19%	100%

- There is a statistically significant association between at least two of the variables
- There is a statistically significant association between each **pair** of the three variables
- . . . but . . .the three-way association is **not** statistically significant (p-value = 0.332)
- The association is dominated by the association between BCA & GEO (36%) followed by RES & BCA (32%)

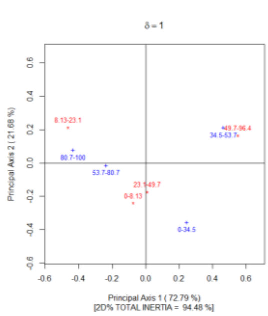
59

Example: Sustainable Development Goals Data

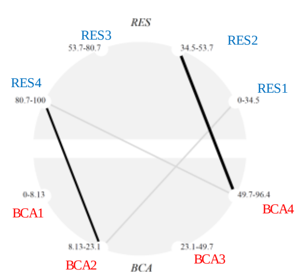
Multi-Way Correspondence Analysis

RES vs BCA

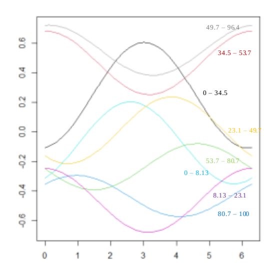
Pearson



Cobweb Diagram



Andrews' Plot



$$X^2_{IJ} = 34.97$$

(12% of total inertia)

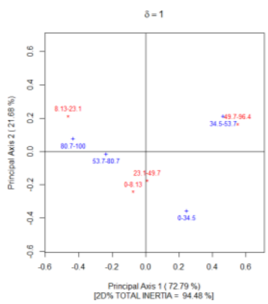
60

Example: Sustainable Development Goals Data

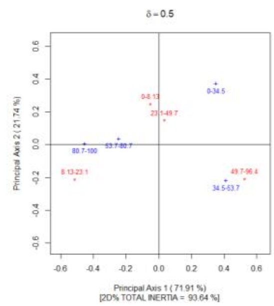
Multi-Way Correspondence Analysis

RES vs BCA

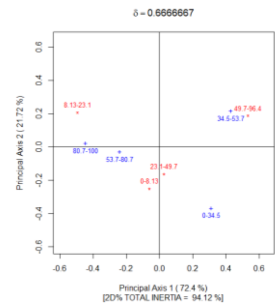
Pearson



Freeman-Tukey



Cressie-Read



$X^2_{IJ} = 34.97$

(12% of total inertia)

$T^2_{IJ} = 35.112$

(14% of total inertia)

$CR_{IJ} = 34.998$

(13% of total inertia)

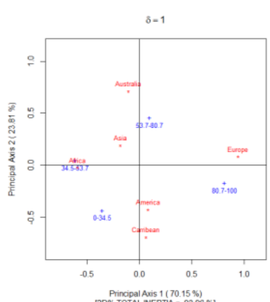
61

Example: Sustainable Development Goals Data

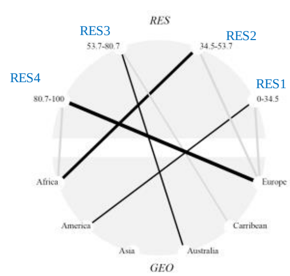
Multi-Way Correspondence Analysis

RES vs GEO

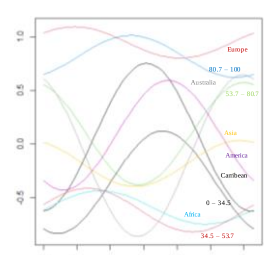
Pearson



Cobweb Diagram



Andrews' Plot



$X^2_{IK} = 82.816$

(29% of total inertia)

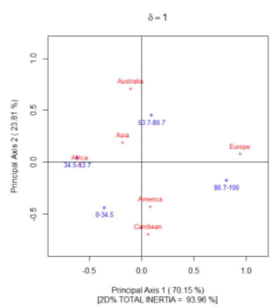
62

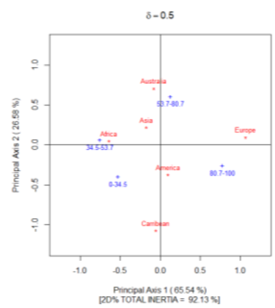
Example: Sustainable Development Goals Data

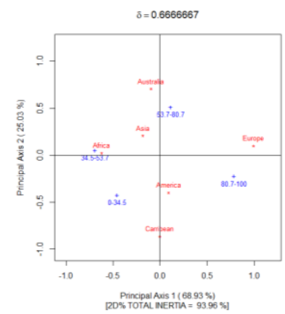
Multi-Way Correspondence Analysis

RES vs GEO

Pearson	Freeman-Tukey	Cressie-Read







$X^2_{IK} = 82.816$ (29% of total inertia)

$T^2_{IK} = 83.340$ (35% of total inertia)

$CR_{IK} = 82.653$ (32% of total inertia)

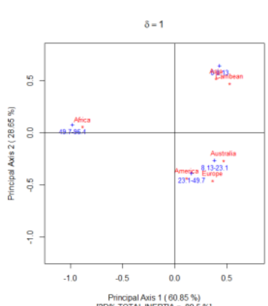
63

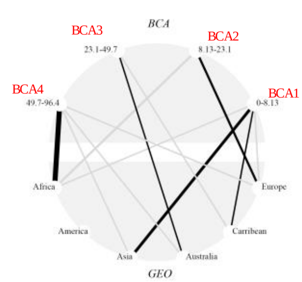
Example: Sustainable Development Goals Data

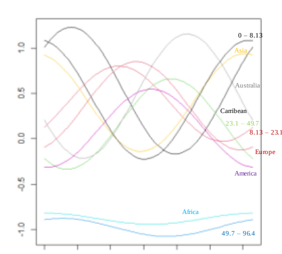
Multi-Way Correspondence Analysis

BCA vs GEO

Pearson	Cobweb Diagram	Andrews' Plot







$X^2_{JK} = 95.677$

(33% of total inertia)

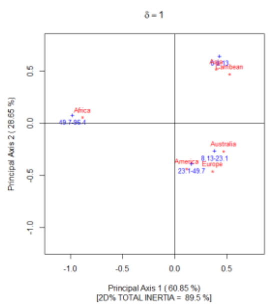
64

Example: Sustainable Development Goals Data

Multi-Way Correspondence Analysis

BCA vs GEO

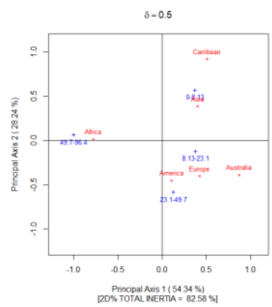
Pearson



$X_{JK}^2 = 95.677$

(33% of total inertia)

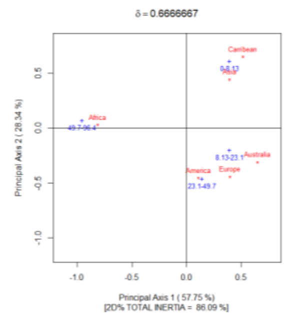
Freeman-Tukey



$T_{JK}^2 = 92.421$

(37% of total inertia)

Cressie-Read



$CR_{JK} = 93.105$

(36% of total inertia)

65

References

Bibliographic Information

1. Beh, E. J. & Lombardo, R. (2014), *Correspondence Analysis: Theory, Methods and New Strategies*, Wiley.
2. Beh, E. J. & Lombardo, R. (2021), *An Introduction to Correspondence Analysis*, Wiley.
3. Bro, R. & Kiers, H. A. L. (2003), A new efficient method for determining the number of components in PARAFAC models, *Journal of Chemometrics*, 17, 274 – 286.
4. Burt, C. (1950) The factorial analysis of qualitative data. *British Journal of Statistical Psychology*, 3, 166 – 185.
5. Carrol, J. and Chang, J.J. (1970) Analysis of individual differences in multidimensional scaling via an *n*-way generalization of Eckart--Young decomposition. *Psychometrika*, 35, 283 – 319.
6. D'Ambra, L., Amenta, P., Lombardo, L., Gallo, M. & Sarnacchiaro, P. (2004), Analisi statistica multivariate per la valutazione della patient satisfaction. In *Qualita e Valutazione delle Strutture Sanitarie* (eds A. Pagano & G. Vittadini), Etas, Milano.
7. Gray, L. N. & Williams, J. S. (1981), Goodman and Kruskal τ_b multiple and partial analogs. *Proceedings of the Social Statistics Section*, 10, 50--62.
8. Harshman, R. A. (1970), Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multi-modal factor analysis, *UCLA Working Papers in Phonetics*, 16, 1 – 84.
9. Kiers, H. A. L. (2004), Bootstrap confidence intervals for three-way methods, *Journal of Chemometrics*, 18, 22 – 36.
10. Kroonenberg, P. M. (1989), Singular value decomposition of interactions in three-way contingency tables. In *Multiway Data Analysis* (eds. R. Coppi and S. Bolasco), pp. 169 – 184. Elsevier
11. Kroonenberg, P. M. (2008), *Applied Multiway Data Analysis*, Wiley.
12. Lancaster, H. O. (1951), Complex contingency tables treated by the partition of χ^2, *Journal of the Royal Statistical Society (Series B)*, 13, 242 – 249.
13. Lombardo, R. (2011), Three-way association measure decompositions: the Delta index, *Journal of Statistical Planning and Inference*, 141, 1789 – 1799.
14. Lombardo, R., Beh, E. J. & Kroonenberg, P. M. (2021), Symmetrical and non-symmetrical variants of three-way correspondence analysis for ordered variables, *Statistical Science*, 36, 542 – 561

66

References

Bibliographic Information

15. Lombardo, R. and Beh, E. J. (2022), Partitioning the Cressie-Read divergence statistic for three-way contingency tables: a study on environmental sustainability data, *Proceedings of the IES2002* (eds. R. Lombardo, I. Camminatiello and V. Simonacci), 491 – 497.
16. Lombardo, R, van de Velden M.. and Beh, E. J. (2022) Three-way Correspondence Analysis in R. Submitted.
17. Marcotorchino, F. (1984), *Utilisation des comparaisons par paires en statistique des contingences: Partie II*, Report # F 071, Etude du Centre Scientifique, IBM, France.
18. Paatero, P. & Andersson, C. A. (1999), Further improvements of the speed of the Tucker3 three-way algorithm, *Chemometrics and Intelligent Laboratory Systems*, 47, 17 – 20.
19. Pardo, M. C. (1996), An empirical investigation of Cressie and Read tests for the hypothesis of independence in a three-way contingency table, *Kybernetika*, 32, 175 – 183.
20. Pravdova, V., Estienne, F., Walczak, B. & Massart, D. L. (2001), A robust version of the Tucker3 model, *Chemometrics and Intelligent Laboratory Systems*, 59, 75 – 88.
21. Tucker, L. R. (1963), Implications of factor analysis of three-way matrices for measurement of change. In *Problems in Measuring Change* (ed. C. W. Harris), pp. 122 – 137. University of Wisconsin Press.
22. Tucker, L. R. (1964) The extension of factor analysis to three-dimensional matrices. In *Contributions to Mathematical Psychology* (eds. N. Frederiksen & H. Gulliksen), pp. 109 – 127. Holt, Rinehart & Winston.
23. Tucker, L. R. (1966), Some mathematical notes on three mode factor analysis, *Psychometrika*, 31, 279 – 311.
24. Upton, G. J. G. (2017), *Categorical Data Analysis by Example*, Wiley.

Appendix A: Indicator Matrix

R Code

```
indicator.exe <- function(N) {

################################################################################################
#            Converts an m-way contingency table, N, into its indicator matrix form    #
################################################################################################

nvars <- dim(N)                                          # number of variables cross-classified to form N
ncats <- length(nvars)                      # number of categories
catnames <- NULL
for(l in 1:ncats) {                         # list of category names
     catnames <- c(catnames, dimnames(N)[[l]])
}
b <- c()
for(ii in 1:ncats) {
     aaap <- prod(nvars[(ii + 1):ncats])
     paaa <- prod(nvars[1:(ii - 1)])
     if(ii == 1) paaa <- 1
     if(ii == ncats) aaap <- 1
     b <- cbind(b, rep(rep(1:nvars[ii], rep(aaap, nvars[ii])), paaa))
}
bb <- matrix(0, dim(b)[1], sum(nvars))
for(jj in 1:dim(b)[1]) {
     eee <- c(b[jj, 1])
          for(ii in 2:length(nvars))
              eee <- c(eee, eval(parse(text = paste("b[jj,", ii, "] + sum(nvars[seq(1, ii-1)])", sep = ""))))
              bb[jj, eee] <- 1
          }
          Z <- c()
          for(jj in 1:dim(b)[1]) {
               aaaa <- c()
              for(ii in 1:dim(b)[2]) {
                   aaaa <- paste(aaaa, b[jj, ii], sep = ",")
              }
              bbbb <- eval(parse(text = paste("N[", substring(aaaa, 2), "]", sep = "")))
              if (bbbb > 0){
                   Z <- rbind(Z, matrix(rep(bb[jj,  ], bbbb), bbbb, byrow = T))
              }
          }
          dimnames(Z) <- list(paste("Individual", 1:sum(N)), paste(catnames))

          Z                  #             The indicator matrix of size n x ncats
}
```

Appendix B: Burt Matrix

R Code

```
burt.exe <- function(N) {

##################################################################################
#         Converts an m-way contingency table, N, into its Burt matrix form    #
##################################################################################

          Z <- indicator.exe(N)
          Burt <- as.matrix(t(Z) %*% Z)
          Burt
}
```

69