Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v16n2p423

Confidence Regions for Simple Correspondence Analysis using the Cressie-Read Family of Divergence Statistics

By Alzahrani, Beh and Stojanovski

14 October 2023

Electronic Journal of Applied Statistical Analysis
Vol. 16, Issue 2, 2023, 423-448
DOI: 10.1285/i20705948v16n2p423

Confidence Regions for Simple Correspondence Analysis using the Cressie-Read Family of Divergence Statistics

Asma A. Alzahrani[a,b], Eric J. Beh*[c,d], and Elizabeth Stojanovski[a]

[a]*School of Information and Physical Sciences, University of Newcastle, Australia*
[b]*Faculty of Science, Al-Baha University, Saudi Arabia*
[c]*National Institute for Applied Statistics Research Australia (NIASRA), University of Wollongong, Australia*
[d]*Centre for Multi-Dimensional Data Visualisation (MuViSU), Stellenbosch University, South Africa*

14 October 2023

When examining the association between symmetrically associated categorical variables, correspondence analysis provides a visual means of identifying the structure of this association. An important and sometimes overlooked feature that can help the analyst determine whether these categories provide a statistically significant contribution to the association is the confidence region. When constructing these regions, correspondence analysis traditionally (but not always) considers Pearson's chi-squared statistic as the core measure of association between the variables. Such a statistic is a special case of the Cressie-Read family of divergence statistics as is the log-likelihood ratio statistic, Freedman-Tukey statistic, and other such measures. Therefore, this paper will consider the construction of confidence regions in correspondence analysis where this family of divergence statistics is used as the measure of association. Doing so provides a means of simply constructing confidence regions for each category of a contingency table and allows for such regions to be constructed when log-ratio analysis (LRA) or the Hellinger distance decomposition (HDD) method is applied to the contingency table.

*Corresponding authors: ericb@uow.edu.au

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

keywords: Confidence circle, confidence ellipse, correspondence analysis, log-ratio analysis, eccentricity, semi-major axes.

1 Introduction

Correspondence analysis (CA) is a popular technique that can be used for visualising the association between two or more categorical variables that are cross-classified to form a contingency table. Such a visualisation assesses the nature of the association by depicting the position of a category's profile in a low-dimensional space, typically consisting of two dimensions. If it is found that a statistically significant association exists between the two variables, it is also important to understand how each row and column category of the table contributes to this dependence structure. To determine whether a row or column plays a statistically significant role in the association between the variables, confidence regions have been a topic of much discussion in the correspondence analysis literature. Lebart et al. (1984, pp. 182-186) proposed a simple and direct means of constructing a $100\,(1-\alpha)\%$ confidence circle for each category. Beh (2001) showed that the confidence circles of Lebart et al. (1984) are also applicable for the correspondence analysis involving ordered categorical variables (Beh, 1997) while Beh and D'Ambra (2009) demonstrated their use for non-symmetrical correspondence analysis. More computationally intensive approaches have been described by Ringrose (1992, 1996, 2012), Markus (1994), Linting et al. (2007), Gower et al. (1975), Greenacre (2017) and Lombardo and Ringrose (2012). See Beh and Lombardo (2014, Chapter 8) and Beh and Lombardo (2021, Section 2.9) for further details on the role of confidence regions in correspondence analysis.

While being simple to construct, the confidence circles described by Lebart et al. (1984), Beh (2001) and Beh and D'Ambra (2009) suffer from two drawbacks. The first is that the circular nature of the regions implies that the axes of a correspondence plot are identically weighted (they are not) and that they ignore the information of a category contained in the third and higher dimensions. An approach that overcame these drawbacks was proposed by Beh (2010) who developed a simple and direct means of constructing correspondence ellipses. The elliptical shape of these regions is reflected by the different weights given to the first two dimensions of a correspondence plot while the information contained in all dimensions can be incorporated into their construction.

One point that is worth highlighting is that, except for the regions described by Beh and D'Ambra (2009), the techniques of those described above use Pearson's chi-squared statistic (Pearson, 1904) as the core measure of association. However, this chi-squared statistic is a special case of the Cressie-Read family of divergence statistics (Cressie and Read, 1984) which depends on the parameter $\delta \in (-\infty, \infty)$. Some of the popular special cases of this divergence statistic include the log-likelihood ratio statistic ($\delta = 0$), the Freeman-Tukey statistic ($\delta = -1/2$), the modified log-likelihood ratio statistic ($\delta = -1$) and, of course, Pearson's statistic ($\delta = 1$). Recently, Beh and Lombardo (2023) briefly showed how the Cressie-Read divergence statistic can be used to perform a correspondence analysis on a two-way contingency table, thereby providing a more general framework under which the analysis can be performed. A far more comprehensive examination of this method and its features is given by Beh and Lombardo (2022). They demonstrated that two special cases of this framework include log-ratio analysis (LRA) (Greenacre, 2009, 2010b) and the Hellinger distance decomposition (HDD) (Cuadras and Cuadras, 2006, 2015; Cuadras et al., 2006) techniques. However, neither of these variants, or the approach of Beh and Lombardo (2022, 2023), considered the construction of confidence ellipses. It is therefore the purpose of this paper to demonstrate how the confidence ellipses of the type described by Beh (2010) can be adapted for a correspondence analysis that uses the Cressie-Read family of divergence statistics as the core measure of association between the categorical variables of a two-way contingency table. This shall be achieved in the following sections. Section 2 provides an overview of correspondence analysis using the family of divergence statistics while the construction of confidence ellipses is described in Section 3. Section 4 demonstrates the applicability of these regions by examining the association between the mother's attachment to her child, and the child's response to their mother's level of attachment. The data that we consider comes from the extensive study undertaken by van IJzendoorn (1995) and has been subsequently analysed in the correspondence analysis literature by, for example, Kroonenberg and Lombardo (1999), Beh (2010) and Ringrose (2012). In doing so, we describe how changes in δ impact on the various features of the confidence ellipses we consider for performing a correspondence analysis on a two-way contingency table. Such features include the eccentricity, semi-axis lengths, and area of the ellipses. Some final comments are made in Section 5.

2 The Cressie-Read Divergence Statistic

2.1 Correspondence Analysis and the Divergence Statistic

Consider an $r \times c$ two-way contingency table, $\mathbf{N}$, where the (i, j)'th cell entry is given by n_{ij} for $i = 1, 2, \ldots, r$ and $j = 1, 2, \ldots, c$. Denote the grand total of $\mathbf{N}$ by n and the (i, j)'th relative frequency by $p_{ij} = n_{ij}/n$. Define the i'th row relative marginal frequency by $p_{i\bullet} = \sum_{j=1}^{c} p_{ij}$ and the j'th column relative marginal frequency by $p_{\bullet j} = \sum_{i=1}^{r} p_{ij}$. To quantify whether there is a statistically significant association between the row and column variables, any number of chi-squared measures can be computed. Five such measures include Pearson's chi-squared statistic, X^2 (Pearson, 1904), the log-likelihood ratio statistic, G^2 (Wilks, 1938), the Freeman-Tukey statistic, T^2 (Freeman and Tukey, 1950), the modified chi-squared statistic, N^2 (Neyman, 1940, 1949) and the modified log-likelihood ratio statistic, M^2 (Kullback, 1959). All these statistics can be represented in terms of $p_{ij}/\left(p_{i\bullet}p_{\bullet j}\right)$ which is the *Pearson ratio* of the (i, j)'th cell of the contingency table. Each of these five statistics is a chi-squared random variables with $(r-1)(c-1)$ degrees of freedom. Moreover, these measures of association are special cases of the Cressie-Read family of divergence statistics (Cressie and Read, 1984) which is defined as

$$\mathrm{CR}^*(\delta) = \frac{2n}{\delta(\delta+1)} \sum_{i=1}^{r} \sum_{j=1}^{c} p_{ij} \left[\left(\frac{p_{ij}}{p_{i\bullet} p_{\bullet j}} \right)^{\delta} - 1 \right] \tag{1}$$

for $\delta \in (-\infty, \infty)$ such that $X^2 = \mathrm{CR}^*(1)$, $G^2 = \mathrm{CR}^*(0)$, $T^2 = \mathrm{CR}^*(-1/2)$, $N^2 = \mathrm{CR}^*(-2)$ and $M^2 = \mathrm{CR}^*(-1)$. For all values of δ, (1) is a chi-squared random variable with $(r-1)(c-1)$ degrees of freedom. One may also consider Cressie and Read (1989) and Cressie and Pardo (2002) for more details on (1) and its properties.

Despite the extensive family of chi-squared statistics that are members of (1), correspondence analysis largely considers $X^2 = \mathrm{CR}^*(1)$ as the measure of association between the variables of $\mathbf{N}$. There are exceptions including the variants LRA (Greenacre, 2009, 2010b) and the HDD technique of Cuadras and Cuadras (2006). While neither technique ties a chi-squared statistic to their method, Beh and Lombardo (2023) showed that LRA uses M^2 while HDD uses T^2. Beh et al. (2018) also demonstrate the link between the Freeman-Tukey statistic and the correspondence analysis of a two-way contingency table and so is a special case of the framework explored in detail by Beh and Lombardo (2023).

To perform a correspondence analysis on $\mathbf{N}$ using the Cressie-Read family of divergence statistics, Beh and Lombardo (2023) consider the second order Taylor series approximation of (1) around $\left(p_{ij}/\left(p_{i\bullet}p_{\bullet j}\right)\right)^{\delta} = 1$ which they define as

$$\mathrm{CR}^*(\delta) \approx \mathrm{CR}(\delta) = n \sum_{i=1}^{r} \sum_{j=1}^{c} p_{i\bullet} p_{\bullet j} \left[\frac{1}{\delta} \left(\left(\frac{p_{ij}}{p_{i\bullet} p_{\bullet j}} \right)^{\delta} - 1 \right) \right]^2 \tag{2}$$

where exact measures of some of the above chi-squared statistics can be obtained; for example, $X^2 = \mathrm{CR}^*(1) = \mathrm{CR}(1)$, $T^2 = \mathrm{CR}^*(-1/2)) = \mathrm{CR}(1/2)$ and $M^2 = \mathrm{CR}^*(-1) = \mathrm{CR}(0)$. Read and Cressie (1988, Section 6.6) show how (2) approximates (1). Another measure of association that we shall describe in Section 4 is the second order approximation to the Cressie-Read statistic which is obtained by substituting $\delta = 2/3$ into (2).

Beh and Lombardo (2023) show that the correspondence analysis of a two-way contingency table can be performed using (2) by first defining the $(i,\, j)$'th divergence residual such that

$$r_{ij}(\delta) = \frac{1}{\delta} \left(\left(\frac{p_{ij}}{p_{i\bullet} p_{\bullet j}} \right)^{\delta} - 1 \right). \tag{3}$$

Interestingly, this residual is not new. Bishop et al. (1975, Example 14.6.3) showed that (3) is a standard normal random variable. When $\delta \to 0$ the Box-Cox transformation can be applied leading to the residual

$$r_{ij}(0) = \ln \left(\frac{p_{ij}}{p_{i\bullet} p_{\bullet j}} \right).$$

Additionally, Anscombe (1953, pp. 229-230) and McCullagh and Nelder (1984, p. 38) considered (3) when $\delta = 2/3$ for "better" normalising a random variable by minimising the skewness of the data at the expense of a non-constant variance which can be achieved when $\delta = 1/2$ (Freeman and Tukey, 1950).

In the correspondence analysis literature, the association between the variables of $\mathbf{N}$ are quantified using Pearson's phi-squared statistic and is commonly referred to as the total inertia of the contingency table. When using (2), this is defined as

$$\phi(\delta) = \frac{\mathrm{CR}(\delta)}{n} = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{i\bullet}p_{\bullet j}r_{ij}^{2}(\delta)\ .$$

The components necessary to obtain a visual display of the association can then be determined by applying a generalised singular value decomposition (GSVD) to the matrix of $r_{ij}(\delta)$ values such that

$$r_{ij}(\delta) = \sum_{m=1}^{M^*} a_{im}(\delta)\,\lambda_m(\delta)\,b_{jm}(\delta)$$

where

$$\lambda_m(\delta) = \sum_{i=1}^{r}\sum_{j=1}^{c} p_{ij}a_{im}(\delta)\,b_{jm}(\delta)$$

is the m'th singular value of the matrix of $r_{ij}(\delta)$ values which are arranged in descending order such that $1 > \lambda_1(\delta) > \lambda_2(\delta) > \cdots > \lambda_{M^*}(\delta) > 0$. The square of these singular values gives the principal inertia associated with the m'th dimension of the correspondence plot. The optimal correspondence plot consists of M^* dimensions; here $M^* = \min(r, c) - 1$ for all δ except $\delta = 0,\ 1/2$ where $M^* = \min(r, c)$. From the GSVD of (3), the components $a_{im}(\delta)$ and $b_{jm}(\delta)$ have the property

$$\sum_{i=1}^{r} p_{i\bullet}a_{im}(\delta)\,a_{im'}(\delta) = \begin{cases} 1 & m = m' \\ 0 & m \neq m' \end{cases} \qquad (4)$$

$$\sum_{j=1}^{c} p_{\bullet j}b_{jm}(\delta)\,b_{jm'}(\delta) = \begin{cases} 1 & m = m' \\ 0 & m \neq m' \end{cases}\ . \qquad (5)$$

A visual summary of the association between the rows and columns of $\mathbf{N}$ can then be obtained by jointly depicting along the m'th dimension

$$f_{im}(\delta) = a_{im}(\delta)\,\lambda_m(\delta) \qquad g_{jm}(\delta) = b_{jm}(\delta)\,\lambda_m(\delta)$$

which is the principal coordinate for the i'th row and j'th column, respectively. From (4) and (5), the total inertia of $\mathbf{N}$ can be expressed in terms of the weighted sum-of-squares of these coordinates such that

$$\sum_{i=1}^{r} p_{i\bullet}f_{im}(\delta)\,f_{im'}(\delta) = \begin{cases} \lambda_m^2(\delta) & m = m' \\ 0 & m \neq m' \end{cases}$$

$$\sum_{j=1}^{c} p_{\bullet j}b_{jm}(\delta)\,b_{jm'}(\delta) = \begin{cases} \lambda_m^2(\delta) & m = m' \\ 0 & m \neq m' \end{cases}\ .$$

Therefore, the total inertia can be expressed in terms of the squared $\lambda_m(\delta)$ values or in terms of the principal coordinates such that

$$\phi(\delta) = \sum_{m=1}^{M^*} \lambda_m^2(\delta) = \sum_{i=1}^{r} p_{i\bullet} f_{im}^2(\delta) = \sum_{j=1}^{c} p_{\bullet j} g_{jm}^2(\delta) \ .$$

2.2 $100(1-\alpha)\%$ Confidence Regions and X^2

In situations where there is a statistically significant association between the variables of a contingency table, correspondence analysis provides a visual means of summarising this association; either a biplot (Greenacre, 2010a; Gower et al., 1975; Beh and Lombardo, 2014) or the more traditional correspondence (aka "symmetric" or "French") plot can be constructed, but we shall confine our attention to the latter of the two. Confidence regions provide additional insight into the structure of the association by allowing the analyst to identify those categories that contribute to this association and those that do not. We provide here an overview of $100(1-\alpha)\%$ confidence circles and confidence ellipses of the type described by Lebart et al. (1984) and Beh (2010) who use Pearson's statistic as the measure of association. We shall then discuss the construction of $100(1-\alpha)\%$ confidence ellipses using (2).

For the two-way contingency table, $\mathbf{N}$, Lebart et al. (1984, p. 183) showed that the radii length of the $100(1-\alpha)\%$ confidence circle for the i'th category in a two-dimensional correspondence plot is

$$r_{i(\alpha)} = \sqrt{\frac{\chi_\alpha^2}{np_{i\bullet}}}$$

where χ_α^2 is the $1-\alpha$ percentile of a chi-squared distribution with 2 degrees of freedom. Note that such lengths are generated where Pearson's chi-squared statistic is used as the measure of association between the variables such that the total inertia is equivalent to $\phi(1) = \mathrm{CR}(1)/n = X^2/n$.

The confidence circles of Lebart et al. (1984) have the disadvantage that they do not consider the unequal weighting of the axes of a correspondence plot; this is apparent since the first singular value is larger than the second singular value. The $100(1-\alpha)\%$ confidence circles generated from these radii lengths also do not take into account the information reflected in the higher dimensions of a correspondence plot. To overcome these features, Beh (2010) proposed a simple method of constructing confidence ellipses and, like the circles of Lebart et al. (1984), considered Pearson's chi-squared statistic as the core measure of association between the variables. Using the notation adopted above, Beh (2010) showed that the semi-major and semi-minor axis lengths of the $100(1-\alpha)\%$

confidence ellipse for the i'th row category is

$$x_{i(\alpha)} = \lambda_1(1)\sqrt{\frac{\chi^2_\alpha}{X^2}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}(1)\right)} \quad (6)$$

$$y_{i(\alpha)} = \lambda_2(1)\sqrt{\frac{\chi^2_\alpha}{X^2}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}(1)\right)} \quad (7)$$

respectively, so that $x_{i(\alpha)} > y_{i(\alpha)}$ since $\lambda_1(1) > \lambda_2(1)$. Therefore, Beh's (2010) confidence ellipses are structured so that the semi-major axis is parallel with the first principal axis and the semi-minor axis is parallel with the second principal axis. They are equivalent to Lebart et. al's (1984) circular regions only when the information in the first two axes is taken into consideration and $\lambda_1(1) = \lambda_2(1)$. More computationally intensive procedures for generating confidence ellipses/hulls have been proposed in the past and Ringrose (2012, Section 5.2) compares his approach with Beh's (2010) indicating "the two sets of ellipses are based on quite different assumptions and objectives, so too much should not be read into any similarities or differences in a particular case" (page 1412).

Since the construction of confidence regions with a semi-axis length given by (6) and (7) is undertaken when Pearson's chi-squared statistic is used to assess the statistical significance between the categorical variables of a two-way contingency table, this suggests that a more general framework can be developed where (2) is used instead. We shall now turn our attention to how this can be achieved.

3 On the Construction of Elliptical Regions

3.1 The Semi- Axis Lengths

Suppose we consider performing a correspondence analysis on a two-way contingency table where the second order approximation of the Cressie-Read family of divergence statistics given by (2) is used as the measure of association between the categorical variables. Then the semi-major axis length of the $100(1-\alpha)\%$ confidence ellipse is

$$x_{i(\alpha)}(\delta) = \lambda_1(\delta)\sqrt{\frac{\chi^2_\alpha}{\mathrm{CR}(\delta)}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}(\delta)\right)} \quad (8)$$

while its semi-minor axis length is

$$y_{i(\alpha)}(\delta) = \lambda_2(\delta)\sqrt{\frac{\chi^2_\alpha}{\mathrm{CR}(\delta)}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}(\delta)\right)}. \quad (9)$$

The derivation of these lengths is given in the Appendix and follows the same lines as the derivation that Beh (2010) gave for (6) and (7).

The semi- axis lengths of (8) and (9) are constructed to reflect not just the information contained in the first two dimensions of the correspondence plot, but to reflect the information in all optimal (M^*) dimensions. They can be alternatively expressed in terms of the row profile coordinates and principal inertia values, $\lambda_m^2(\delta)$, of the third and higher dimensions by

$$\begin{aligned} x_{i(\alpha)} &= \lambda_1(\delta)\sqrt{\frac{\chi^2_\alpha}{\mathrm{CR}\,(\delta)}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*}\frac{f^2_{im}\,(\delta)}{\lambda^2_m\,(\delta)}\right)} \\ y_{i(\alpha)} &= \lambda_2(\delta)\sqrt{\frac{\chi^2_\alpha}{\mathrm{CR}\,(\delta)}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*}\frac{f^2_{im}\,(\delta)}{\lambda^2_m\,(\delta)}\right)}. \end{aligned}$$

3.2 Special Cases

Suppose we consider the case where $\delta = 1$. Then, since $\mathrm{CR}(1) = X^2$ is Pearson's chi-squared statistic, (8) and (9) simplify to (6) and (7), respectively. Therefore, the confidence ellipses presented in Beh (2010) are a special case of the confidence ellipses constructed using (8) and (9) as the semi axis lengths.

Suppose we now consider the case where $\delta = 1/2$. Then the semi-major and semi-minor axis lengths are

$$x_{i(\alpha)}\left(\frac{1}{2}\right) = \lambda_1\left(\frac{1}{2}\right)\sqrt{\frac{\chi^2_\alpha}{T^2}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}\left(\frac{1}{2}\right)\right)} \tag{10}$$

$$y_{i(\alpha)}\left(\frac{1}{2}\right) = \lambda_2\left(\frac{1}{2}\right)\sqrt{\frac{\chi^2_\alpha}{T^2}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}\left(\frac{1}{2}\right)\right)} \tag{11}$$

respectively. These axis lengths are for a correspondence analysis where the Freeman-Tukey statistic is used to assess the statistical significance of the two variables. They are therefore appropriate to use with the HDD variant described by Cuadras and Cuadras (2006, 2015) and Cuadras et al. (2006) and the approach of Beh et al. (2018).

Confidence ellipses can also be constructed for Greenacre's (2009, 2010b) LRA variant of correspondence analysis. Since Beh and Lombardo (2023) showed that this variant uses the modified log-likelihood ratio statistic, $M^2 = \mathrm{CR}(0)$, then $100(1-\alpha)\%$ confidence ellipses can be constructed for LRA where the semi-axis lengths are

$$x_{i(\alpha)}\,(0) = \lambda_1\,(0)\sqrt{\frac{\chi^2_\alpha}{M^2}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}\,(0)\right)} \tag{12}$$

$$y_{i(\alpha)}\,(0) = \lambda_2\,(0)\sqrt{\frac{\chi^2_\alpha}{M^2}\left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}\,(0)\right)}. \tag{13}$$

Confidence ellipses can also be constructed for other variants of correspondence analysis where (2) is the measure used to assess the association structure of the variables although we shall confine our attention to the special cases described above.

3.3 Eccentricity of the Ellipses

Eccentricity is an important feature of any ellipse. It takes on a value that is bounded by 0 and 1 where a perfect circle has an eccentricity of zero while an ellipse with $\lambda_1(\delta) \gg \lambda_2(\delta)$ will be increasingly dominated by the first principal axis and have an eccentricity approaching 1. Suppose we consider the i'th row category of $\mathbf{N}$. Then the eccentricity of its $100(1-\alpha)\%$ confidence ellipse with a semi-major and semi-minor axis length defined by (8) and (9) is

$$E_{i(\alpha)}(\delta) = \sqrt{1 - \frac{\lambda_2^2(\delta)}{\lambda_1^2(\delta)}} \tag{14}$$

where the ratio under the square root sign is just the ratio of the principal inertia values for the first two dimensions of the correspondence plot. Since the principal inertia values are arranged in descending order (a feature evident by knowing the singular values are arranged in the same way) this confirms that the eccentricity calculated from (14) will range between 0 and 1. Note that a perfectly elliptical confidence region will only ever be observed when the second principal inertia value, $\lambda_2^2(\delta)$, is equal to zero; this very rarely happens. Also, since it very rarely occurs that the first two principal inertia values will be identical when analysing a two-way contingency table the confidence region will rarely be perfectly circular either.

It can also be shown that (14) can be assessed in terms of the semi- axis lengths such that

$$E_{i(\alpha)}(\delta) = \sqrt{1 - \left(\frac{x_{i(\alpha)}(\delta)}{y_{i(\alpha)}(\delta)}\right)^2}$$

so that the semi-major and semi-minor axis lengths are related to each other through the eccentricity of the ellipse.

This measure, and that of (14), is also the eccentricity of the $100(1-\alpha)\%$ confidence ellipse for the j'th column category of the two-way contingency table, $E_{j(\alpha)}(\delta)$. When $\delta = 1$, we obtain the same eccentricity measure of Beh (2010, eq (8)) who showed the eccentricity of his ellipses in terms of Pearson's chi-squared statistic. Like the ellipses of Beh (2010), the eccentricity of the ellipses is the same for all row and column categories. However, in the case of (14), the eccentricity changes only for changes in the value of δ that is adopted to define the chi-squared statistic used to numerically, and visually, assess the association structure in the correspondence analysis of $\mathbf{N}$.

3.4 Area of the Ellipses

Another important feature of the confidence ellipse is its area. Suppose we again consider the i'th row category. Then, given the semi- axis lengths of (8) and (9), the area of its $100(1-\alpha)\%$ confidence ellipse is

$$\begin{aligned} A_{i(\alpha)}(\delta) &= \pi x_{i(\alpha)}(\delta)\, y_{i(\alpha)}(\delta) \\ &= \pi\chi^2_\alpha \frac{\lambda_1(\delta)\,\lambda_2(\delta)}{\mathrm{CR}(\delta)} \left(\frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a^2_{im}(\delta)\right). \end{aligned} \tag{15}$$

3.5 Link with Confidence Circles

Suppose we now consider the case where the first two dimensions of a correspondence plot provide a near comprehensive visual summary of the association so that the third and higher dimensions contribute very little to the association structure. In this case, the eccentricity of the row and column categories is still calculated using (14) since this measure considers the principal inertia values associated with the first two dimensions of the optimal correspondence plot. However, if it is assumed, or known (in rare cases), that $\lambda_1^2(\delta) = \lambda_2^2(\delta)$ then (14) simplifies to zero showing that equality of the first two principal inertia values will coincide with confidence regions that are circular in shape.

When only the first two principal axes are used to construct the confidence regions then the semi-major and semi-minor axis lengths are

$$x_{i(\alpha)} = \lambda_1(\delta)\sqrt{\frac{\chi^2_\alpha}{p_{i\bullet}\mathrm{CR}(\delta)}} = \lambda_1(\delta)\sqrt{\frac{\chi^2_\alpha}{np_{i\bullet}\phi(\delta)}}$$

$$y_{i(\alpha)} = \lambda_2(\delta)\sqrt{\frac{\chi^2_\alpha}{p_{i\bullet}\mathrm{CR}(\delta)}} = \lambda_2(\delta)\sqrt{\frac{\chi^2_\alpha}{np_{i\bullet}\phi(\delta)}}$$

respectively. These results can be obtained from (8) and (9) when $M^* = 2$. Therefore, the sum-of-squares of these lengths gives

$$x^2_{i(\alpha)}(\delta) + y^2_{i(\alpha)}(\delta) = \frac{\left(\lambda_1^2(\delta) + \lambda_2^2(\delta)\right)}{\phi(\delta)}\left(\frac{\chi^2_\alpha}{np_{i\bullet}}\right)$$

where $100 \times \left(\lambda_1^2(\delta) + \lambda_2^2(\delta)\right)/\phi(\delta)$ is the percentage of the total inertia that is reflected in the first two dimensions of the optimal correspondence plot. When the optimal correspondence plot consists of only two dimensions so that $\left(\lambda_1^2(\delta) + \lambda_2^2(\delta)\right)/\phi(\delta) = 1$ then

$$x^2_{i(\alpha)}(\delta) + y^2_{i(\alpha)}(\delta) = r^2_{i(\alpha)}(\delta)$$

where $r_{i(\alpha)}(\delta) \equiv r_{i(\alpha)}$ is just the radii length of the confidence circles described by Lebart et al. (1984). Interestingly, in this case, the value of δ has no impact on the radii length of confidence circles when a correspondence analysis is performed using the Cressie-Read family of divergence statistics.

If $0 < \left(\lambda_1^2(\delta) + \lambda_2^2(\delta)\right)/\phi(\delta) < 1$ (as will be the case in most situations) ignoring the association that is depicted in the third and higher dimensions will increase the magnitude of the semi- axis lengths and the area of an ellipse, however the eccentricity will remain unchanged. In this case, there is the possibility that an ellipse that does not overlap the origin will do so leading to a spurious conclusion about the contribution of that category to the association structure between the variables.

4 Application

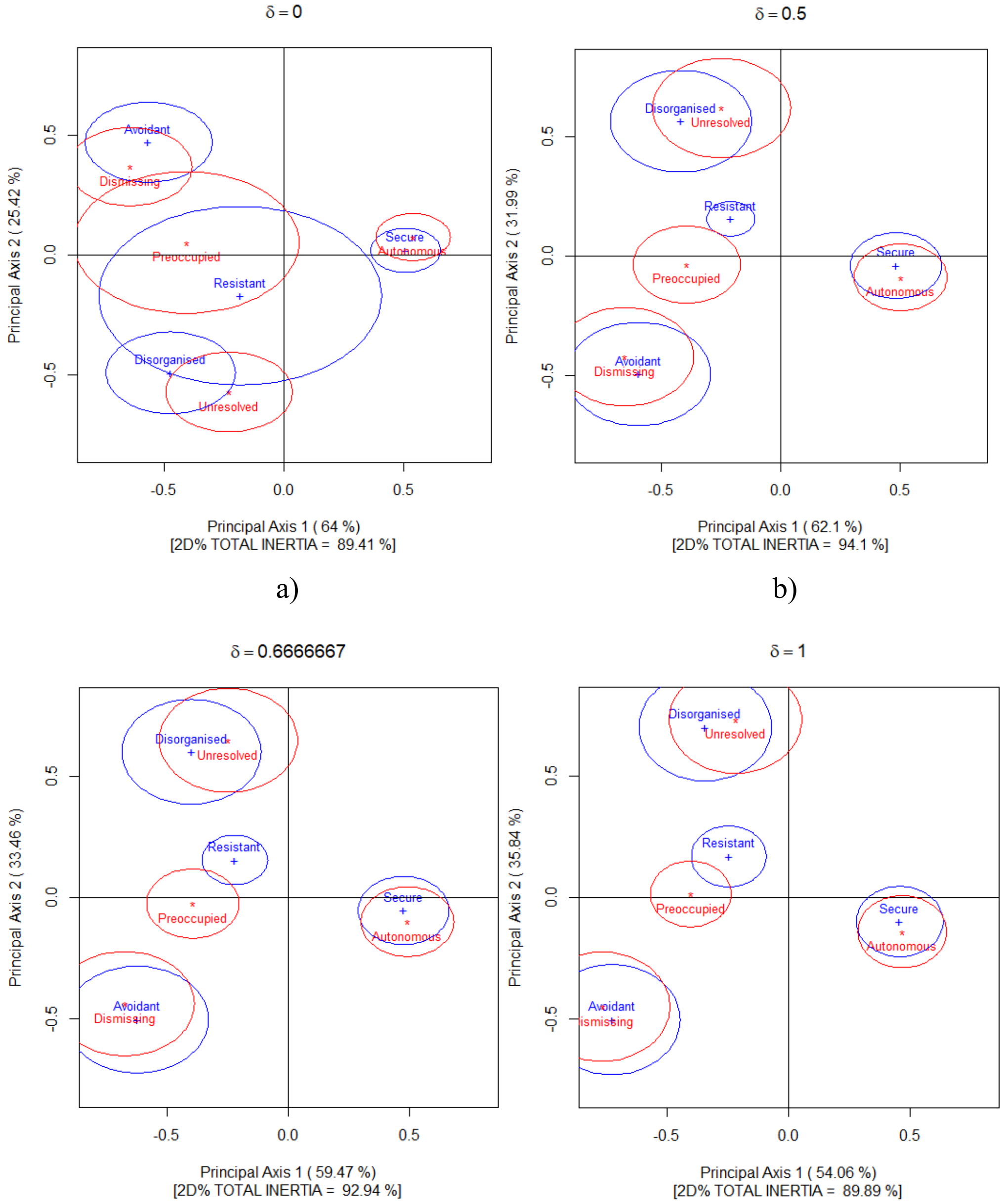
4.1 A Numerical Inspection of the Association

Consider the 4×4 contingency table of Table 1 that summarises the cross-classification of 548 mothers by their attachment with their parents and their child's attachment to them. The data are based on an extensive study of mother–child attachment conducted by van IJzendoorn (1995). The four column categories are a result of the adult attachment interview (George et al., 1985), while the four row categories are observed from the Ainsworth strange situation (Ainsworth et al., 1978). Table 1 was analysed using non-symmetrical correspondence analysis by Kroonenberg and Lombardo (1999). It was also considered by Beh (2010) and Ringrose (2012) in their derivation of elliptical confidence regions where a symmetric association between the two variables was assumed.

Table 1: Cross-classification of the attachment classification of a mother and her infant Mother's Attachment Classification

Infant	*Mother's Attachment Classification*				
Response	Dismissing	Autonomous	Preoccupied	Unresolved	Total
Avoidant	62	29	14	11	116
Secure	24	210	14	39	287
Resistant	3	9	10	6	28
Disorganised	19	26	10	62	117
Total	108	274	48	118	548

Pearson's chi-squared test of independence of Table 1 shows that there exists a statistically significant association between the two variables ($X^2 = 252.40$, p-value < 0.0001); the statistical significance of this association can also be verified using, for example, the log-likelihood ratio statistic, its modified version, the Freeman-Tukey statistic or the Cressie-Read statistic.



a)

b)

c)

d)

Figure 1: 95% confidence ellipses for correspondence analysis of Table 1 using the Cressie-Read divergence statistic with a) $\delta = 0$, b) $\delta = 1/2$ c) $\delta = 2/3$ and d) $\delta = 1$

4.2 A Visual Inspection of the Association

The nature of this association can be investigated further by applying a correspondence analysis on Table 1. Typically, such an analysis can be performed by using the technique described above for $\delta = 1$. However, a more general framework can be performed by applying a GSVD to the matrix of divergence residuals of Table 1 for a given δ. Figure 1 shows the two-dimensional correspondence plot when the modified log-likelihood statistic ($\delta = 0$), Freeman–Tukey statistic ($\delta = 1/2$), a second order approximation of the Cressie-Read statistic ($\delta = 2/3$) and Pearson's statistic ($\delta = 1$) are used to assess the nature of the association. Identifying those categories that play a statistically significant role in the structure of this association can be achieved by constructing a 95% confidence ellipse for each category of Table 1 and these are superimposed upon each of the plots in Figure 1; more will be said on the interpretation of these ellipses shortly but we point out the scale of each axis in the four plots is identical, ranging from -0.8 to $+0.8$. When $\delta = 0$ these regions are applicable for Greenacre's (2009, 2010b) LRA while confidence ellipses for the HDD method of Cuadras and Cuadras (2006) is generated when $\delta = 1/2$. As we have described above, the confidence ellipses for the classical approach to correspondence analysis is obtained when $\delta = 1$. In all four cases, the semi- axis lengths of (8) and (9) are used so that information in the third dimension of the correspondence plot is reflected in the two-dimensional correspondence plot.

Figure 1 shows that all four analyses provide a very effective and informative visual display of the association, depicting at least 89% of the total inertia calculated for each δ considered. When the modified log-likelihood ratio statistic is used, Figure 1a) displays 89.4% of the association while Figure 1b), which uses the Freeman-Tukey statistic, accounts for 94.1% of the association. The Pearson and second-order approximation to the Cressie-Read statistic visualise 89.89% and 92.94% of the association using the first two dimensions, respectively. Thus, the Freeman-Tukey statistics provides the best quality visualisation of the association between the two variables of Table 1 by maximising the percentage of the inertia explained by the first two dimensions. Using this statistic also has the advantage of being able to stabilise any over-dispersion that exists in the contingency table (Beh et al., 2018) and has long been advocated in the correspondence analysis literature (Domenges and Volle, 1979; Rao, 1995; Cuadras and Cuadras, 2006). Another reason for their preference is that distances of row points, say, in the correspondence plot are not influenced by the column marginal information.

4.3 Interpretation of the Ellipses

To help identify the important contributors to this association, consider the 95% confidence ellipses are superimposed upon each of the four plots of Figure 1. A region is interpreted such that if it overlaps the origin, then its category does not have a statistically significant impact on the association between the two variables. For ellipses that do not overlap the origin, their category does have an impact. Figure 1a), which is produced using $\delta = 0$ and is an analysis that is equivalent to Greenacre's (2009, 2010b) LRA, shows that all row and column categories, except "Preoccupied" and "Resistant" contribute to the association structure between the variables. For the remaining three values of δ, all confidence regions do not overlap the origin and so play a statistically significant role in defining the association between the variables of Table 1. On the other hand, when $\delta = 1/2$, $2/3$ and 1, all categories play a statistically significant role in defining how the association is structured since all of the ellipses do not overlap with the origin. The four plots in Figure 1 take into consideration all of the information contained in the optimal number of dimensions, not just the first two as displayed.

While the 95% confidence ellipses appear quite similar for the four plots of Figure 1, there are some key differences in the semi-major axis lengths and areas which we shall discuss. First though, we shall comment on their eccentricity and how they change as δ changes value.

4.4 On the Inertia and Eccentricity

The shape and size of the ellipses between the four plots does differ in places indicating that changes in δ have impacted on the semi-axis lengths and, therefore, the eccentricity and area of the ellipses. For example, consider Table 2. It summarises the $\mathrm{CR}(\delta)$ value, given by (2). It also summarises the first two principal inertia values, $\lambda_1^2(\delta)$ and $\lambda_2^2(\delta)$, and their percentage contribution to $\mathrm{CR}(\delta)$, for $\delta = 0$, $1/2$, $2/3$ and 1. When $\delta = 0$, Table 2 shows that the first principal inertia is the largest of the four δ while the second principal inertia value is the smallest. This suggests that the confidence ellipses for $\delta = 0$ is more elliptical than for any of the three other values considered; hence the LRA of Table 1 produces a more elliptical confidence region for each of the categories of Table 1 than the HDD method ($\delta = 1/2$) or the traditional approach to correspondence analysis ($\delta = 1$). This concurs with the finding that the smallest difference between these two inertia values occurs when $\delta = 1$. Such a result suggests that, of the four variants of correspondence analysis considered here, the traditional approach yields a more circular confidence ellipse than the three other variants considered. Figure 2 summarises the eccentricity of the ellipses for $\delta \in [0, 2]$ and shows that it decreases steadily from 0.7764 at $\delta = 0$ to 0.4892 at $\delta = 1.57$ and then slowly increases to 0.5298 at $\delta = 2$. When one

Table 2: First two principal inertia values, their percentage contribution to the total inertia values and $\mathrm{CR}(\delta)$ for $\delta = 0$, $1/2$, $2/3$ and $\delta = 1$ of Table 1

Summary	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$
$\lambda_1^2(\delta)$	0.2518	0.2361	0.2368	0.2490
$\lambda_1^2(\delta)$	0.1000	0.1216	0.1332	0.1651
% Cont.	89.41	94.10	92.94	89.89
$\mathrm{CR}(\delta)$	215.59	208.94	218.16	252.40

considers larger values of δ (not shown here), the eccentricity of the regions is almost perfectly elliptical (with an eccentricity of nearly 1) for $\delta > 6$ since $\lambda_2^2(\delta) \to 0$ as δ increases beyond 1.5.

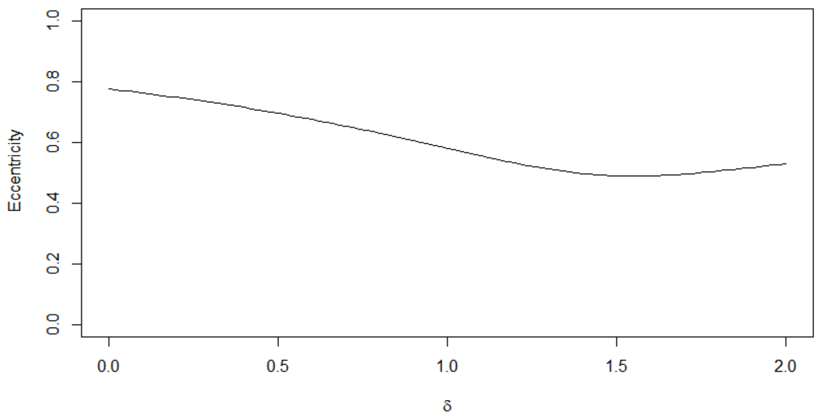


Figure 2: The eccentricity of the 95% confidence ellipses for $\delta \in [0,\ 2]$

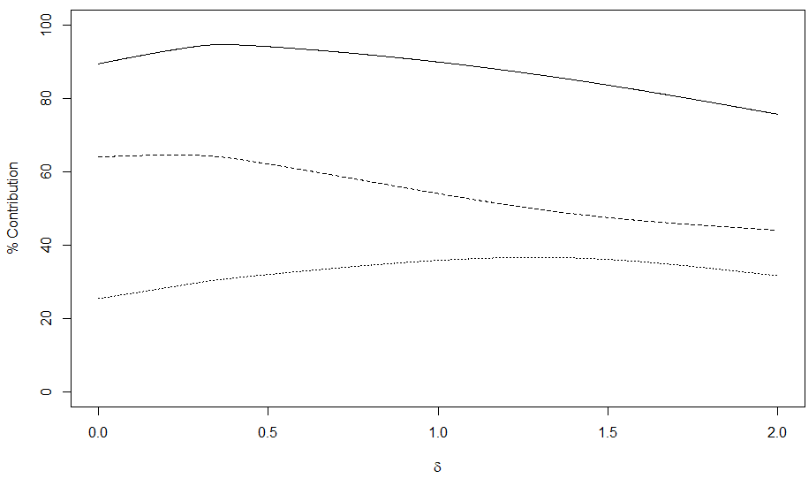


Figure 3: Percentage contribution of the first (dashed line) and second (dotted line) to the total inertia and the quality of the two-dimensional plot (solid line) for $\delta \in [0,\ 2]$

We can also view the impact of the first two principal inertia values on the eccentricity of the ellipses by observing Figure 3. It shows the percentage contribution of both dimensions to the total inertia and the quality of the two-dimensional correspondence plot for $\delta \in [0, 2]$. Figure 3 reveals that for the δ values that lie in this interval, $\lambda_1^2(\delta)$ is at its largest compared to $\lambda_2^2(\delta)$ when $\delta = 0$ highlighting that the eccentricity of the ellipses at this value of δ is at its maximum. For values of δ that lie between 0 and 2, these two principal inertia values are closest when $\delta = 1.57$ highlighting that this is the value of δ where the regions are the most circular.

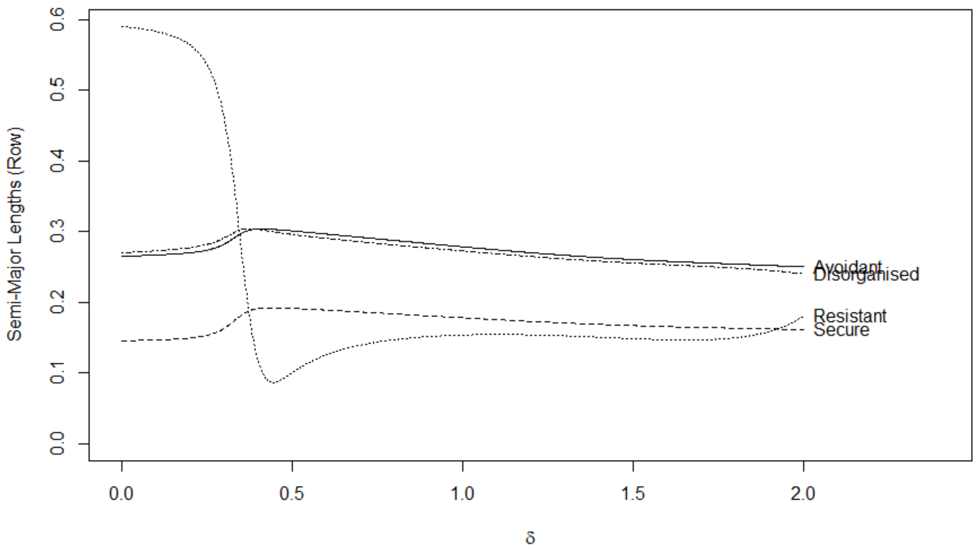
These shifts in the principal inertia values of Table 1 account for the shifts in the semi-major and semi-minor axis lengths for the row and column confidence ellipses and their eccentricity. They also impact upon the area of each of these ellipses and so we shall now turn our attention this feature.

4.5 On the Semi- Axis Lengths

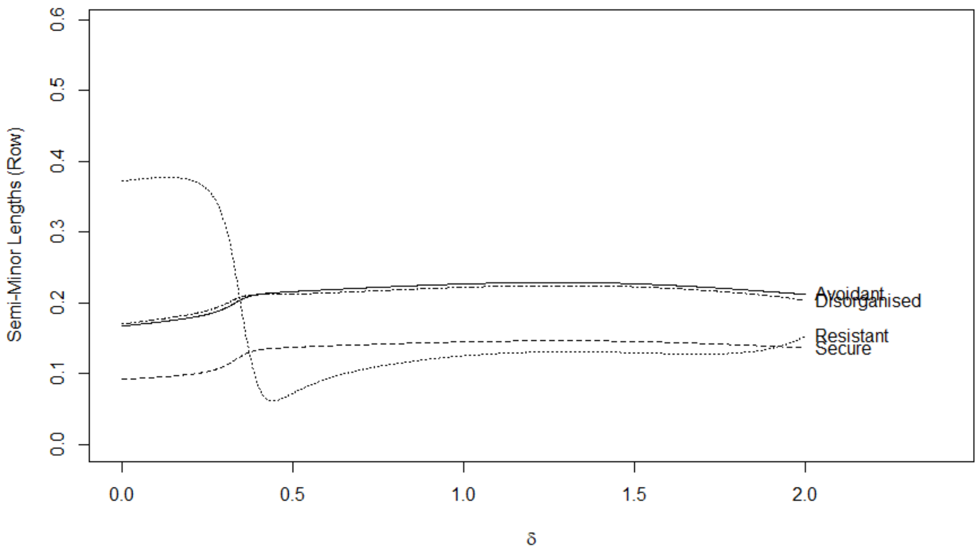
The shift in the principal inertia values not only impacts on the eccentricity of the confidence ellipses but also on their semi-major and semi-minor axis lengths. Table 3 summarises these lengths for each row category of Table 1 for our four values of δ while Table 4 summarises these lengths for the column categories. Suppose we consider first the values summarised in Table 3. They show that the semi-major axis lengths for each of the row categories, while different between the categories, remains quite stable as δ increases from 0 to 1; although there is one exception. The semi-major axis length for "Resistant" is 0.5897 when $\delta = 0$ but then stabilises to around 0.13 for the remaining three values of δ. The row categories "Avoidant", "Secure" and "Disorganised" are relatively stable with a semi-major axis length centred at about 0.28, 0.18 and 0.28 respectively. There are similar disparities in magnitude of the semi-minor axis lengths. Again, the greatest variation in length is for the row category "Resistant"; when $\delta = 0$ its length is 0.3716 but then stabilises to around 0.10 for the other values of δ.

A more comprehensive overview of the changes in semi-axis lengths can be seen by examining Figure 4 and Figure 5. Figure 4a) visualises the semi-major axis length of the row categories of Table 1 for $\delta \in [0, 2]$ while the semi-minor axis lengths are depicted by Figure 4b). It shows that, across this range of δ values, the semi- axis lengths of the rows are quite stable except for "Resistant" which is relatively unstable for δ values ranging from 0 to about 0.6. Interestingly, while Figure 4 shows a big drop in semi-major and semi-minor axis lengths for this category around $\delta \in (0.2, 0.4)$, these lengths for the remaining row categories experience a moderate jump and then plateau in the region $\delta \in (0.4, 2)$. A similar behaviour of the semi- axis lengths for the column categories can be seen from Figure 5.

Suppose we now consider Table 4. It shows that the semi-major axis lengths remain quite stable for each of the column categories except for "Preoccupied". When $\delta = 0$ its length is 0.4687 and decreases to 0.1661 when $\delta = 1$. A similar observation is also made for this category's semi-minor axis length which is 0.2953 when $\delta = 0$ and drops to 0.1352 when $\delta = 1$.

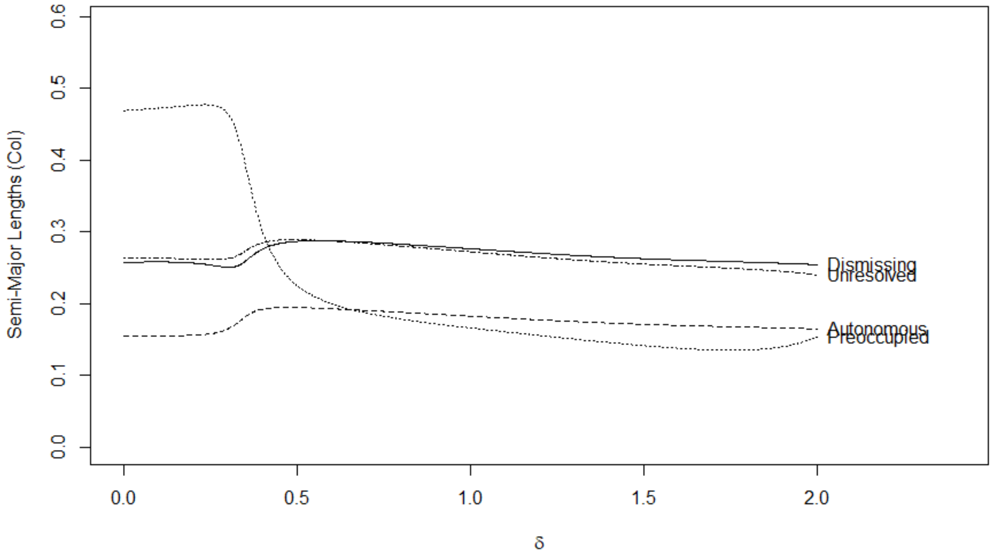


a)

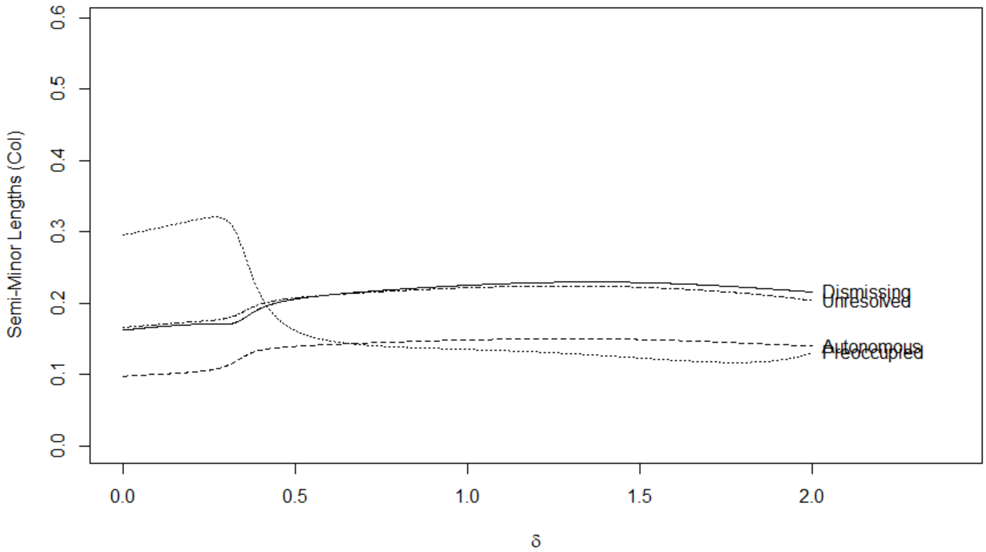


b)

Figure 4: The a) semi-major and b) semi-minor axis lengths for each of the row categories of Table 1 for $\delta \in [0,\, 2]$



a)



b)

Figure 5: The a) semi-major and b) semi-minor axis lengths for each of the column categories of Table 1 for $\delta \in [0, 2]$

Table 3: Semi-major ($x_{i(0.05)}(\delta)$) and a semi minor ($y_{i(0.05)}(\delta)$) axis lengths for the 95% confidence ellipse of the row categories of Table 1; $\delta = 0, 1/2$, 2/3 and $\delta = 1$

Row	$x_{i(0.5)}(\delta)$				$y_{i(0.05)}(\delta)$			
Categories	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$
Avoidant	0.2651	0.3008	0.2933	0.2784	0.1670	0.2159	0.2200	0.2267
Secure	0.1454	0.1913	0.1871	0.1780	0.0917	0.1373	0.1403	0.1449
Resistant	0.5897	0.1007	0.1357	0.1537	0.3716	0.0723	0.1018	0.1252
Disorganised	0.2699	0.2958	0.2871	0.2726	0.1701	0.2123	0.2154	0.2219

Table 4: Semi-major ($x_{j(0.05)}(\delta)$) and a semi minor ($y_{j(0.05)}(\delta)$) axis lengths for the 95% confidence ellipse of the column categories of Table 1; $\delta = 0, 1/2$, 2/3 and $\delta = 1$

Column	$x_{j(0.05)}(\delta)$				$y_{j(0.05)}(\delta)$			
Categories	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$
Dismissing	0.2573	0.2866	0.2861	0.2764	0.1621	0.2057	0.2146	0.2251
Autonomous	0.1552	0.1945	0.1912	0.1824	0.0978	0.1396	0.1434	0.1485
Preoccupied	0.4687	0.2249	0.1906	0.1661	0.2953	0.1614	0.1430	0.1352
Unresolved	0.2635	0.2889	0.2850	0.2722	0.1661	0.2074	0.2138	0.2216

4.6 On the Area of the Confidence Ellipses

We now turn our attention to investigating the changes in the area of the 95% confidence ellipses for $\delta \in [0, 2]$. The area for the row ellipses is defined using (15) while the column ellipse areas can be calculated in a similar manner. Table 5 summarises these values for the row categories while the area of the ellipses for the column categories appears in Table 6. Tables 5 and 6 show that the area of the ellipse for the row category "Resistant" and column category "Preoccupied", respectively, is very large for $\delta = 0$ than for any of the three other values. This is not surprising given the large size of their ellipse in Figure 1 and their relatively large semi-axis lengths described in Section 4.5.

Figure 6 provides a visual inspection of the area of the 95% confidence ellipses for the row categories while a similar inspection can be made of the area of the ellipses for the column categories (see Figure 7); for both figures, values of δ that lie in the interval [0, 2] are considered. They show that the area of the ellipse for the row category "Resistant" and the column category "Preoccupied" are very large in comparison to the other categories for δ values that fall within the interval [0, 0.3] and then drop

Table 5: Area of the 95% confidence ellipse for each of the row categories of Table 1 the 95% confidence ellipse of the row categories of Table 1; $\delta = 0, 1/2$, $2/3$ and $\delta = 1$

Row	*Area*			
Categories	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$
Avoidant	0.1391	0.2040	0.2028	0.1983
Secure	0.0419	0.0825	0.0825	0.0810
Resistant	0.6884	0.0229	0.0434	0.0604
Disorganised	0.1443	0.1973	0.1943	0.1900

Table 6: Area of the 95% confidence ellipse for each of the column categories of Table 1 the 95% confidence ellipse of the row categories of Table 1; $\delta = 0, 1/2$, $2/3$ and $\delta = 1$

Column	*Area*			
Categories	$\delta = 0$	$\delta = 1/2$	$\delta = 2/3$	$\delta = 1$
Dismissing	0.1311	0.1853	0.1930	0.1954
Autonomous	0.0477	0.0853	0.0862	0.0851
Preoccupied	0.4349	0.1141	0.0856	0.0706
Unresolved	0.1375	0.1883	0.1914	0.1895

quickly for δ between 0.4 and 0.6 and remain the smallest ellipse for all δ values up to 2. This suggests that while LRA, and similar analyses where δ is small (but not necessarily tiny) in size lead to "Resistant" and "Preoccupied" not being statistically significant contributors to the association that exists between the variables of Table 1. Other analyses, including those ranging from HDD ($\delta = 0.5$) to the classical approach to correspondence analysis ($\delta = 1$) say otherwise.

5 Discussion

Correspondence analysis is commonly applied using Pearson's chi-squared statistic as the foundation on which all numerical and visual summaries of association depend. However, Pearson's statistic is a special case of the Cressie-Read family of divergence statistics as are many other chi-squared random variables including, but not confined to, the modified log-likelihood ratio statistic, Freeman-Tukey statistic and the Cressie-

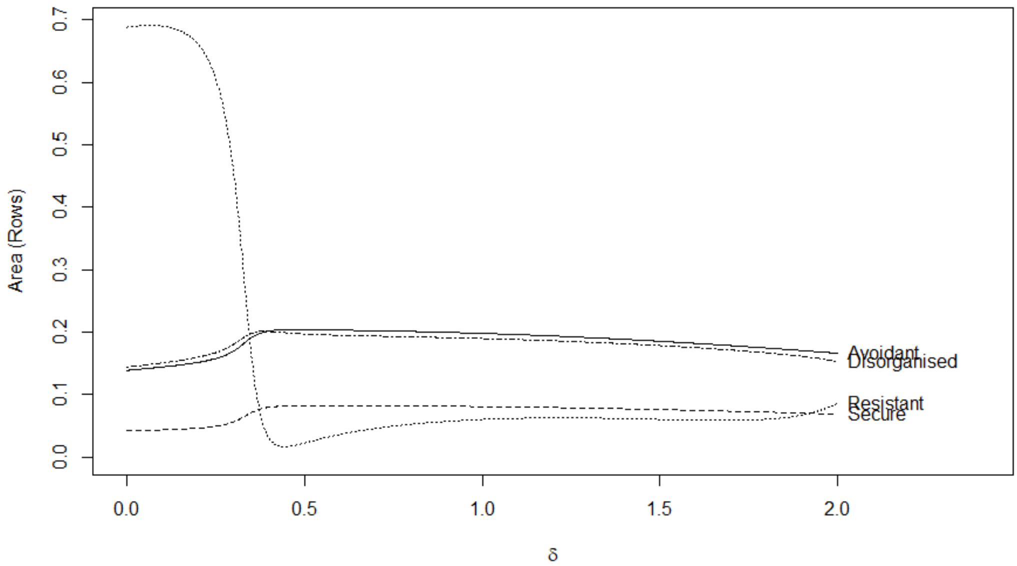


Figure 6: The area of the 95% confidence ellipses for each of the row categories of Table 1 for $\delta \in [0, 2]$

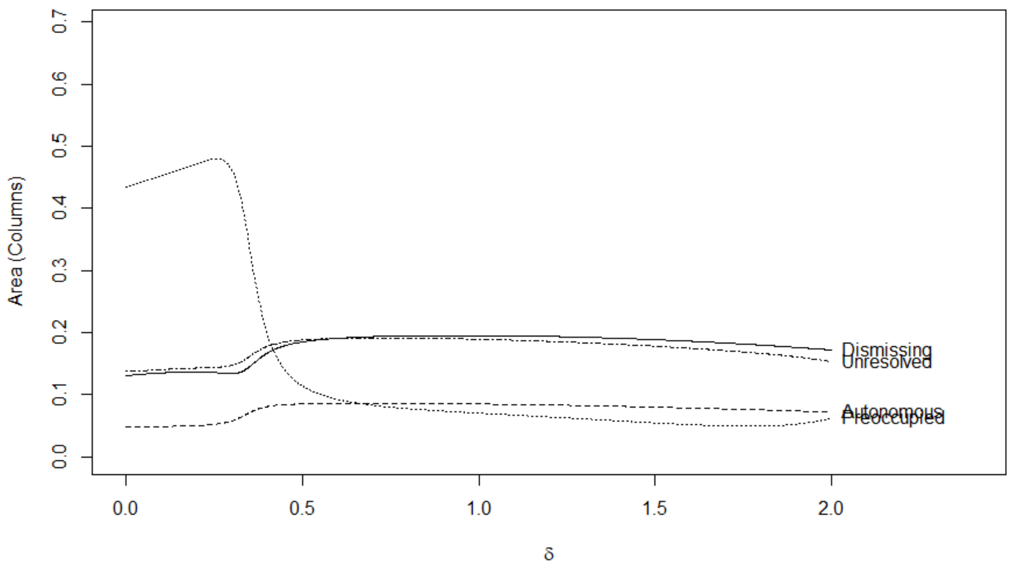


Figure 7: The area of the 95% confidence ellipses for each of the column categories of Table 1 for $\delta \in [0, 2]$

Read statistic. With the recent work of Beh and Lombardo (2023) it is now possible to include Greenacre's (2009) LRA and the HDD method of Cuadras and Cuadras (2006), and the related technique described by Beh et al. (2018), under a single framework where the family divergence statistics is used as the measure of association. The contribution outlined in this paper has been to demonstrate that the simple to calculate confidence ellipses of Beh (2010) can be extended so that they can be included within this framework. Doing so allows for one to identify those row and column categories that contribute to the association structured defined by (2).

Our aim here has been to derive the semi-major and semi-minor axis lengths of these ellipses (derived in the Appendix) and gain an understanding of several of their key features so that it provides a clear interpretation on the nature of the association structure between the variables of a contingency table. We have highlighted in our application of van IJzendoorn's (1995) data that the structure of these ellipses depends on the special case of the Cressie-Read divergence statistic considered. When using the modified log-likelihood ratio statistic ($\delta = 0$), so that a LRA is being performed on the contingency table, the confidence ellipses identified that all categories, except for one row and one column category, contributed to the association. This appears to be a unique feature of the data since, for larger values of δ that exceed about 0.4, all of the categories contribute to the association structure. It is unclear precisely why "Resistant" and "Preoccupied" behave in this manner for small values of δ and the other categories do not. However, it may be a result of the role played by the natural logarithm function (when $\delta = 0$) and the impact of small values of δ. These two categories also have relatively small cell and marginal frequencies which introduces some uncertainty as to the contribution of these categories to the association between the two variables. Irrespective of the possible cause of this feature, the application shows that methods of correspondence analysis ranging from HDD ($\delta = 1/2$) to the classical approach to correspondence analysis ($\delta = 1$) yield confidence ellipses with similar features. This does not mean that all features of the confidence ellipses will be similar for all contingency tables that are analysed, or for different values of δ. The structure of the ellipses depends entirely on the elements obtained from the GSVD of the matrix of divergence residuals with a generic element defined by (3).

More work still needs to be undertaken to explore the statistical significance of a category to the association structured measured using the Cressie-Read family of divergence statistics, or its second order series expansion. Such work has been dominated by computationally intensive procedures that use Pearson's chi-squared statistic as the measure of association although such procedures could be amended to incorporate any member of the family of the Cressie-Read divergence statistic. Confidence ellipses for alternative visualisation strategies, such as the biplot, are another natural extension of the work considered here but we shall leave this, and other avenues of research, for future consideration.

References

Ainsworth, M. D., Blehar, M. C., Waters, E., and Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Elbaum.

Anscombe, F. J. (1953). Discussion of "new light on the correlation coefficient and its transforms" (hotelling, h.). *Journal of the Royal Statistical Society, Series B*, 15(2):229–230.

Beh, E. J. (1997). Simple correspondence analysis of ordinal cross-classifications using orthogonal polynomials. *Biometrical Journal*, 39(5):589–613.

Beh, E. J. (2001). Confidence circles for correspondence analysis using orthogonal polynomials. *Journal of Applied Mathematics and Decision Sciences*, 5(1):35–45.

Beh, E. J. (2010). Elliptical confidence regions for simple correspondence analysis. *Journal of Statistical Planning and Inference*, 140:2582–2588.

Beh, E. J. and D'Ambra, L. (2009). Some interpretative tools for non-symmetrical correspondence analysis. *Journal of Classification*, 26:55–76.

Beh, E. J. and Lombardo, R. (2014). *Correspondence Analysis: Theory, Practice and New Strategies*. Wiley.

Beh, E. J. and Lombardo, R. (2021). *An Introduction to Correspondence Analysis*. Wiley.

Beh, E. J. and Lombardo, R. (2022). Correspondence analysis and the Cressie-Read family of divergence statistics. Technical report, Working Paper Series 06-22, National Institute for Applied Statistics Research Australia (NIASRA), University of Wollongong, Australia.

Beh, E. J. and Lombardo, R. (2023). Correspondence analysis and the Cressie-Read family of divergence statistics. *International Statistical Review*, 51:1–26.

Beh, E. J., Lombardo, R., and Albert, G. (2018). Correspondence analysis and the freeman–tukey statistic: A study of archaeological data. *Computational Statistics & Data Analysis*, 128:73–86.

Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975). *Discrete Multivariate Analysis*. MIT Press.

Cressie, N. and Pardo, L. (2002). Phi-divergence statistic. In El-Shaarawi, A. H. and Piegorsch, W. W., editors, *Encyclopedia of Environmetrics*, pages 1551–1555. Wiley.

Cressie, N. A. C. and Read, T. R. C. (1984). Multinomial goodness-of-fit tests. *Journal of the Royal Statistical Society, Series B*, 46(3):440–464.

Cressie, N. A. C. and Read, T. R. C. (1989). Cressie-read statistic. In Kotz, S. and Johnson, N. L., editors, *Encyclopedia of Statistical Sciences, Supplement Volume*, pages 37–39. Wiley.

Cuadras, C. M. and Cuadras, D. (2006). A parametric approach to correspondence analysis. *Linear Algebra and its Applications*, 417:64–74.

Cuadras, C. M. and Cuadras, D. (2015). A unified approach for the multivariate analysis of contingency tables. *Open Journal of Statistics*, 5:223–232.

Cuadras, C. M., Cuadras, D., and Greenacre, M. (2006). A comparison of different methods for representing categorical data. *Communications in Statistics – Simulation and Computation*, 35(2):447–459.

Domenges, D. and Volle, M. (1979). Analyse factorielle sphérique: une exploration. *Annales de l'inséé*, 35:3–84.

Freeman, M. F. and Tukey, J. W. (1950). Transformations related to the angular and square root. *The Annals of Mathematical Statistics*, 21(4):607–611.

George, C., Kaplan, N., and Main, M. (1985). Adult attachment interview. Technical report, Unpublished manuscript, University of California, Berkeley.

Gower, J., Gardner-Lubbe, S., and le Roux, N. (1975). *Understanding Biplots.* Wiley.

Greenacre, M. (2009). Power transformations in correspondence analysis. *Computational Statistics and Data Analysis*, 53(8):3107–3116.

Greenacre, M. (2010a). *Biplots in Practice.* Fundación, BBVA, Bilbao.

Greenacre, M. (2010b). Log-ratio analysis is a limiting case of correspondence analysis. *Mathematical Geosciences*, 42:129–134.

Greenacre, M. (2017). *Correspondence Analysis in Practice.* Chapman & Hall/CRC Press, 3 edition.

Kroonenberg, P. M. and Lombardo, R. (1999). Nonsymmetric correspondence analysis: A tool for analysing contingency tables with a dependence structure. *Multivariate Behavioral Research*, 34(3):367–396.

Kullback, S. (1959). *Information Theory and Statistics.* Wiley.

Lebart, L., Morineau, A., and Warwick, K. M. (1984). *Multivariate Descriptive Statistical Analysis.* Wiley.

Linting, M., Meulman, J., Groenen, P., and van der Kooij, A. (2007). Stability of nonlinear principal component analysis: An empirical study using the balanced bootstrap. *Psychological Methods*, 12(3):359–379.

Lombardo, R. and Ringrose, T. (2012). Bootstrap confidence regions in non-symmetrical correspondence analysis. *Electronic Journal of Applied Statistical Analysis*, 5(3):413–417.

Markus, M. (1994). *Bootstrap Confidence Regions in Non-Linear Multivariate Analysis.* DSWO Press.

McCullagh, P. and Nelder, J. A. (1984). *Generalized Linear Models.* Chapman & Hall, 2 edition.

Neyman, J. (1940). Contribution to the theory of certain test criteria. *Bulletin de L'Institut International de Statistique*, 24:44–86.

Neyman, J. (1949). Contributions to the theory of the χ^2 test. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, 1:239–273.

Pearson, K. (1904). On the theory of contingency and its relation to association and normal correlation. Technical report, Drapers Memoirs, Biometric Series, Vol 1, London.

Rao, C. R. (1995). A review of canonical coordinates and an alternative to correspondence analysis using hellinger distance. *Qüestiió*, 19:23–63.

Read, T. R. C. and Cressie, N. A. C. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Springer-Verlag.

Ringrose, T. (1992). Bootstrapping and correspondence analysis in archaeology. *Journal of Archaeological Science*, 19(6):615–629.

Ringrose, T. (1996). Alternative confidence regions for canonical variate analysis. *Biometrika*, 83(3):575–587.

Ringrose, T. (2012). Bootstrap confidence regions for correspondence analysis. *Journal of Statistical Computation and Simulation*, 82(10):1397–1413.

van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment. a meta-analysis on the predictive validity of the adult attachment interview. *Psychological Bulletin*, 117(3):387–403.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9(1):60–62.

Appendix

Here we derive the semi-major and semi-minor axis lengths defined by (8) and (9). The following derivation is analogous to the derivation of the lengths outlined in Beh (2010) when performing a correspondence analysis using Pearson's chi-squared statistic. The derivation of (8) and (9) is therefore given for the sake of completeness.

Suppose we consider (4) which is the property of the elements of the i'th row singular vector. Then when M^* is of optimal dimension

$$\sum_{m=1}^{M^*} a_{im}^2(\delta) = \frac{1}{p_{i\bullet}} .$$

Therefore, if we confine our attention to the first two dimensions this property may be alternatively expressed as

$$a_{i1}^2(\delta) + a_{i2}^2(\delta) = \frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a_{im}^2(\delta) . \tag{A.1}$$

Since the i'th row principal coordinates is defined as

$$f_{im}(\delta) = a_{im}(\delta)\,\lambda_m(\delta)$$

then

$$a_{im}(\delta) = \frac{f_{im}(\delta)}{\lambda_m(\delta)} . \tag{A.2}$$

Thus, substituting (A.2) into (A.1) yields

$$\frac{f_{i1}^2(\delta)}{\lambda_1^2(\delta)} + \frac{f_{i2}^2(\delta)}{\lambda_2^2(\delta)} = \frac{1}{p_{i\bullet}} - \sum_{m=3}^{M^*} a_{im}^2(\delta)$$

so that

$$\frac{f_{i1}^2(\delta)}{\lambda_1^2(\delta)\left(\frac{1}{p_{i\bullet}}-\sum_{m=3}^{M^*}a_{im}^2(\delta)\right)}+\frac{f_{i2}^2(\delta)}{\lambda_2^2(\delta)\left(\frac{1}{p_{i\bullet}}-\sum_{m=3}^{M^*}a_{im}^2(\delta)\right)}=1\,.$$

This is the equation for an ellipse centred at the point $(f_{i1}(\delta)\,,\, f_{i2}(\delta))$ in a two-dimensional correspondence plot with a semi-major and semi-minor axis length of

$$x_i(\delta)=\lambda_1(\delta)\sqrt{\frac{1}{p_{i\bullet}}-\sum_{m=3}^{M^*}a_{im}^2(\delta)} \tag{A.3}$$

$$y_i(\delta)=\lambda_2(\delta)\sqrt{\frac{1}{p_{i\bullet}}-\sum_{m=3}^{M^*}a_{im}^2(\delta)} \tag{A.4}$$

respectively. Semi- major and minor lengths for the j'th column coordinate in a two-dimensional space can be obtained in a similar manner.

To construct a $100(1-\alpha)\%$ confidence ellipse, recall that $\lambda_1^2(\delta)>\lambda_2^2(\delta)$ and that each of these quantities contributes towards the total inertia, $\phi(\delta)=\mathrm{CR}(\delta)/n$. If the $1-\alpha$ percentile of the chi-squared statistic with $(r-1)(c-1)$ degrees of freedom is denoted by χ_α^2 and $n\tilde{\lambda}_{m(\alpha)}^2$ is a percentage of χ_α^2 such that

$$\frac{\lambda_m^2(\delta)}{\mathrm{CR}(\delta)/n}=\frac{\tilde{\lambda}_{m(\alpha)}^2}{\chi_\alpha^2/n}$$

then

$$\tilde{\lambda}_{m(\alpha)}=\lambda_m(\delta)\sqrt{\frac{\chi_\alpha^2}{\mathrm{CR}(\delta)}}\,. \tag{A.5}$$

Thus, by substituting (A.5) when $m=1$, 2 for $\lambda_1(\delta)$ $\lambda_2(\delta)$ in (A.3) and (A.4), the semi-major and semi-minor axis lengths of the $100(1-\alpha)\%$ confidence ellipse for the ith row category is

$$x_{i(\alpha)}(\delta)=\lambda_1(\delta)\sqrt{\frac{\chi_\alpha^2}{\mathrm{CR}(\delta)}\left(\frac{1}{p_{i\bullet}}-\sum_{m=3}^{M^*}a_{im}^2(\delta)\right)}$$

$$y_{i(\alpha)}(\delta)=\lambda_2(\delta)\sqrt{\frac{\chi_\alpha^2}{\mathrm{CR}(\delta)}\left(\frac{1}{p_{i\bullet}}-\sum_{m=3}^{M^*}a_{im}^2(\delta)\right)}$$

which is (8) and (9), respectively, and completes the derivation.