DERLEME REVIEW

Configural Frequency Analysis as A Statistical Tool: Review

İstatistiksel Bir Araç Olarak Konfigüral Frekans Analizi

İsmet DOĞAN,[a]
Nurhan DOĞAN[a]

"Department of Bioistatistics and Medical Informatics, Afyon Kocatepe University Faculty of Medicine, Afyonkarahisar

Geliş Tarihi/*Received:* 16.03.2016
Kabul Tarihi/*Accepted:* 26.05.2016

Yazışma Adresi/*Correspondence:*
Nurhan DOĞAN
Afyon Kocatepe University
Faculty of Medicine,
Department of Bioistatistics and Medical Informatics,
Afyonkarahisar,
TÜRKİYE/TURKEY
nurhandogan@hotmail.com

ABSTRACT Configural Frequency Analysis (CFA) is a method for cell-wise testing in contingency tables whether some model is contradicted. The original idea that lead to CFA is that the cell frequencies in contingency tables can not only be used to evaluate statements about the association structure of the variables that span the table, but are worthy of consideration in their own right. CFA is a method that allows researchers to identify patterns (configurations) of variable categories that occur more often or less often than expected based on some chance model. Each cell in this cross-classification is described by a profile of variable categories. These profiles are called configurations. CFA relates individual observed cell frequencies to their expected counterparts, thereby looking for surprising, that is, statistically significant deviations. When a cross-classification cell contains more cases than expected according to some base model, it is said to constitute a CFA "type". When there are fewer cases than expected, a cell is said to constitute a CFA "antitype". Using CFA, researchers inspect either each configuration in this cross-classification or an a priori specified selection of configurations. A large number of tests has been introduced for CFA which differ in power, in their capabilities to detect types and antitypes. Lehmacher's test always has more power than the other tests. The aim of this study is to provide an introductory review for CFA, a method of categorical data analysis originally introduced as a heuristic method the so-called CFA recently has been developed into an inferential method.

Key Words: Frequency analysis; contingency tables; cross-classification; categorical data analysis; discrete multivariate analysis

ÖZET Konfigüral Frekans Analizi (KFA) çapraz tablolardaki gözelerde çelişki olup olmadığını test etmek için kullanılan hücre tabanlı bir yöntemdir. Yöntemin orijinal fikri, çapraz tablolardaki hücre frekanslarının sadece tabloda yer alan değişkenlerin ortak yapıları hakkında değerlendirme yapmak amacıyla kullanılması değil bunun yanında her bir hücrenin kendi başına dikkate alınmasının değerliliğidir. KFA araştırmacılara, bazı şans modeline dayalı ortaya çıkan değişken kategorilerine ait desenlerin (konfigürasyonların) beklenenden daha sık ya da daha az sıklıkta olup olmadıklarını tanımlamak için izin veren bir yöntemdir. Çapraz sınıflandırmadaki her bir hücre, değişken kategorilerinin bir profili tarafından tanımlanmaktadır. Bu profiller konfigürasyon olarak isimlendirilmektedir. KFA her bir hücredeki gözlenen ve beklenen frekanslar ile ayrı ayrı ilgilenmekte, dolayısıyla her bir hücre için istatistiksel olarak anlamlı sapmalar aramaktadır. Çapraz sınıflandırma tablosunun bir hücresinde gözlenen frekans, beklenen frekanstan fazla ise bu durum "type" olarak gözlenen frekans beklenen frekanstan az ise bu durum "antitype" olarak ifade edilir. KFA kullanarak, araştırmacılar çapraz sınıflandırmaki her bir konfigürasyonu ya inceler ya da önceden belirlenmiş konfigürasyonları seçer. KFA için her bir hücreyi "type" veya "antitype" olarak belirleme yetenekleri bakımından farklı güçlere sahip çok sayıda yöntem önerilmiştir. Lehmacher testinin diğer testlere göre her zaman daha güçlü olduğu ifade edilmektedir. Bu çalışmanın amacı, başlangıçta sezgisel bir yöntem olarak ileri sürülen, son zamanlarda ise çıkarımsal bir yöntem haline gelmiş olan ve kategorik veri analizinde kullanılan KFA yönteminin tanıtılmasıdır.

Anahtar Kelimeler: Frekans analizi; çapraz tablo; çapraz sınıflandırma; kategorik veri analizi; kesikli çok değişkenli analiz

doi: 10.5336/biostatic.2016-51363

Turkiye Klinikleri J Biostat 2016;8(2):172-9

The empirical investigation of causal relationships typically requires knowledge about where cause-effect relationships might exist. If this konowledge is not available, researchers have no statistical methods at their disposal. Usage of Configural Frequency Analysis (CFA) was proposed as a searching device for the identification of possible causal relationships.[1] To track the complex intervention pathways, rigorous yet flexible assesment and evaluation methods captured multicomponenet and dynamic community trends. CFA examines the level of key resources in communities and how they are arranged. CFA can identify potential differences in communities because it allows for a case oriented, as opposed to variable oriented, approach to analyzing community level data. Variable oriented analyses seek to explain associations between variables across communities, whereas case oriented analyses can identify clusters of communities having different levels of variables. CFA is similar to cluster analysis and latent growth curve analysis in that it can detect configurations of cases that deviate from what is expected. These deviations are the result of a system that "pushes" certain cases in a direction away from the general pattern. CFA provides a way to identify community patterns that may be associated with different underlying systems.[2]

The aim of this study is to provide an introductory review for CFA, a method of categorical data analysis originally introduced as a heuristic method the so-called CFA recently has been developed into an inferential method.

In empirical sciences, many data are categorical, that is nominal or ordinal in nature. The states of categorical variables are mutually exclusive. Therefore, every individual can be assigned to only one state per variable at a time. When two or more categorical variables are measured they can be cross-classified so that every variable may be observed under every state of every other variable.[3] CFA is a multivariate statistical method for the analysis of cross-classifications that can be used in a wide range of multivariate experimental designs. CFA allows researchers to identify patterns of categories that contradict expectations from CFA specific base models. Many base models of CFA use methods of estimation known from log-linear modeling. These methods yield parameter estimates and expected cell frequencies that can be trusted only if the ratio of sample size to number of cells is reasonably large. In addition, the cell-wise tests typically employed in CFA can be trusted only if the expected cell frequency is sufficiently large.[4]

CFA is a method for the analysis of bivariate or multivariate cross-classifications of categorical variables. In contrast to such methods as log-linear modeling, which express results mostly in terms of relationships among variables, CFA allows one to look for effects at the level of individual cells, or groups of cells, in a table. The patterns of categories that define a cell, that is, the cell indices, are called configurations.[5]

In the garden of classification methods, CFA plays a particular role. In contrast to such methods as cluster analysis or latent class analysis, both of which create a priori unknown groups from raw data, CFA asks whether clusters of existing groups contain more or fewer cases than expected. CFA shares the characteristic of analyzing existing groups with discriminant analysis and with logistic regression. However, in contrast to these two methods, the typical application of CFA is exploratory in nature, as is the case for many other methods of classification.[6]

The CFA is a test procedure for the analysis of multidimensional contingency tables. It computes the probability of whether a given pattern could be expected to occur by chance, according to the margin frequency distributions of the binary variables A, B, and C pattern, or whether it is significant, i.e., overrepresented, for a given

stage. If more than 1 pattern becomes significant for a given stage, this suggests that the stage contains different identifiable substages. The CFA allows one to define types, by a statistical probability above the chosen level of significance, as well as antitypes, which occur less frequently than expected by chance.[7]

CFA is a well established simple analysis technique used for detection of syndromes in psychology and medicine. The identification of types or syndromes in contingency tables is a frequent and important problem in psychology and medicine. CFA is widely used for this purpose. Krauth and Lienert (1973) propose to test cellwise the null hypothesis of local independence, i.e. testing cell by cell the residuals of the model of total independence against zero.[8] Apart from using a purely descriptive approach, several variants of binomial and hypergeometrical as well as χ^2 statistics have been used to test if the configurations of parameter values occur more (or less) frequently than expected by chance alone. The applications of this method can be used in various sciences. For most variants of CFA tests, the configurations of inputs cannot be tested completely independent from each other while the test statistics are estimated only.[9]

CFA is a widely used methods of explorative data analysis. It tries to detect patterns in the data that occur significantly more or significantly less often than expected by chance. Patterns which occur more often than expected by chance are called CFA types, while those which occur less often than expected by chance are called CFA antitypes. CFA types and antitypes are defined symmetrically, but in practical applications of CFA, researchers are mainly interested in detecting CFA types. For example, in clinical studies the interest is generally in detecting symptom combinations which are indicators of a disease. These are by definition symptom combinations which occur more often than expected by chance, i.e. types.[10] The definition of types by Lienert corresponds to this concept: Let subjects be characterized by a given set of attributes (traits) each exhibiting few categories. A "type" is a pattern (=configuration) of categories of the attributes which exhibit more subjects than expected (than "normal"); analogously, an "antitype" is a configuration which exhibits fewer subjects than expected. This definition makes only sense if the number of (anti)types is small compared to the number of possible configurations. Hence, we have to analyse multi-variate categorical data where the configurations are nothing else than the set of all possible outcomes.[8]

CFA has been developed intensively since its first presentation in 1968 and is now among the more popular methods of data analysis. One of the reasons for the increasing popularity of CFA is that the results of this method of data analysis are deemed easy to interpret. CFA is performed in five steps. The first step involves selecting a base model. The second step of CFA involves the selection of a concept of deviation from independence. The third step involves the selection of a significance test. The fourth step involves estimating (or determining) the expected cell frequencies, performing the significance tests, and identifying those configurations that constitute types or antitypes. The fifth step involves interpreting types and antitypes.[6]

Among the advantages of CFA often mentioned are:[10]

- it can be used in a wide range of multivariate experimental designs,
- it requires only variables on a nominal scale level,
- it is distribution-free,
- it is extremely easy to compute,
- it allows a profile-oriented analysis, i.e. the unit of analysis is the profile (configuration) of all observed values of a person.

Further, CFA aims at identification of outstanding cells of multivariate cross classifications instead of fitting a model. It seeks to identify patterns that stand out as more frequent or less frequent than expected by chance. Another advantage is that unlike many other methods, CFA does not use a similarity or distance measure to identify types, but takes only types into consideration with identical attribute patterns, which makes its type definition more accurate than any other previous type definition.[11]

There has been a large number of tests proposed for use in CFA. The best known of these tests include the binomial test, the χ^2 component test, and Lehmacker's hypergeometric tests. The best known and most widely used asymptotic test in CFA is the Pearson χ^2 test. To illustrate application of the CFA, consider the following example. A researcher investigates the three variables, A, B, and C. Variable A has I categories. Variable B has J categories, and variable C has K categories: that is, we have indexes $n_{i..}$, $n_{.j.}$, and $n_{..k}$, with $i = 1, \dots, I$, $j = 1, \dots, J$, and $k = 1, \dots, K$. For the cross-tabulation of the three variables A, B, and C, χ^2 test statistic is as follows;

$$\chi^2 = \frac{\left(o_{ijk} - e_{ijk}\right)^2}{e_{ijk}}$$

o_{ijk}: the observed cell frequency for cell ijk of the cross tabulation,

e_{ijk}: the expected cell frequency for cell ijk of the cross tabulation,

The expected cell frequency for cell ijk of the cross-tabulation of A, B, and C is estimated using

$$e_{ijk} = \frac{n_{i..} n_{.j.} n_{..k}}{N^{d-1}}$$

where periods indicate variables summed across. In other words, $n_{i..}$ denotes the frequency of the ith category of variable A, $n_{.j.}$ denotes the frequency of the jth category of variable B and $n_{..k}$ denotes the frequency of the kth category of variable C. N denotes the sample size, and d is the number of variables.[12] The null (H_0) and alternative (H_1) hypotheses of total independence of the variables are defined by,[13]

$$H_0 = p_{ijk} = p_{i..} p_{.j.} p_{..k} \quad for\ all \quad (i,j,k)$$

$$H_1 = p_{ijk} \neq p_{i..} p_{.j.} p_{..k} \quad for\ all \quad (i,j,k)$$

$p_{i..} = \frac{n_{i..}}{N}$ $p_{.j.} = \frac{n_{.j.}}{N}$ $p_{..k} = \frac{n_{..k}}{N}$: marginal probabilities,

p_{ijk}: the cell probability for cell ijk of the cross tabulation.

A well-known test statistic for H_0 is Pearson's χ^2. If the χ^2 test for one degree of freedom suggests that the difference is significant for a given α level, then

- Configuration is called CFA *type* if observed value > expected value,
- Configuration is called CFA *antitype* if observed value < expected value.

Types indicate that the states that define the cell under study "go together well" such that more subjects display this pattern of states than expected under the assumption of total independence. *Antitypes* indicate a conceptual misfit between the independence assumption that underlies the estimation of expected frequencies and the data.[3] If there is no significant difference between observed value and expected value, then *configuration* is neither a *type* nor an *antitype*. Thus, each configuration can in principle have three different states. It can be a *type*, an *antitype*, or *not classified*.

The χ^2 approximation to the z-statistic is calculated as follows:[14]

$$z_{ijk} = \frac{\left(o_{ijk} - e_{ijk}\right)}{\sqrt{e_{ijk}}}$$

TABLE 1: CFA results for standart normal approximation of the χ^2 test

$DUSG$	o_{ijk}	e_{ijk}	χ^2	z_{ijk}	$p(z_{ijk})$	Type/Antitype*
1111	11	10.477	0.026	0.161	0.4539	
1112	19	7.079	20.071	4.480	0.0000	Type
1113	3	9.345	4.308	-2.075	0.0190	Antitype
1121	13	10.814	0.441	0.664	0.2531	
1122	9	7.306	0.392	0.626	0.2655	
1123	6	9.645	1.377	-1.173	0.1203	
1211	3	7.538	2.732	-1.653	0.0492	Antitype
1212	13	5.093	12.273	3.503	0.0002	Type
1213	0	6.723	6.723	-2.593	0.0048	Antitype
1221	4	7.780	1.836	-1.355	0.0877	
1222	12	5.256	8.649	2.941	0.0016	Type
1223	1	6.939	5.083	-2.254	0.0121	Antitype
2111	30	31.879	0.110	-0.332	0.3696	
2112	14	21.540	2.639	-1.624	0.0521	
2113	44	28.433	8.522	2.919	0.0018	Type
2121	38	32.902	0.789	0.888	0.1871	
2122	11	22.231	5.674	-2.382	0.0086	Antitype
2123	23	29.345	1.372	-1.171	0.1207	
2211	18	22.935	1.062	-1.030	0.1514	
2212	9	15.497	2.724	-1.650	0.0494	Antitype
2213	23	20.456	0.316	0.562	0.2869	
2221	31	23.671	2.268	1.506	0.0660	
2222	13	15.994	0.560	-0.748	0.2270	
2223	32	21.112	5.614	2.369	0.0089	Type

The following example presents a re-analysis of data published by von Eye (2002).[15] In a study on the relationships between the three psychiatric symptoms Depression (D), Feeling of Insecurity (U), and Mood Swings (S) on the one hand and the three psychiatric diagnoses (G) Cyclothymia (C), Anxiety Neuroticism (A), and Neurotic Depression (N). 380 inpatients were diagnosed as either displaying (=1) or not displaying (=2) a symptom. Each patient had been diagnosed as falling under C, A or N. Crossed, these four variables form a $2x2x2x3$ contingency table. The complete CFA results appear in Table 1.

The χ^2 test can be replaced by other tests, for example, the binomial test or the hypergeometric test. A standart normal approximation of the binomial test that can be applied when $Np_{ijk} \geq 10$, with $p_{ijk} = e_{ijk}/N$. The approximation is defined as follows: [12]

$$z_{ijk} = \frac{(o_{ijk} - e_{ijk})}{\sqrt{e_{ijk} q_{ijk}}}$$

where $q_{ijk} = 1 - p_{ijk}$. The following relationship holds for degree of freedom = 1: $z^2(\alpha/2) = \chi^2(\alpha)$.

The results obtained from the example for this method is shown in Table 2.

The Lehmacher's approximation to the z-statistic is calculated as follows: [16]

$$z_{ijk}^L = \frac{(o_{ijk} - e_{ijk})}{\sigma_{ijk}}$$

$$\sigma_{ijk}^2 = Np_{ijk}\left[1 - p_{ijk} - (N-1)\left(p_{ijk} - \tilde{p}_{ijk}\right)\right]$$

$$p_{ijk} = \frac{n_{i..} n_{.j.} n_{..k}}{N^d}$$

$$\tilde{p}_{ijk} = \frac{(n_{i..} - 1)\left(n_{.j.} - 1\right)(n_{..k} - 1)}{(N-1)^d}$$

The results obtained from the example for Lehmacher's test is shown in Table 3.

Anscombe's z-approximation,[17]

$$z_{ijk}^A = \frac{3\left[o_{ijk}^{2/3} - \left(e_{ijk} - \frac{1}{6}\right)^{2/3}\right]}{2e_{ijk}^{1/6}}$$

The results obtained from the example for Anscombe's test is shown in Table 4.

TABLE 2: CFA results for standart normal approximation of the binomial test.

$DUSG$	o_{ijk}	e_{ijk}	z_{ijk}	$p(z_{ijk})$	Type/Antitype*
1111	11	10.477	0.163	0.4350	
1112	19	7.079	4.522	0.0000	Type
1113	3	9.345	-2.101	0.0178	Antitype
1121	13	10.814	0.674	0.2500	
1122	9	7.306	0.632	0.2635	
1123	6	9.645	-1.188	0.1173	
1211	3	7.538	-1.669	0.0475	Antitype
1212	13	5.093	3.527	0.0002	Type
1213	0	6.723	-2.616	0.0044	Antitype
1221	4	7.780	-1.369	0.0854	
1222	12	5.256	2.961	0.0015	Type
1223	1	6.939	-2.275	0.0114	Antitype
2111	30	31.879	-0.347	0.3640	
2112	14	21.540	-1.672	0.0472	Antitype
2113	44	28.433	3.035	0.0012	Type
2121	38	32.902	0.929	0.1762	
2122	11	22.231	-2.454	0.0070	Antitype
2123	23	29.345	-1.219	0.1113	
2211	18	22.935	-1.063	0.1438	
2212	9	15.497	-1.685	0.0460	Antitype
2213	23	20.456	0.578	0.2816	
2221	31	23.671	1.555	0.0599	
2222	13	15.994	-0.765	0.2221	
2223	32	21.112	2.438	0.0074	Type

$\alpha = 0.05$

TABLE 3: CFA results for standart normal approximation of the Lehmacher's test.

$DUSG$	o_{ijk}	e_{ijk}	p_{ijk}	$\tilde{p}_{ijk}$	σ^2_{ijk}	z^L_{ijk}	$p(z^L_{ijk})$	Type/Antitype*
1111	11	10.477	0.0276	0.0271	8.3613	0.062	0.4751	
1112	19	7.079	0.0186	0.0183	5.9529	2.002	0.0226	Type
1113	3	9.345	0.0246	0.0242	7.5909	-0.835	0.2016	
1121	13	10.814	0.0285	0.0280	8.5787	0.254	0.3994	
1122	9	7.306	0.0192	0.0188	6.1153	0.276	0.3909	
1123	6	9.645	0.0254	0.0249	7.7916	-0.467	0.3200	
1211	3	7.538	0.0198	0.0195	6.3440	-0.715	0.2372	
1212	13	5.093	0.0134	0.0131	4.4653	1.770	0.0383	Type
1213	0	6.723	0.0177	0.0174	5.7369	-1.172	0.1206	
1221	4	7.780	0.0205	0.0201	6.5179	-0.580	0.2810	
1222	12	5.256	0.0138	0.0135	4.5923	1.468	0.0710	
1223	1	6.939	0.0183	0.0179	5.8961	-1.007	0.1569	
2111	30	31.879	0.0839	0.0831	19.4814	-0.096	0.4616	
2112	14	21.540	0.0567	0.0560	14.3834	-0.524	0.3001	
2113	44	28.433	0.0748	0.0740	17.9122	0.869	0.1924	
2121	38	32.902	0.0866	0.0858	19.8746	0.256	0.3988	
2122	11	22.231	0.0585	0.0578	14.6893	-0.764	0.2223	
2123	23	29.345	0.0772	0.0764	18.2806	-0.347	0.3643	
2211	18	22.935	0.0604	0.0597	15.5974	-0.316	0.3758	
2212	9	15.497	0.0408	0.0402	11.3737	-0.571	0.2839	
2213	23	20.456	0.0538	0.0532	14.2785	0.178	0.4293	
2221	31	23.671	0.0623	0.0616	15.9472	0.459	0.3229	
2222	13	15.994	0.0421	0.0415	11.6441	-0.257	0.3985	
2223	32	21.112	0.0556	0.0549	14.6055	0.745	0.2280	

$\alpha = 0.05$

CFA significance tests are typically applied to each cell of a cross-tabulation. Therefore, the risk of committing an α-error can be high. In addition, tests will be dependent. Simultaneous CFA testing of more than one configuration entails two interrelated problems. The first is the problem of mutual dependence of multiple test of the same data set. The second is the problem of multiple testing. Both problems lead to the alpha level, nominally set at an a priori level, becom-

TABLE 4: CFA results for standart normal approximation of the Anscombe's test.

$DUSG$	o_{ijk}	e_{ijk}	z^A_{ijk}	$p(z^A_{ijk})$	Type/Antitype*
1111	11	10.477	0.211	0.4162	
1112	19	7.079	3.779	0.0001	Type
1113	3	9.345	-2.380	0.0086	Antitype
1121	13	10.814	0.695	0.2435	
1122	9	7.306	0.666	0.2526	
1123	6	9.645	-1.209	0.1132	
1211	3	7.538	-1.829	0.0337	Antitype
1212	13	5.093	3.011	0.0013	Type
1213	0	6.723	-3.824	0.0001	Antitype
1221	4	7.780	-1.438	0.0751	
1222	12	5.256	2.596	0.0047	Type
1223	1	6.939	-2.801	0.0025	Antitype
2111	30	31.879	-0.306	0.3795	
2112	14	21.540	-1.702	0.0444	Antitype
2113	44	28.433	2.733	0.0031	Type
2121	38	32.902	0.896	0.1850	
2122	11	22.231	-2.612	0.0045	Antitype
2123	23	29.345	-1.187	0.1175	
2211	18	22.935	-1.036	0.1499	
2212	9	15.497	-1.752	0.0399	Antitype
2213	23	20.456	0.588	0.2782	
2221	31	23.671	1.471	0.0705	
2222	13	15.994	-0.732	0.2319	
2223	32	21.112	2.238	0.0126	Type

$\alpha = 0.05$

ing, in fact, inflated. Many scientest suggest that the local level alpha should be controlled. Therefore, various methods of adjusting the significance level to these risks have been proposed. Alternative procedures for controlling the local and the multiple level alphas in simultaneous testing have been discussed.[3] Bonferroni adjustment seems to be the currently most popular method of protection.[1]

CONCLUSION

A different approach to classification analysis is taken in CFA. It is a set of methods for analyzing all possible value patterns. Instead of relying on a method for sorting patterns into groups according to their similarity, all patterns are studied directly. Each theoretically possible pattern is called configuration. To carry out a CFA the involved variables have to be discrete often dichotomized or trichotomized to make manageable the number of configurations for which the observed frequencies are to be examined.[18]

CFA takes a perspective of data analysis that differs from the perspective taken by most statistic methods. CFA does not ask whether variables are associated with each other, interact, or are predictive of each other. CFA asks whether a particular pattern of categorical predictor variables allows one to predict the above or the below expectancy occurence rate of a particular criterion pattern.[19] For the decision as to whether a cell constitutes a CFA type or antitype, a number of tests has been proposed. Each of these tests can be used to examine individual cells of a cross-classification. Tests for the examination of groups of cells have also been proposed. CFA tests are either exact or asymptotic, and they either can be used under any sampling scheme or require product-multinomial sampling. The binomial test is exact and can be used under any sampling scheme. The z-test and the χ^2 test are asymptotic and can also be used under any sampling scheme. The exact and asymptotic hypergeometric tests require product-multinomial sampling. These tests are the most powerful of all current CFA tests, by far.[20] Lehmacher's test has the most balanced antitype to type ratio, followed by the z-test and the χ^2 test. Each of these tests typically

detects more types than antitypes when samples are small, and more antitypes than types when samples are large. Anscombe's z-approximation almost always detects more antitypes than types. Lehmacher's test always has more power than the z-test and the χ^2 test. Anscombe's z lies between the z and the χ^2 tests for types, and between Lehmacher's test and the z-test for antitypes.[17]

REFERENCES

1. von Eye A, Brandtstädter J. Configural frequency analysis as a searching device for possible causal relationships. Methods of Psychological Research 1997;2(2):7-23.
2. Brennan LK, Brownson RC, Hovmand P. Evaluation of Active Living by Design: implementation patterns across communities. Am J Prev Med 2012;43(5 Suppl 4):S351-66.
3. von Eye A. Concepts of configural frequency analysis: the analysis of contingency tables. Introduction to Configural Frequency Analysis, 1st ed. Cambridge: Cambridge University Press; 1990, p. 3-39.
4. von Eye A, Pena EG. Configural frequency analysis of large sparse cross-classifications. Psychology Science 2005; 47(3/4):356-76.
5. von Eye A, Mair P, Mun EY. Introduction. Advances in Configural Frequency Analysis, 1st ed. New York: The Guilford Press; 2010. p. 1-15.
6. von Eye A. Base models for configural frequency analysis. Psychology Science 2004;46(1):150-70.
7. Bettina M, Wolf DG, Hartmut S. A taxonomic analysis of sleep stages. SLEEP 2006;29(7):967-74.
8. Victor N, Kieser M. Configural frequency analysis and association analyses in contingency tables. Computational Statistics & Data Analysis 2003;44(1-2):419-29.
9. Harloff J. An efficient algorithm for Lindner's test (configural frequency analysis). Quality and Quantity 2012;46(1):371-8.
10. Schrepp M. The use of configural frequency analysis for explorative data analysis. Br J Math Stat Psychol 2006;59(Pt 1):59-73.
11. Loeffert S, Ommen O, Kuch C, Scheibler F, Woehrmann A, Baldamus C, et al. Configural frequency analysis as a method of determining patients' preferred decision-making roles in dialysis. BMC Med Inform Decis Mak 2010;10(47):1-11.
12. von Eye A, Spiel C, Wood PK. Configural frequency analysis in applied psychological research. Applied Psychology 1996;45(4):301-52.
13. Lienert GA, Oeveste HZ. Configural frequency analysis as a statistical tool for developmental research. Educational and Psychological Measurement 1985;45(2):301-7.
14. Stemmler M. Significance testing in CFA: Chi-square approximation to the z-test. Person-Centered Methods: Configural Frequency Analysis (CFA) and other Methods for the Analysis of Contingency Tables, 1st ed. Switzerland: Springer International Publishing; 2014. p.19-23.
15. von Eye A. The CFA specialty file and alternative approaches to CFA: more facets of CFA. Configural Frequency Analysis: Methods, Models and Applications, 1st ed. New Jersey: Lawrence Erlbaum Associates, Inc. Press; 2002. p. 287-9.
16. Lienert GA, Lehmann E. Differential drug effects identified by 3-way configural frequency analysis. Neuropsychobiology 1984;11(4):247-50.
17. Lehmacher W. A more powerful simultaneous test procedure in configural frequency analysis. Biometrical Journal 1981;23(5):429-36.
18. von Eye, A. The odds favor antitypes- a comparison of tests for the identification of configural types and antitypes. Methods of Psychological Research 2002;7(3):1-29.
19. Bergman LR. The application of a person-oriented approach: types and clusters. In: Bergman LR, Cairns RB, Nilsson LG, Nystedt L, eds. Developmental Science And The Holistic Approach. 1st ed. New Jersey: Lawrence Erlbaum Associates, Inc. Press; 2000. p.137-49.
20. von Eye A, Mair P, Bogat GA. Prediction models for configural frequency analysis. Psychology Science 2005;47(3/4):342-55.
21. von Eye A, Mair P. A functional approach to configural frequency analysis. Austrian Journal of Statistics 2008;37(2):161-73.