

Glossary

Antitype Represents or indicates an under-frequented cell ($f_{(o)} < f_{(e)}$).

Base Model Is the underlying model to calculate the expected frequencies; it usually the independence model. Different base models result in different types or *antitypes*.

Bonferroni Alpha Adjustment In case of multiple test procedures a Bonferroni alpha adjustment is necessary. It divides the alpha level by the number of tests, e.g., with a two-sided test of $\alpha = 0.05$ and three tests, the new alpha level would be $\alpha^* = 0.025/3 = 0.00833$.

Chi-Square Automatic Interaction Detection (CHAID) In case that we have many categorized variables and we define one variable as a dependent. CHAID partitions the data into mutually exclusive, exhaustive, subsets that best describes the dependent variable. One applies a CHAID-model for contrasting groups. The use of this program results in a graphical illustration called a tree diagram.

Coefficient of Precision Is a coefficient which can be interpreted similar to the determination coefficient R^2 in multiple regression. This is the statistic Q, which is a coefficient of precision or a coefficient of *the pregnancy of a type*.

Configural Cluster Analysis (CCA) Is a zero-order CFA where the underlying model includes no main effects or interactions; that is, each cell has the same expected frequency.

Configural Frequency Analysis (CFA) Is a statistical method that looks for over- and under-frequented cells or patterns in a contingency table. Over-frequented means, that the observations in this cell or configuration are observed more often than expected, under-frequented means that this configurations is observed less often than expected.

Confreq An R-package using Configural Frequency Analysis and the log-linear modeling approach for analyzing contingency tables.

Correspondence Analysis (CA) Is another statistical tool that investigates the relationship between persons or objects in contingency tables. CA examines the relationships between categories of nominal data in a cross-tabulation based on their associations and CA presents the results in a graphical description called

a perceptual map. In a perceptual map, persons or objects are plotted such that their proximity represent closeness or strong relationships. Sometimes CA is also referred to as homogeneity analysis.

Design Matrix X is design matrix containing the effect-coded main effect and interaction terms plus the constant. The design matrix X has as many rows as there are cells or configurations, and $m + 1$ columns. m is the number of weights; the first weight is always the constant, coded with ones.

Discrimination Type Is a type that differentiates significantly between two (originally) independent samples.

Global Chi-square Is a statistic referring to the whole contingency table. It is the sum of all cell-wise deviations between observed and expected frequencies using the chi-square statistic.

Local Chi-square Is a statistic referring to one cell or configuration, where it represents the calculated deviation between the observed and expected frequencies using the chi-square statistic.

Lancaster decomposition Lancaster (1969) found out, that the global chi-square of a first order CFA is composed additively of all possible interactions.

Log-linear modeling Is a statistical tool to investigate the underlying structure of dependency in a contingency table. The logarithm of the expected frequencies can be expressed in a linear function of parameters. The parameters indicate the impact of main effects and interactions on the data in the contingency table.

Latent Class Analysis; LCA Is most often seen as an equivalent to factor analysis (FA). While FA extracts latent continuous factors from a pool of continuous variables, LCA extracts latent categorical factors or classes from a pool of categorical variables. Both statistical tools aim at data reduction.

Longitudinal CFA This version of CFA tests the stability or instability of configurations over time. Configurations define observations of patterns of one sample over time.

Meehl's paradox A data example constructed by Meehl where there are no bivariate associations or correlations but higher order associations which allow the exact prediction of a group membership in a 2 by 2 by 2 contingency table.

Quasi-Independence Means, that after blanking out a certain cell, the remaining contingency table has to be independent and therefore, the respective chi-square must not be significant.

R Software R is an open source software which is suitable for Linux, MacOS X, and Windows. R (R Development Core Team, 2011) is a program for data analysis, data manipulation and graphical display.

Saturated Model Is a model that reproduces the observed values perfectly. It includes all main and interaction effects.

Structural Zeros Are usually cells which cannot be observed (e.g., a pattern of heavy rain together with a beautiful blue sky).

Test of Marginal Homogeneity This table searches for equal probabilities (i.e., $p_{i.} = p_{.i}$) and therefore equal pairs of marginals (i.e., $f_{i.} = f_{.i}$) in a square symmetric contingency table.

Type Represents or indicates an over-frequented cell ($f_{(o)} > f_{(e)}$).

Index

A

Answer Tree, 60
antitype, 6

B

base model, 4, 45
binomial test, 19
biprediction type, 78
Bonferroni alpha adjustment, 6

C

CFA of First Differences, 51
cfa.exe, 13
CHAID, 60
chi-square test, 21
coefficient of precision, 42
conditional response probabilities, 65
Configural Cluster Analysis (CCA), 39
Configural Frequency Analysis (CFA), 1
corrected standardized residuals (csr), 44
correspondence analysis (CA), 68

D

degree of freedom, 5, 7, 36
Deming-Stephan-Algorithm, 46
design matrix, 26
discrimination type, 55

F

first differences, 51
Fisher's exact test, 77
Fuchs-Kenett Test, 44
functional CFA, 45

G

General Linear Model (GLM), 26
global chi-square, 3, 36
goodness-of-fit, 36

H

hierarchical log-linear modeling, 35

I

index-point notation, 4
inertia, 71
Interaction Structure Analysis (ISA), 75

K

Kimball's decomposition test, 78

L

Lancaster decomposition, 36
Latent Class Analysis (LCA), 64
local chi-square value, 5
log-linear models, 25
Longitudinal CFA, 51

M

main effects model, 26
Meehl's paradox, 9
multinomial sampling, 2
multiple correspondence analysis (MCA), 69

O

outlandish cells, 6

P

Prediction-CFA, 73

Q

quasi-independence, 47

R

R-Package confreq, 15

R-Studio, 15

regional contingency, 78

S

saturated model, 32

significance testing in CFA, 19

standardized residuals, 31

structural zero, 46

T

test of marginal homogeneity, 53

two-sample CFA, 79

type, 6

V

Victor-cells, 45

visual shape patterns, 52

Y

Yates' continuity correction, 23