

Korrespondenssianalyysi - graafinen ja geometrinen data-analyysin menetelmät

Jussi Hirvonen

Versio 0.01, tulostettu 2018-08-08

Sisältö

Alkutoimia	5
1 Johdanto	7
1.1 Tutkielman tavoite (tutkimusongelma?)	7
2 Data	9
3 Yksinkertainen korrespondenssianalyysi	11
4 Yksinkertaisen korrespondenssianalyysi - tulkinnan syventäminen	13
5 Yksinkertaisen korrespondenssianalyysin laajennuksia	15

Alkutoimia

Ladataan r-paketit, ei tulosteta dokumenttiin. Pelkkä YAML- ‘front matter’, lisäkonfiguroinnit tiedostoissa `__bookdown.yml` ja `__output.yml`.

Dokumenttiin kuuluvat Rmd-tiedostot luetellaan eksplisiittisesti (ei vielä).

Ideoita

1. Ehkä automaattista R-kirjastojen dokumentointia voisi harkita?
2. Saako gitbook-tulosteessa päälle asetuksen `code_folding: hide`? Vaatii teeman (theme), jos tarpeen voi lisätä `__output.yml` - tiedostoon esim. `html_book` - formaatin.

Luku 1

Johdanto

xyz Kirjoitetaan disposition pohjalta, keräillään kaikki yleiset ca-luonnehdinnat yhteen paikkaan eli johdantoon.

Mahdollisia lisäyksiä

1. Lyhyt esitys CA:n historiasta (vai omaksi luvuksi, luku 2)?
2. Käytetyt ohjelmistot, tekninen ympäristö ml. bookdown-asetukset. Ehkä paremmin omaksi liitteeksi?
3. Tavoitteet, sisältö, rajaukset (jota voi myöhemmin täydentää)
4. Muutamat puutteet, onko kerrottava tässä?
 - data: ei huomioida sitä, että otoskoot vaihtelevat aika paljon eli “maapainot” eri suuruisia
 - ei huomioida muitakaan otantaan liittyviä asioita (tämä ainakin mainittava data-osuudessa)
 - kuvaileva menetelmä, mutta mikä on tutkimusongelma? Sellainen pitäisi olla.

****zxy*** Mitä on korrespondenssianalyysi? Muutamalla kappaleella. Yksi kappale historiasta.

1.1 Tutkielman tavoite (tutkimusongelma?)

ks Esitellään korrespondenssianalyysin käsitteet ja graafisen analyysin periaatteet.

ks Tämän voi tehdä yksinkertaisen korrespondenssianalyysin avulla. Yksinkertainen kahden luokittelumuuttujan korrespondenssianalyysi antaa graafisen analyysin “...perussäännöt tulkinalle. Kaikki muut korrespondenssianalyysin muodot ovat saman algoritmin soveltamista toisen tyyppiisiin datamatriiseihin, ja tulkintaa sovelletaan vastaavasti (with the consequent adaptation of the interpretation)” (Greenacre and Hastie, 1987, , s. 437)

Luku 2

Data

xyz Voisi miettiä paremman otsikon

Luku 3

Yksinkertainen korrespondenssianalyysi

xyz Tässä yksi kysymys, kuusi maata, peruskäsitteet lopussa

Luku 4

Yksinkertaisen korrespondenssianalyysi - tulkinnan syventäminen

xyz Tarkasti läpi keskeiset tulokset ja niiden tulkinta, kaavat, ja ytimenä eri kuvat eli kartat.

Luku 5

Yksinkertaisen korrespondenssianalyysin laajennuksia

xyz Yksinkertainen korrespondenssianalyysi on menetelmän tulkinnan perusta. Perusasetelmaa kahden luokittelumuuttujan ristiintaulukoinnista voidaan laajentaa monipuolisempiin tutkimusasetelmiin. Varsinainen useamman muuttujan korrespondenssianalyysi (MCA - multiple correspondence analysis) esitellään seuraavassa luvussa.

Lähteet

Greenacre, M. and Hastie, T. (1987). The geometric interpretation of correspondence analysis. *Journal of the American Statistical Association*, 82(398):437–447. doi: 10.1080/01621459.1987.10478446.