

Kvantitatiiviset menetelmät

Pienryhmäkokoontumisissa tarvitaan EK03-aineiston haastattelulomake. Sen voi tulostaa verkosta. Linkki löytyy kurssin kotisivulta: <http://www.valt.helsinki.fi/staff/mmattila/kvanti/>

© Mikko Mattila 2005

1

Soveltuvan menetelmän valinta

SELITETTÄVÄ MUUTTUJA

SELITTÄVÄ MUUTTUJA	Laatuero- tai järjestysasteikko	Välimatka- tai suhdelukeasteikko
	Laatuero- tai järjestysasteikko	Välimatka- tai suhdelukeasteikko
	- Ristiintaulukointi - Log-lineaariset mallit	- Varianssianalyysi
	- Logistinen regressio - Multinomiaalinen regressio	- Regressioanalyysi

© Mikko Mattila 2005

2

Faktoriansalyysi

- Joidenkin käsitteiden mittaaminen yhdellä muuttujalla mahdotonta
 - tarvitaan useita samaa ilmiötä mittaavia muuttujia
- Faktori on piilevä (eli latentti), ei suoraan havaittavissa oleva, konstruktio, jonka olemassaolo päätellään mitattavissa olevista muuttujista
- Esimerkkejä faktoreista: konservatiivisuus, sosiaalisuus, vieraantuminen, älykyys
- Faktoreita tutkitaan faktoriansalyysin avulla

© Mikko Mattila 2005

3

Faktoriansalyysi

- Eksploriatiivinen faktoriansalyysi
 - Aineistolähtöinen menetelmä
 - Ei vahvoja etukäteisoletuksia faktoreiden määrästä tai niihin liittyvistä muuttujista
- Konfirmatorinen faktoriansalyysi
 - Teorialähtöinen menetelmä
 - Tutkijalla etukäteen hypoteesit faktorirakenteesta

© Mikko Mattila 2005

4

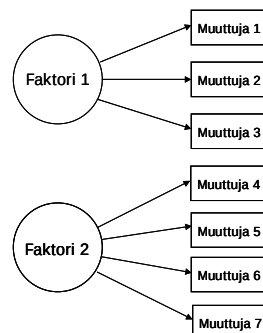
Faktoriansalyysi

- Eksploriatiivinen faktoriansalyysi
 - tavoitteena mittauksen reliabiliteetin ja validiteetin parantaminen sekä informaation tiivistäminen
 - kertoo kuinka monta faktoria muuttujajoukosta löytyy
 - mitkä muuttujat muodostavat ("latautuvat") kunkin faktoriin → tulkinta
 - kuinka paljon faktorit pystyvät selittämään havaittujen muuttujien hajontaa

© Mikko Mattila 2005

5

Faktoriansalyysin perusidea



Faktorit selittävät havaittuja muuttujia.

Kuvio on visualisoinnin vuoksi yksinkertaistettu, oikeastaan kummastakin faktorista tulisi lähteä nuoli jokaiseen havaittuun muuttujaan.

Faktoriansalyysin tuloksena saadaan tieto siitä, kuinka monta faktoria tarvitaan selittämään havaittuja muuttujia ja mitkä muuttujat latautuvat kullakin faktorilla.

© Mikko Mattila 2005

6

Faktorilataukset

- Faktorilataukset kuvaavat havaittujen muuttujien ja faktorin yhteyden vahvuutta (eli edellisen kuvion nuolien "vahvuutta")
- Saavat arvoja -1 ja 1 välillä
- Negatiivinen lataus kertoo että ko. muuttuja korreloi negatiivisesti faktorin kanssa
- Mitä lähempänä faktorilatauksen itseisarvo on yhtä, sitä voimakkaammin havaittu muuttuja latautuu faktorilla
- Ei ole yksikäsitteistä sääntöä sille, kuinka suuri latauksen tulee olla, jotta se olisi merkittävä

© Mikko Mattila 2005

7

Faktorin ominaisarvo

- Ominaisarvot (engl. eigenvalue) kuvaavat sitä, kuinka hyvin faktorit pystyvät selittämään havaittujen muuttujien vaihtelua
- Jatkoanalyysiin otetaan yleensä vain sellaiset faktorit, joiden ominaisarvo on yhtä suurempi (>1)
- Kun ominaisarvo jaetaan muuttujien määrällä saadaan faktorin selitysosuus
- Mitä suurempi selitysosuus on, sitä paremmin faktori pystyy kuvaamaan havaittujen muuttujien hajontaa

© Mikko Mattila 2005

8

Muuttujan kommunaliteetti

- Kommunaliteetit kertovat, kuinka suuri osuus havaittujen muuttujien vaihtelusta pystytään selittämään kaikkien faktoreiden avulla
- Mitä lähempänä yhtä (1) kommunaliteetti on, sitä paremmin faktorit selittävät muuttujan vaihtelua
- Jos muuttujan kommunaliteetti on pieni, kannattaa harkita sen pudottamista pois analyysistä, koska mikään faktoreista ei pysty selittämään sen arvoja

© Mikko Mattila 2005

9

Rotaatio

- Rotaation avulla faktorilatausten saadaan tulkinta yksinkertaisemmaksi
- Rotaatio ei vaikuta tuloksiin, se tekee ne vain helpommin tulkittaviksi → kannattaa aina käyttää
- Rotaatiolla tarkoitetaan faktoriaksien "kiertämistä" siten, että faktorirakenne tulee selkeämmäksi
- Kaksi eri lähestymistapaa: suorakulma- ja vinokulmarotaatio
- Ns. varimax-rotaatio yleisimmin käytetty menetelmä

© Mikko Mattila 2005

10

Esimerkki

Lyhenne	Koko kysymys
sosiaaliturva	Vaatia sairauskorvaus tai sosiaaliturvaetu, johon ei ole oikeutta
kulkuneuvo	Jättää maksamatta julkisessa kulkuneuvossa
verovilppi	Tehdä verovilppiä, jos tilaisuus sallii
varastettu tavara	Ostaa tavaraa, jonka tietää olevan varastettua
lahjukset	Lahjusten ottaminen virkatehtävien hoidossa
homoseksuaalisuus	Homoseksuaalisuus
prostituutio	Prostituutio
abortti	Abortti
avioero	Avioero
eutanasia	Eutanasia, parantamattomasti sairaiden elämän lopettaminen
itsemurha	Itsemurha

"Voisitteko sanoa jokaisesta seuraavaksi luettelemastani asiasta, ovatko ne aina hyväksyttäviä, ei koskaan hyväksyttäviä vai jotain siltä väliltä?"

Vastaajat ilmaisivat hyväksymisensä asteen valitsemalla vastauksensa väliltä 1-10 niin, että arvo 1 tarkoitti "ei koskaan hyväksyttävä" ja arvo 10 "aina hyväksyttävä".

© Mikko Mattila 2005

11

Esimerkki

Muuttuja	Faktori 1 'Vapaamielisyys'	Faktori 2 'Lainkuuliaisuus'	Kommunaliteetti
sosiaaliturva	0,02	0,52	0,27
kulkuneuvo	0,24	0,69	0,53
verovilppi	0,18	0,68	0,50
varastettu tavara	0,14	0,75	0,58
lahjukset	0,03	0,55	0,30
homoseksuaalisuus	0,64	0,03	0,41
prostituutio	0,63	0,23	0,45
abortti	0,80	0,09	0,65
avioero	0,75	0,02	0,56
eutanasia	0,61	0,14	0,39
itsemurha	0,50	0,13	0,27
Ominaisarvo	2,75	2,17	
Selitysosuus	25,0%	19,8%	

Faktorianalyysi tuotti kaksi faktoria.

Ensimmäinen faktori selittää 25% havaittujen muuttujien hajonnasta, toinen 20% eli yhteensä 45%.

Faktorianalyysin keskeisin haaste on löytää faktoreille mielekäs sisällöllinen tulkinta.

© Mikko Mattila 2005

12

Esimerkki

Muuttuja	Faktori 1 'Vapaamielisyys'	Faktori 2 'Lainkuuliaisuus'	Kommunaliteetti
sosiaalitytö	0,02	0,52	0,27
kulkuneuvo	0,24	0,69	0,53
verovilppi	0,18	0,68	0,50
varastettu tavara	0,14	0,75	0,58
lahjukset	0,03	0,55	0,30
homoseksuaalisuus	0,64	0,03	0,41
prostituutio	0,63	0,23	0,45
abortti	0,80	0,09	0,65
avioero	0,75	0,02	0,56
eutanasia	0,61	0,14	0,39
itsemurha	0,50	0,13	0,27
Ominaisarvo	2,75	2,17	
Selitysosuus	25,0%	19,8%	

Faktorin ominaisarvo saadaan laskemalla yhteen kaikkien latausten neliöt.

Ensimmäinen faktori:
 $0,02^2 + 0,24^2 + 0,18^2 + \dots + 0,50^2 = 2,75$

Faktorin selitysosuus saadaan jakamalla ominaisarvo muuttujien määrällä:
 $2,75 / 11 = 0,25$

© Mikko Mattila 2005

13

Esimerkki

Muuttuja	Faktori 1 'Vapaamielisyys'	Faktori 2 'Lainkuuliaisuus'	Kommunaliteetti
sosiaalitytö	0,02	0,52	0,27
kulkuneuvo	0,24	0,69	0,53
verovilppi	0,18	0,68	0,50
varastettu tavara	0,14	0,75	0,58
lahjukset	0,03	0,55	0,30
homoseksuaalisuus	0,64	0,03	0,41
prostituutio	0,63	0,23	0,45
abortti	0,80	0,09	0,65
avioero	0,75	0,02	0,56
eutanasia	0,61	0,14	0,39
itsemurha	0,50	0,13	0,27
Ominaisarvo	2,75	2,17	
Selitysosuus	25,0%	19,8%	

Kommunaliteetti kertoo, kuinka hyvin kaikki faktorit pystyvät selittämään yksittäistä havaittua muuttujaa.

"Sosiaalitytö"-muuttujan kommunaliteetti:
 $0,02^2 + 0,52^2 = 0,27$

© Mikko Mattila 2005

14

Faktoripisteet

- Faktoripisteet kuvaavat yksittäisten havaintojen sijoittumista faktoreilla
- Pisteet lasketaan painotettuna keskiarvona havaittujen muuttujien arvoista, painoina käytetään faktorilatauksia (lisäksi pisteet "standardisoidaan" niin, että keskiarvo=0 ja keskihajonta=1)
- Voidaan käyttää jatkoanalyysissä joko selitettävänä tai selittävinä muuttujina
- Muuttujista voidaan myös muodostaa summamuuttujia faktorianalyysin tulosten pohjalta

© Mikko Mattila 2005

15

Faktoripisteet

Ryhmä	Keskiarvo ensimmäisellä faktorilla ('Vapaamielisyys')	Keskiarvo toisella faktorilla ('Lainkuuliaisuus')
Alle 35-vuotiaat	0,17	0,17
60-vuotta tai yli	-0,43	-0,34
Nainen	-0,01	-0,14
Mies	0,02	0,13
Ei ammatillista koulutusta	-0,15	0,20
Korkeakoulututkinto	0,20	-0,14

© Mikko Mattila 2005

16

Konfirmatorinen faktorianalyysi

- Lähtökohtana teoriaan perustuva oletus aineiston faktorirakenteesta
- Etukäteishypoteesit faktorien määrästä ja niillä latautuvista muuttujista
- Analyysin tuloksien avulla voidaan arvioida, kuinka hyvin oletukset pitivät paikkansa

© Mikko Mattila 2005

17

Esimerkki

- Eurobarometri 55 (2001):
- In general, do you pay attention to news about each of the following?
 - Politics, Social issues (such as education, health care...), European Union, Economy, Sport, Environment, International affairs, Culture
 - 1 = a lot of attention
 - 2 = a little attention
 - 3 = no attention at all

© Mikko Mattila 2005

18

Esimerkki

Communalities		
	Initial	Extraction
Q12 ATTENTION TO NEWS: POLITICS	.398	.520
Q12 ATTENTION TO NEWS: SOCIAL ISSUES	.290	.335
Q12 ATTENTION TO NEWS: EUROPEAN UNION	.386	.479
Q12 ATTENTION TO NEWS: ECONOMY	.303	.396
Q12 ATTENTION TO NEWS: SPORT	4.579E-02	2.936E-02
Q12 ATTENTION TO NEWS: ENVIRONMENT	.248	.287
Q12 ATTENTION TO NEWS: INTERNAT AFFAIRS	.442	.536
Q12 ATTENTION TO NEWS: CULTURE	.262	.572

Extraction Method: Principal Axis Factoring.

"Extraction" –sarake kertoo muuttujien kommunaliteetin sen jälkeen kun valittu määrä faktoreita on "uutettu" aineistosta.

© Mikko Mattila 2005

19

Esimerkki

Total Variance Explained									
Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.257	40.716	40.716	2.714	33.919	33.919	1.671	20.885	20.885
2	1.063	13.292	54.008	.441	5.510	39.429	1.484	18.545	39.429
3	.927	11.581	65.590						
4	.670	8.369	73.959						
5	.610	7.621	81.579						
6	.559	6.991	88.570						
7	.498	6.230	94.800						
8	.416	5.200	100.000						

Extraction Method: Principal Axis Factoring.

Alussa SPSS-ohjelma laskee aineistosta yhtä monta faktoria kuin analyysiin on valittu muuttujia. Näistä faktoreista kahdella on ominaisarvo suurempi kuin yksi. Aineistoa kuvaa siis parhaiten kahden faktorin ratkaisu. Nämä kaksi faktoria pystyvät selittämään 39% havaittujen muuttujien hajonnasta.

© Mikko Mattila 2005

20

Esimerkki

	Factor	
	1	2
Q12 ATTENTION TO NEWS: INTERNAT AFFAIRS	.731	
Q12 ATTENTION TO NEWS: POLITICS	.680	
Q12 ATTENTION TO NEWS: EUROPEAN UNION	.679	
Q12 ATTENTION TO NEWS: ECONOMY	.593	
Q12 ATTENTION TO NEWS: SOCIAL ISSUES	.572	
Q12 ATTENTION TO NEWS: CULTURE	.559	.509
Q12 ATTENTION TO NEWS: ENVIRONMENT	.503	
Q12 ATTENTION TO NEWS: SPORT		

Extraction Method: Principal Axis Factoring.

a. Attempted to extract 2 factors. More than 25 iterations required. (Convergence=2.684E-03). Extraction was terminated.

Tässä on faktorianalyysin tulokset ennen rotaatiota. Niitä ei kuitenkaan kannata tuijotella liikaa, koska tulokset ovat paljon selkeämmät rotaation jälkeen.

© Mikko Mattila 2005

21

Esimerkki

	Factor	
	1	2
Q12 ATTENTION TO NEWS: CULTURE	.757	
Q12 ATTENTION TO NEWS: ENVIRONMENT	.496	
Q12 ATTENTION TO NEWS: SOCIAL ISSUES	.479	
Q12 ATTENTION TO NEWS: POLITICS		.637
Q12 ATTENTION TO NEWS: EUROPEAN UNION	.411	.557
Q12 ATTENTION TO NEWS: ECONOMY		.556
Q12 ATTENTION TO NEWS: INTERNAT AFFAIRS	.503	.532
Q12 ATTENTION TO NEWS: SPORT		

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Tuloksista voi päätellä, että suomalaiset uutisten seuraajat jakautuvat niihin, jotka seuraavat "kovia" uutisia ja niihin jotka seuraavat "pehmeitä" uutisia.

Urheilun seuraajat eivät kuulu kumpaankaan joukkoon. He muodostavat luultavasti oman faktorinsa. Jos näin on, oheinen analyysi ei pysty sitä kertomaan, koska faktorianalyysissä etsitään muuttujia, jotka korreloivat keskenään.

© Mikko Mattila 2005

22

Esimerkki2

- EB55 (2001)
- Muodostaako suomalaisten luottamus eri EU-instituutioihin yhden vai useita faktoreita?
- Vastajilta kysyttiin luottamusta yhdeksään EU:n toimielimeen (vaihtoehdot: luottaa=1, ei luota=2)
 - Euroopan parlamentti, komissio, neuvosto, oikeusistuin, oikeusasiamies, keskuspankki, tilintarkastustuomioistuin, alueiden komitea, sosiaali- ja talouskomitea

© Mikko Mattila 2005

23

Esimerkki2

Total Variance Explained						
Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7.060	78.440	78.440	6.824	75.825	75.825
2	.526	5.843	84.284			
3	.330	3.671	87.955			
4	.274	3.040	90.995			
5	.237	2.629	93.624			
6	.189	2.101	95.724			
7	.164	1.825	97.549			
8	.152	1.693	99.243			
9	6.814E-02	.757	100.000			

Extraction Method: Principal Axis Factoring.

Kaikki havaitut muuttujat latautuvat samalla faktorilla, joka selittää 76% hajonnasta. Siis suomalaiset eivät erottele instituutioita niihin kohdistuvan luottamuksen perusteella. Jos luottaa yhteen, todennäköisesti luottaa myös muihin ja päinvastoin.

© Mikko Mattila 2005

24

Esimerkki2

Factor Matrix^a

	Factor 1
Q24 SOC&ECONOM COMMITTEE - TRUST	.917
Q24 COMMITTEE OF REGIONS - TRUST	.907
Q24 COUNCIL OF MINISTERS - TRUST	.904
Q24 EUROPEAN PARLIAMENT - TRUST	.892
Q24 EU COURT OF AUDITORS - TRUST	.888
Q24 EUROPEAN COMMISSION - TRUST	.875
Q24 EUROPEAN CENTRAL BANK - TRUST	.849
Q24 EU COURT OF JUSTICE - TRUST	.811
Q24 EUROPEAN OMBUDSMAN - TRUST	.785

Extraction Method: Principal Axis Factoring.
a. 1 factors extracted. 4 iterations required.

Koska tuloksena oli vain yksi faktori, rotaatiota ei tarvita.

© Mikko Mattila 2005

25

Esimerkki3

- Ympäristökysely 2000
- Mitä mieltä olette seuraavista väitteistä? (skaala 1-5, ääripäät "täysin samaa mieltä" ja "täysin eri mieltä")
 - Minun kaltaiseni ihmisen on kerta kaikkiaan liian vaikeaa tehdä juuri mitään ympäristön hyväksi
 - Teen mikä on oikein ympäristön kannalta, vaikka se maksaa enemmän ja vie enemmän aikaa
 - Elämässä on tärkeämpiäkin asioita kuin ympäristönsuojelu
 - Minun ei ole järkevää toimia ympäristön puolesta, jos muut eivät toimi samoin
 - Monet väitteet ympäristöuhista ovat liioiteltuja

© Mikko Mattila 2005

26

Esimerkki3

Total Variance Explained

Fact	Initial Eigenvalues			Total Sums of Squared Loadings			Total Sums of Squared Loadings		
	Total	of Variance	Initial Eigenvalues	Total	of Variance	Initial Eigenvalues	Total	of Variance	Initial Eigenvalues
1	1.928	38.567	38.567	1.329	26.578	26.578	.919	18.384	18.384
2	1.019	20.387	58.954	.484	9.680	36.258	.894	17.874	36.258
3	.861	17.213	76.167						
4	.617	12.334	88.500						
5	.575	11.500	100.000						

Extraction Method: Principal Axis Factoring.

© Mikko Mattila 2005

27

Esimerkki3

Rotated Factor Matrix^a

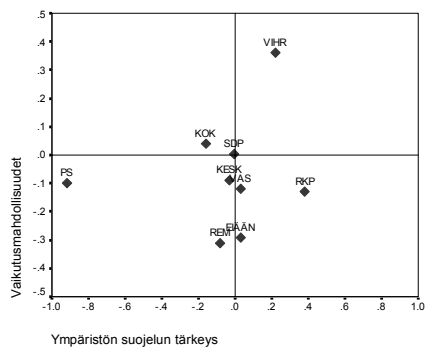
	Factor	
	1	2
Minun ei ole järkevää toimia ympäristön puolesta, jos muut eivät toimi samoin?	.678	
Minun kaltaiseni ihmisen on kerta kaikkiaan liian vaikeaa tehdä juuri mitään ympäristön hyväksi?	.576	
Elämässä on tärkeämpiäkin asioita kuin ympäristönsuojelu?		.762
Monet väitteet ympäristöuhista ovat liioiteltuja?		.426
Teen mikä on oikein ympäristön kannalta, vaikka se maksaa enemmän ja vie enemmän aikaa?		

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

© Mikko Mattila 2005

28

Esimerkki3



© Mikko Mattila 2005

29