



KÄYTTÄYTYMISTIEIDEIN

.....

TILASTOLLISET
MENETELMÄT

HARJOITUSKIRJA

.....

LAURI NUMMENMAA

TAMMI



Lauri Nummenmaa

Käyttäytymistieteiden tilastolliset menetelmät

HARJOITUSKIRJA

Kustannusosakeyhtiö Tammi
Helsinki

Lauri Nummenmaa

Käyttäytymistieteiden tilastolliset menetelmät

HARJOITUSKIRJA

Kustannusosakeyhtiö Tammi
Helsinki

TAMMI
Oppimateriaalit
www.tammi.fi/oppimateriaalit

1. painos (uud. digitaalinen laitos)

© Lauri Nummenmaa ja Kustannusosakeyhtiö Tammi, 2023

Kopiointikielto

Tämä teos on työkirja, jota ei saa tekijänoikeuslain nojalla valokopioida lainkaan. Myös teoksen tai sen osan kopioiminen tai muuntelu digitaalisesti on ehdottomasti kielletty. Lisätietoja antaa Kopiosto ry.

ISBN 978-951-31-5153-9

Digitaalinen uusintapainos
Helsinki 2023

Sisällys

Lukijalle	5
-----------------	---

HARJOITUKSET

I Tutkimusaineisto	7
1 Mitä tilastotiede on?	8
2 Tutkimuksen suunnitteleminen	8
3 Mittaaminen ja aineisto	10
5 Aineiston kuvaaminen numeerisesti	11
6 Aineiston kuvaaminen graafisesti	13
II Todennäköisyys ja jakaumat	15
7 Todennäköisyyslaskennan perusteet	16
8 Todennäköisyysjakaumat	18
9 Tilastolliset testit ja hypoteesien testaaminen	21
III Keskiarvoja koskeva päättely	23
10 Aineiston valmistelu	24
11 T-testit	24
12 Yksisuuntainen varianssianalyysi	27
13 Useampisuuntainen varianssianalyysi	29
14 Toistettujen mittausten varianssianalyysi	32
15 Epäparametriset menetelmät	34
IV Yhteisvaihtelun tutkiminen	37
16 Kovarianssi ja korrelaatio	38
17 Kategoristen muuttujien välisten yhteyksien tarkasteleminen	41
18 Lineaarinen regressio	43
19 Logistinen regressio	46
V Latentit muuttujat	49
20 Faktorianalyysi	50
21 Moniulotteinen asteikointi	52
22 Klusterianalyysi	52
23 Rakenneytälömallit	54

VASTAUKSET

I Tutkimusaineisto	58
1 Mitä tilastotiede on?	58
2 Tutkimuksen suunnitteleminen	58
3 Mittaaminen ja aineisto	59
5 Aineiston kuvaaminen numeerisesti	60
6 Aineiston kuvaaminen graafisesti	61
II Todennäköisyys ja jakaumat	63
7 Todennäköisyyslaskennan perusteet	63
8 Todennäköisyysjakaumat	64
9 Tilastolliset testit ja hypoteesien testaaminen	65
III Keskiarvoja koskeva päättely	67
10 Aineiston valmistelu	67
11 T-testit	67
12 Yksisuuntainen varianssianalyysi	69
13 Useampisuuntainen varianssianalyysi	71
14 Toistettujen mittausten varianssianalyysi	74
15 Epäparametriset menetelmät	76

IV Yhteisvaihdelun tutkiminen.....	78
16 Kovarianssi ja korrelaatio	78
17 Kategoristen muuttujien välisten yhteyksien tarkasteleminen	81
18 Lineaarinen regressio.....	82
19 Logistinen regressio	84
V Latentit muuttujat	86
20 Faktorianalyysi	86
21 Moniulotteinen asteikointi	88
22 Klusterianalyysi	89
23 Rakenneyhtälömallit.....	89
LÄHTEET	92

AINEISTOT

Liite 1: Lukeminen ja työmuisti	94
Liite 2: Koulukiusaaminen.....	95
Liite 3: Tarkkaavaisuuden suuntautuminen katsevihjeiden perusteella.....	96
Liite 4: Emootioiden kokeminen.....	97
Liite 5: Työn kuormittavuus	98
Liite 6: Emootiot ja visuaalinen tarkkaavaisuus	99
Liite 7: Stroopin värisanatehtävä	100
Liite 8: Depression hoitaminen.....	101
Liite 9: Sokeiden työmuisti	102
Liite 10: Kasvoinversio	103
Liite 11: Työuupumus	104
Liite 12: Prosessointikontekstin vaikutus luetun ymmärtämiseen.....	105

Lukijalle

Vaikka tutkimusaineistojen analyysi tehdään nykyisin poikkeuksetta tietokoneella, tilastollisten menetelmien matemaattisen perustan ymmärtäminen on välttämätöntä, jotta menetelmiä osaa käyttää oikein. Periaatteita opeteltaessa käsin laskeminen onkin yhä erinomainen tapa opiskella tilastollisen analyysin perusteita.

Nämä harjoitustehtävät on laadittu *Käyttätymistieteiden tilastolliset menetelmät* -kirjaa täydentäväksi oppimateriaaliksi. Harjoitusten järjestys noudattaa kirjan rakennetta. Pääpaino on päättelytehtävissä ja käsinlaskuharjoituksissa, ja harjoitusten laatimisessa on lähdetty siitä, että kaikkien harjoitusten tekeminen on mahdollista ilman tietokoneohjelmistoja. Täten aivan kaikista *Käyttätymistieteiden tilastolliset menetelmät* -kirjassa käsitellyistä aihepiireistä ei ole harjoituksia. Toisaalta suurin osa aineistojen analysoimistehtävistä soveltuu myös SPSS:llä tai vastaavalla tilasto-ohjelmistolla tehtäväksi. Kirjaa voi siis käyttää apuna myös tilastollisten ohjelmistojen käyttöä opiskeltaessa.

Näitä harjoituksia tekemällä oppii sekä tilastollisen päättelyn että käyttätymistieteellisen tutkimuksen periaatteita. Harjoitustehtäväpakettiin kuuluu 12 tutkimusesimerkkiä ja harjoitusaineistoa, joita käytetään suurimmassa osassa laskutehtäviä. Tutkimusesimerkit on laadittu käyttätymistieteiden alan todellisia tutkimusraportteja ja tuloksia mukaillen, mutta kaikki kirjassa käytetyt aineistot on muodostettu simuloimalla. Laskettaessa harjoituksia tietokoneella aineistot voi syöttää suoraan esimerkiksi SPSS-ohjelmaan. Vaihtoehtoisesti aineistopakettin voi ladata kirjan verkkosivulta

<http://emotion.utu.fi/KTM>

Aineistot on saatavana sekä SPSS-muodossa että tabulaattorierotettuina tekstitiedostoina, joita lähes kaikki tilastolliset ohjelmat osaavat lukea. Tietokoneella laskettaessa ainakin ensimmäiset aineistot kannattaa syöttää harjoitusmielessä käsin ja kokemuksen karttuessa loput aineistot voi ladata suoraan koneelle. Ennen laskuharjoitusten tekemistä on syytä tutustua kaikkiin aineistoihin ja niissä tarkasteltuihin tutkimuskysymyksiin huolellisesti, jotta ymmärtää, millaisilla tutkimusasetelmilla aineistot on kerätty ja millaisia analyyseja aineistoille voi tehdä.

Turussa 3.4.2023

Tekijä

Tutkimusaineisto



Mitä tilastotiede on?

1. Mikä on tilastotieteen tutkimuskohde?
2. Mikä on tieteellisen tutkimuksen tavoite?
3. Miksi tilastollisia menetelmiä tarvitaan tutkimusaineiston analysoimiseen?
4. Mieti esimerkkejä tutkimuksista, joiden tulosten analysoiminen olisi hankalaa ilman tilastollisia menetelmiä.
5. Mihin kolmeen luokkaan tilastolliset menetelmät voidaan karkeasti jakaa?

2 Tutkimuksen suunnitleminen

1. Millainen on tiedon hierarkkinen rakenne?
2. Mikä on Platonin perinteinen tiedon määritelmä?
3. Mieti esimerkkejä tutkimuksista, jotka ovat
 - a) tieteellisiä tutkimuksia
 - b) epätieteellisiä tutkimuksia.
4. Miksi tilastollinen tutkimus perustuu usein ilmiöiden tarkastelemiseen otoksessa populaation sijaan?
5. Miksi ja millä edellytyksellä otoksen perusteella voi tehdä populaatiota koskevia päätelmiä?
6. Tutkimuksen populaationa ovat kaikki 250 000 depressiosta kärsivää suomalaista. Poimitaan heistä
 - a) 250
 - b) 500
 - c) 5 000 hengen suuruiset otokset.Mitkä ovat otantasuhteet?

7. Tarkastellaan lukemistutkimuksen aineistoa (liite 1). Poimitaan aineistosta

- a) kymmenen
- b) viiden hengen suuruiset otokset systemaattisella otannalla.

Aloitetaan otanta aina ensimmäisestä koehenkilöstä. Millaiset poimintavälit saadaan? Ketkä aineiston henkilöistä tulevat poimituiksi otoksiin?

8. Millaisissa tilanteissa tarvitaan suuria otoksia? Milloin taas riittävät pienet otokset?

9. Tarkastele

- a) työmuistitutkimusta (liite 9)
- b) lukemistutkimusta (liite 1)
- c) koulukiusaamistutkimusta (liite 2)
- d) tarkkaavaisuuden suuntautumistutkimusta (liite 3).

Mitkä tutkimuksista on tehty kokeellisella ja mitkä korrelatiivisella asetelmalla? Määritä tutkimusten riippuvat ja riippumattomat muuttajat.

10. Miten toteuttaisit vaihe vaiheelta tutkimuksen, jossa selvitetään

- a) koululaisten kiusatuksi tulemisen syitä
- b) valaistuksen vaikutusta pupillin kokoon
- c) dementian esiintyvyyden eroja miehillä ja naisilla
- d) tekstin vaikeuden vaikutusta tekstin lukemiseen käytettävään aikaan?

11. Mitkä olisivat edellisten tutkimusten riippuvat ja riippumattomat muuttajat?

3 Mittaaminen ja aineisto

1. **Mitä mittaaminen tarkoittaa?**
2. **Miten kvalitatiiviset ja kvantitatiiviset muuttujat eroavat toisistaan?**
3. **Jaottele seuraavat muuttujat kvalitatiivisiin ja kvantitatiivisiin muuttujiin: sukupuoli, reaktioaika, koulutus, oppilaan luokka-aste.**
4. **Määrittele, millä mitta-asteikolla eri muuttujat on mitattu**
 - a) koulukiusaamistutkimuksessa (liite 2)
 - b) työn kuormittavuustutkimuksessa (liite 5)
 - c) depressiotutkimuksessa (liite 8).
5. **Millaisilla mitta-asteikoilla voidaan mitata**
 - a) laatueroasteikollisia muuttujia
 - b) järjestysasteikollisia muuttujia
 - c) välimatka-asteikollisia muuttujia
 - d) suhdeasteikollisia muuttujia?
6. **Minkä muunnosten suhteen seuraavat mitta-asteikot ovat invariantteja?**
 - a) laatueroasteikko
 - b) järjestysasteikko
 - c) välimatka-asteikko
 - d) suhdeasteikko
7. **Mikä on perussääntö havaintojen syöttämisestä datamatriisiin?**
8. **Tarkastele**
 - a) lukemisen ja työmuistin tutkimusta (liite 1)
 - b) koulukiusaamistutkimusta (liite 2)
 - c) emootioiden kokemistutkimusta (liite 4)
 - d) Stroopin värisanatehtävätutkimusta (liite 7).

Kuinka monta tilastoyksikköä, muuttujaa ja havaintoa kussakin aineistossa on?

5

Aineiston kuvaaminen numeerisesti

1. Mitä tarkoitusta varten kuvailevia tunnuslukuja yleensä lasketaan?

2. Mitä eroa on parametreilla ja estimaateilla?

3. Mitä hajontalukuja voidaan laskea

- a) laatueroasteikollisista muuttujista
- b) järjestysasteikollisista muuttujista
- c) välimatka-asteikollisista muuttujista
- d) suhdeasteikollisista muuttujista?

4. Mitä sijaintilukuja voidaan laskea

- a) laatueroasteikollisista muuttujista
- b) järjestysasteikollisista muuttujista
- c) välimatka-asteikollisista muuttujista
- d) suhdeasteikollisista muuttujista?

5. Laske

a) $\sum_{k=1}^{10} k$

d) $\sum_{i=1}^5 (5i-1)$

b) $\sum_{t=1}^7 5t$

e) $\sum_{u=1}^{10} (10-u)$

c) $\sum_{k=1}^{15} (k-10+2k)$

6. Millaisissa tapauksissa mediaanin ja moodin laskemisesta on hyötyä, vaikka muuttuja olisikin mitattu suhdeasteikolla?

7. Miten jakauman moodi ja keskiarvo suhteutuvat toisiinsa, jos jakauma on yksihuippuinen ja

- a) symmetrinen
- b) oikealle vino
- c) vasemmalle vino?

8. **Tutkimuksessa on selvitetty kymmenen aikuisen leposykkeitä (bpm). Sykkeet ovat 64, 67, 52, 56, 57, 60, 54, 49, 59 ja 60. Määritä sykkeiden**
 - a) mediaani
 - b) otosvarianssi
 - c) populaatiokeskihajonta
 - d) keskiarvo
 - e) variaatiokerroin
 - f) minimi ja maksimi.
9. **Tutkimuksessa selvitettiin kymmenen suomalaisen aikuisen viikoittaiseen liikuntaan käyttämä aika minuutteina. Havainnot ovat 125, 100, 65, 20, 95, 160, 30, 45, 20 ja 10. Laske viikoittaisten liikunta-aikojen**
 - a) moodi
 - b) otoskeskihajonta
 - c) keskiarvo
 - d) populaatiovarianssi
 - e) vaihteluväli.
10. **Tarkastellaan koulukiusaamistutkimusta (liite 2). Määritä**
 - a) tyttöjen ja poikien frekvenssit
 - b) sisarusten lukumäärä -muuttujan vaihteluväli ja mediaani
 - c) sukupuoli-muuttujan moodi
 - d) sisarusten lukumäärä -muuttujan moodi ja keskiarvo
 - e) kiusatuksi tulemisen kokemus-muuttujan keskiarvo ja otoskeskihajonta
 - f) itsetunto-muuttujan keskiarvo ja otosvarianssi.
11. **Tarkastellaan emootioiden kokemus -aineistoa (liite 4). Määritä**
 - a) sukupuolen ja kätisyyden moodit
 - b) iän mediaani ja variaatiokerroin
 - c) mielihyvän kokemisen otosvarianssi, otoskeskihajonta ja variaatiokerroin
 - d) eri emootioiden (mielihyvä, viha, inho, pelko, suru, hämmästys) kokemisen keskiarvot ja mediaanit.

12. Tarkastellaan prosessointikontekstiaineistoa (liite 12).
Määrittele
- a) koulutuksen luokkafrekvenssit
 - b) sukupuolen moodi
 - c) kokonaispalautuksen minimi ja maksimi
 - d) tekstin ymmärtämisen mediaani
 - e) kokonaispalautuksen moodi.
13. Tarkastele kasvoinversioaineistoa. Laske kasvojen tunnistamisprosenttien
- a) mediaanit
 - b) moodit
 - c) keskiarvot eri tilanteissa.
14. Luo viiden hengen muuttujat, jotka täyttävät seuraavat ehdot:
- a) $\bar{x} = 5, MD = 5, MO = 5$
 - b) $\bar{x} = 5, MD = 3, MO = 2$

6 Aineiston kuvaaminen graafisesti

1. Millaisissa tilanteissa
 - a) sektoridiagrammi
 - b) histogrammi soveltuvat käytettäväksi?
2. Millaiset tekijät hankaloittavat kuvaajien tulkitsemista?
3. Mitä kuvaajia voidaan käyttää
 - a) kategoristen muuttujien
 - b) jatkuvien muuttujien kuvaamiseen?
4. Tutkimukseen on osallistunut 60 koululaista, joista 35 on tyttöjä ja 25 poikia. Esitä sukupuolijakauma pylväsdiagrammin avulla.

5. **Tarkastellaan koulukiusaamisaineistoa (liite 2). Esitä**
 - a) sukupuolimuuttujan jakauma pylväsdiagrammin avulla
 - b) itsetuntomuuttuja histogrammin avulla
 - c) kiusatuksi tulemisen kokemus -muuttuja laatikko–jana-kuvion avulla
 - d) poikien ja tyttöjen itsetunnot ja kiusatuksi tulemisen kokemukset samassa laatikko–jana-kuviossa.
6. **Tarkastellaan prosessointikontekstiaineistoa. Piirrä**
 - a) sektoridiagrammi sukupuolen jakaumasta
 - b) tekstin ymmärtämisen ja kokonaispalautuksen jakaumat samassa laatikko–jana-kuviossa.
7. **Tarkastellaan katsevihjeet ja tarkkaavaisuuden suuntautuminen -aineistoa. Piirrä**
 - a) pylväsdiagrammi
 - b) viivadiagrammi reaktioaikojen keskiarvoista eri tilanteissa.

Todennäköisyys ja jakaumat



7 Todennäköisyyslaskennan perusteet

1. **Määrittele, ovatko seuraavat tapahtumat deterministisiä vai stokastisia.**
 - a) Rahan heiton lopputulos
 - b) Syntyvän lapsen sukupuoli, jos hänellä tiedetään olevan x- ja y-kromosomit
 - c) Pakasta satunnaisesti vedettävän kortin maa
 - d) Pelkistä ruutukorteista koostuvasta pakasta vedettävän kortin maa

2. **Sovitaan, että kaikki maailman ihmiset muodostavat perusjoukon E. Tarkastellaan muutamia luonnollisesti olemassa olevia ryhmiä ja käsitellään niitä joukkoina. Onko**
 - a) miesten joukko kaikkien maailman ihmisten osajoukko
 - b) miesten ja naisten joukon unioni tyhjä joukko
 - c) miesten ja naisten joukon unioni perusjoukko?

3. **Määritellään joukot $A = \{q, w, e, r, t\}$ ja $B = \{a, b, c, d, e\}$. Pitävätkö seuraavat väitteet paikkansa?**
 - a) $q \in A$
 - b) $e \notin A$
 - c) $e \notin B$
 - d) $b \in B$

4. **Määritellään joukot $E = \{a, b, c, d, e, f, g, z, x, y\}$, $A = \{b, c, d, f, g, z, x\}$, $B = \{b, c, d, y\}$, ja $C = \{a, e, y\}$. Onko**
 - a) joukko B joukon A osajoukko
 - b) joukko E tyhjä joukko
 - c) joukko C joukon A komplementti
 - d) joukko C joukon B komplementti?

5. **Määritellään joukot $E = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $A = \{2, 4, 6, 8, 10\}$, $B = \{2, 8, 10\}$ ja $D = \{7, 9\}$. Määrittele**
 - a) $\overline{A}$
 - b) $A \cap B$
 - c) $A \cup D$
 - d) $E - B$

6. **Todennäköisyys diagnosoida skitsofrenia oikein Rorschachin mustetahrasteillä on 0.3. Mikä on todennäköisyys sille, että Rorschachin mustetahrasteillä ei diagnosoida skitsofreniaa oikein?**
 - a) on pata
 - b) on herttaässä
 - c) on ruutu tai nelonen
 - d) ei ole herttakymppi eikä ruutu
 - e) ei ole pata tai vitonen?
7. **Mikä on todennäköisyys sille, että satunnaisesti pakasta vedettävä kortti**
 - a) on pata
 - b) on herttaässä
 - c) on ruutu tai nelonen
 - d) ei ole herttakymppi eikä ruutu
 - e) ei ole pata tai vitonen?
8. **Heitetään kahta noppaa samanaikaisesti. Mikä on todennäköisyys saada**
 - a) vähintään yksi kutonen
 - b) tasan kaksi kuutosta
 - c) tasan kolme viitosta
 - d) silmälukujen summaksi vähintään 5
 - e) silmälukujen summaksi vähintään 2?
9. **Psykiatri tietää, että hänen yleisimmin määräämänsä rauhoittava lääke tehoaa 80 %:iin potilaista. Mikä on todennäköisyys sille, että lääke tehoaa**
 - a) kolmeen
 - b) neljään
 - c) viiteen seuraavaan potilaaseen, joille sitä määrätään?
10. **Silmänliikerekisteröinnissä koehenkilöltä kerätty data on käyttökelpoista 97 %:n todennäköisyydellä. Mikä on todennäköisyys sille, että käyttökelpoista dataa saadaan peräkkäin**
 - a) kahdelta
 - b) kuudelta
 - c) kymmeneltä henkilöltä?
11. **Reaktioaikamittauksissa koehenkilöltä kerätty data on käyttökelpoista 94 %:n todennäköisyydellä. Mikä on todennäköisyys sille, että käyttökelpoista dataa saadaan peräkkäin**
 - a) kahdelta
 - b) kolmelta
 - c) viideltäkymmeneltä henkilöltä?

12. Tutkimukseen osallistuu

- a) kolme
- b) viisi
- c) kahdeksan
- d) kymmenen koehenkilöä.

Kuinka monessa eri järjestyksessä tutkittavat on mahdollista testata?

13. Lasten oppimisvaikeuksia arvioitaessa on mahdollista käyttää kymmentä erilaista testiä. Kuinka monta erilaista testien yhdistelmää psykologi voi valita, jos testien tekemisjärjestys huomioidaan ja hän käyttää tutkimuksessaan

- a) kolmea
- b) viittä
- c) kahdeksaa
- d) kymmentä testiä?

14. Kuinka monella eri tavalla psykologi voi suorittaa edellisen esimerkin testauksen, jos testien tekemisjärjestystä ei huomioida ja hän käyttää tutkimuksessaan

- a) neljää
- b) viittä
- c) seitsemää testiä?

8 Todennäköisyysjakaumat

1. Heitetään samanaikaisesti kahta noppaa ja lasketaan tulos yhteen. Mikä on todennäköisyys saada tulokseksi

- a) enintään 10
- b) vähintään 12
- c) mikä tahansa tulos väliltä 6–8
- d) enintään 2?

2. Mikä tulkinta on diskreetin satunnaismuuttujan

- a) odotusarvolla
- b) varianssilla?

3. **Heitetään samanaikaisesti kahta noppaa ja lasketaan tulos yhteen. Mikä on tuloksen**
 - a) odotusarvo
 - b) varianssi
 - c) keskihajonta?

4. **Tarkastellaan, kuinka taloudellisesti kannattavaa neljän eri-hintaisen psykoterapiahoitomuodon tarjoaminen on Kansaneläkelaitokselle (Kela). Jos potilas paranee, Kela säästää tulevana hoitokuluina 10 000 euroa. Jos hän ei parane, Kela menettää ainoastaan hoitoon käytetyt rahat.**

Mikä on Kelan rahankulutuksen tai -säästön odotusarvo seuraavissa hoitomuodoissa? Mitä terapiaa Kelan kannattaisi tulosten perusteella suosia?

 - a) Psykoanalyysi, jonka kustannukset ovat 50 000 € ja paranemisprosentti 80
 - b) Ratkaisukeskeinen terapia, jonka kustannukset ovat 10 000 € ja paranemisprosentti 75
 - c) Kognitiivinen lyhytterapia, jonka kustannukset ovat 5 000 € ja paranemisprosentti 70
 - d) Gestalt-terapia, jonka kustannukset ovat 25 000 € ja paranemisprosentti 80.

5. **Noin 1.5 % suomalaisista kärsii skitsofreniasta. Valitaan tutkimukseen satunnaisesti 200 koehenkilöä. Mikä on otokseen valittujen skitsofreenisten henkilöiden lukumäärän**
 - a) odotusarvo
 - b) varianssi
 - c) keskihajonta?

6. **Sovitaan, että tasan puolet suomalaisista on miehiä ja puolet naisia. Mikä on todennäköisyys sille, että satunnaisesti valitussa sadan hengen otoksessa on**
 - a) 50 naista
 - b) 45 miestä
 - c) 10 miestä
 - d) enintään 3 naista?

7. **Työnantajan tekemissä työuupumuksen seurannoissa diagnosoidaan keskimäärin kymmenen työuupumustapausta kuukaudessa. Mikä on todennäköisyys sille, että kuukaudessa diagnosoidaan**
 - a) 5
 - b) 3
 - c) 12 työuupumustapausta?

8. EEG-rekisteröinnissä käytettäviä elektrodeja hajoaa laboratorioissa keskimäärin 12 kappaletta vuodessa. Mikä on todennäköisyys sille, että vuodessa menee rikki tasan
 - a) 10
 - b) 15
 - c) 100 elektrodia?
9. Noin 0.01 % Suomen väestöstä kärsii prosopagnosiasta (kasvojen tunnistamisen vaikeudesta). Noin 10 000 suomalaista aloittaa vuosittain peruskoulun. Kuinka suuri todennäköistä on, että vuosittain peruskoulun aloittaneista
 - a) 2
 - b) 5
 - c) 10 kärsii prosopagnosiasta?
10. Kuinka paljon standardoidussa normaalijakaumassa on sellaisia havaintoja, jotka ovat
 - a) yli kaksi keskihajontaa keskiarvoa suurempia
 - b) yli keskihajonnan verran keskiarvoa pienempiä
 - c) yli kahden keskihajonnan päässä keskiarvosta
 - d) yli keskihajonnan päässä keskiarvosta?
11. Mitä ilmiöitä mallintavat
 - a) normaalijakauma
 - b) t-jakauma
 - c) χ^2 -jakauma
 - d) F-jakauma?
12. Mikä on
 - a) t-jakauman
 - b) F-jakauman odotusarvo?

9

Tilastolliset testit ja hypoteesien testaaminen

1. Mitä tarkoitetaan otantavirheellä?
2. Miksi keskeinen raja-arvolause on tärkeä populaatiokeskiarvoja koskevassa päättelyssä?
3. Suomalaisten aikuisten miesten veren kolesterolipitoisuus on keskimäärin 5.5 mmol/l keskihajonnan ollessa 2. Poimitaan suomalaisista aikuisista miehistä satunnaisesti
 - a) 10
 - b) 15
 - c) 50
 - d) 100
 - e) 1 000 henkilön otokset.

Laske keskiarvon keskivirheet näille otoksille ja hahmottele viivadiagrammin avulla keskiarvon keskivirheen muuttumista kun otoskoko kasvaa.

4. Viiden hengen afasiapotilaspopulaatiossa keskimääräiset esineiden nimeämisajat ovat 1 110 ms, 1 550 ms, 1 320 ms, 1 400 ms ja 1 450 ms.
 - a) Laske populaation keskiarvo.
 - b) Muodosta populaatiosta kaikki mahdolliset neljän hengen otokset.
 - c) Laske otosten keskiarvot.
 - d) Määritä keskiarvon keskivirhe neljän hengen suuruisille tästä populaatiosta poimituille otoksille.
5. Tarkastellaan populaatioita A ja B, joiden hajonnat ovat $\sigma_A = 10$ ja $\sigma_B = 20$. Poimitaan populaatiosta A 10 hengen ja populaatiosta B 50 hengen otos. Kumman otoksen perusteella voidaan todennäköisesti estimoida tarkemmin sen populaation keskiarvoa, josta otos on poimittu?
6. Selvitä standardoidun normaalijakauman kertymäfunktion taulukoitujen arvojen perusteella, kuinka todennäköistä on, että
 - a) $z > 1$
 - b) $z > 2.57$
 - c) $z > 0.18$
 - d) $z > 2.12$
 - e) $z > 1.65$

7. Tiedetään, että suomalaislapset katsovat televisiota keskimäärin kaksi tuntia vuorokaudessa ja että katseluaikojen hajonta on 0.5 tuntia. Poimitaan suomalaislapsista satunnaisesti sadan hengen otos, jonka keskiarvo on 2.2 tuntia. Testaa, tukeeko aineisto hypoteesia ”suomalaiset lapset katselevat televisiota keskimäärin kaksi tuntia vuorokaudessa”.
8. Tiedetään, että standardoidulla WAIS-R-älykkyystestillä mitattuna kaikkien suomalaisten aikuisten älykkyysosamäärien keskiarvo on 100 ja hajonta 15. Poimitaan satunnaisesti kaikista suomalaisista aikuisista
 - a) kymmenen hengen otos, jonka keskiarvo on 105
 - b) sadan hengen otos, jonka keskiarvo on 105.Testaa, tukevatko otokset hypoteesia ”suomalaisten aikuisten älykkyysosamäärien keskiarvo on sata”.
9. Aikaisemmat tutkimukset ovat osoittaneet, että ihmiset pystyvät seuraamaan kerrallaan keskimäärin seitsemää liikkuvaa esinettä keskihajonnan ollessa kaksi. Tutkimukseen osallistuu kymmenen henkilöä, joilta mitataan, kuinka monta liikkuvaa esinettä he pystyvät kerrallaan seuraamaan.

Koehenkilöiden tulokset ovat 8, 7, 6, 8, 9, 8, 5, 6, 7 ja 8. Tukeeko aineisto hypoteesia ”ihmiset pystyvät seuraamaan kerrallaan keskimäärin seitsemää liikkuvaa esinettä”?
10. Tarkastele työmuistitutkimusta (liite 9). Tiedetään, että suomalaisten aikuisten fonologisen työmuistin kapasiteetti on keskimäärin viisi yksikköä keskihajonnan ollessa 0.4. Testaa, saako hypoteesi ”suomalaisten aikuisten fonologisen työmuistin kapasiteetti on keskimäärin viisi yksikköä” tukea
 - a) koko aineistolta
 - b) sokeiden osa-aineistolta
 - c) näkevien osa-aineistolta.
11. Miten tyypin 1 ja 2 virheet eroavat toisistaan ja miten niiden tekemistä voidaan välttää tilastollisia analyyseja tehtäessä?
12. Miten parametriset ja epäparametriset testit eroavat toisistaan?

Keskiarvoja koskeva päättely



10 Aineiston valmistelu

1. Miksi aineiston huolellinen valmisteleminen on tärkeää?
2. Mistä eri syistä johtuen aineistossa voi olla poikkeavia havaintoja?
3. Miksi ja milloin puuttuvat havainnot voidaan joskus korvata muuttujan keskiarvolla?
4. Miten voidaan testata, onko kato systemaattista muuttujalla A?
5. Millaista transformaatiota kannattaa käyttää, jos jakauma on
 - a) lievästi oikealle vino
 - b) keskinkertaisesti vasemmalle vino
 - c) lievästi vasemmalle vino
 - d) voimakkaasti oikealle vino?

11 T-testit

1. Miksi yhden otoksen t-testi on monissa tapauksissa yhden otoksen z-testiä käyttökelpoisempi menettely?
2. Millainen jakauma on Studentin t-jakauma?
3. Millaisia oletuksia t-testit tekevät aineistosta?
4. Jos muut tekijät pysyvät vakioina, miten yhden otoksen t-testi-suureeseen vaikuttavat
 - a) keskiarvon keskivirhe
 - b) otoskoko
 - c) otoksen keskihajonta?
5. Miten eroavat toisistaan yhden otoksen, riippumattomien otosten ja toistettujen mittausten t-testit?

6. Mieti esimerkkejä tutkimuksista, joissa kerättyjen aineistojen analysoimisessa voitaisiin käyttää
 - a) yhden otoksen
 - b) riippumattomien otosten
 - c) toistettujen mittausten t-testiä.
7. Määritä Studentin t-jakaumaan liittyvät kriittiset arvot ($p < 0.01$), kun
 - a) $df = 750$
 - b) $df = 23$
 - c) $df = 100$
 - d) $df = 60$.
8. Määritä Studentin t-jakaumaan liittyvät kriittiset arvot ($p < 0.005$), kun
 - a) $df = 6$
 - b) $df = 90$
 - c) $df = 11$
 - d) $df = 200$.
9. Mikä on t-suureen vaihteluväli?
10. Minkä suuruiset t-suureen arvot ovat yleisiä, mitkä harvinaisia?
11. Miesten keskimääräinen vuosittainen alkoholinkulutus on Suomessa 8.6 litraa puhdasta alkoholia. Poimitaan 50 hengen otos, jossa alkoholinkulutuksen keskiarvo on 7.7 litraa vuodessa ja kulutuksen keskihajonta 2.2 litraa. Testaa, vastaako otos suomalaisia miehiä alkoholin kulutuksen suhteen.
12. Aikaisemmissa tutkimuksissa on havaittu, että asteikolla 1–10 (1 = erittäin negatiivinen, 10 = erittäin positiivinen) ihmiset kokevat tyypillisesti mielialansa olevansa lievästi positiivinen, keskiarvo 5.8. Mitataan vastaavalla asteikolla kymmenen ihmisen mieliala. Saadaan havainnot 4, 5, 6, 8, 9, 8, 7, 6, 7 ja 5. Testaa, ovatko tutkittavat olleet mittaushetkellä samanlaisessa mielialassa kuin ihmiset keskimäärin.
13. Suomalaisten aikuisten keskimääräinen tulos WAIS-R-älykkyys-testissä on tasan sata pistettä. Testaa, onko seuraava aineisto älykkyyden suhteen edustava otos suomalaisista: 100, 121, 96, 97, 84, 99, 102, 103, 104, 119.

14. **Suomalaisten peruskoulun alaluokkalaisten itsetunto on keskimäärin 4.2 (asteikolla 1–10). Testaa, saako tämä hypoteesi tukea koulukiusaamistutkimuksen aineistolta (liite 2).**
15. **Miksi riippumattomien otosten t-testissä yhdistetään otosten keskiarvojen keskivirheet?**
16. **Tarkastellaan koulukiusaamistutkimuksen aineistoa (liite 2). Havaitaanko aineistossa eroja**
 - a) tyttöjen ja poikien itsetunnossa
 - b) tyttöjen ja poikien kiusatuksi tulemisessa?
17. **Tarkastellaan emootioiden kokemistutkimusta (liite 4). Onko miesten ja naisten välillä eroja**
 - a) mielihyvän
 - b) pelon
 - c) hämmästyksen kokemisessa?
18. **Tarkastele työmuistitutkimusaineistoa (liite 9). Onko sokeiden ja näkevien koehenkilöiden välillä eroa**
 - a) fonologisessa työmuistitehtäväsuorituksessa
 - b) taktiillisessa työmuistisuorituksessa?
19. **Tarkastellaan depressiotutkimusta (liite 8). Onko lääkitys- ja psykoterapiahoitoa saavien ryhmien depressiossa eroa**
 - a) tutkimuksen alussa
 - b) tutkimuksen päättyessä?
20. **Mitä tarkoitetaan**
 - a) riippumattomien otosten asetelmalla
 - b) toistomittausasetelmalla?
21. **Miksi toistomittausasetelma on usein käyttökelpoisempi kuin riippumattomien otosten asetelma?**
22. **Tarkastellaan Stroop-tutkimuksen aineistoa (liite 7). Ovatko reaktioajat olleet nopeampia kongruenteissa vai inkongruenteissa tilanteissa?**
23. **Tarkastellaan depressiotutkimuksen aineistoa (liite 8). Ovatko tutkittavat olleet yhtä depressoituneita alku- ja loppumittauksissa?**

24. Tarkastellaan tutkimusta, jossa on selvitetty tarkkaavaisuuden suuntautumista katsevihjeiden perusteella (liite 3). Laske keskiarvot reaktioajoista eri tilanteissa ja hahmottele ne kuvaajaan. Laske sen jälkeen, ovatko tutkittavat reagoineet yhtä nopeasti
- a) kongruenteissa ja inkongruenteissa
 - b) inkongruenteissa ja neutraaleissa
 - c) kongruenteissa ja neutraaleissa tilanteissa.

12 Yksisuuntainen varianssianalyysi

1. Miksi t-testi ei sovellu käytettäväksi sellaisessa tilanteessa, jossa vertailtavia ryhmiä on useampia kuin kaksi?
2. Miten t-testi ja varianssianalyysi eroavat toisistaan keskiarvojen vertailussa?
3. Mitkä tekijät vaikuttavat
 - a) ryhmien välisen
 - b) ryhmien sisäisen varianssin suuruuteen?
4. Miten F-suhde määritellään ja miten sitä tulkitaan?
5. Tutkimukseen on osallistunut yhteensä kolmekymmentä henkilöä, jotka on jaettu kolmeen kymmenen hengen suuruiseen ryhmään. Estimoi populaatiovarianssi kun ryhmäkeskiarvojen varianssi on ollut
 - a) 2
 - b) 4
 - c) 10.
6. Määritä taulukosta seuraavien F-jakaumien kriittiset ($p < .05$) arvot.
 - a) $F(1,100)$
 - b) $F(3,50)$
 - c) $F(6,500)$
 - d) $F(2,20)$
 - e) $F(1,5)$
7. Mikä on F-suureen vaihteluväli?
8. Millaiset F-suureen arvot ovat yleisiä, millaiset harvinaisia?

9. Tarkastellaan työmuistitutkimusta (liite 9). Vertaa sokeiden ja näkevien taktiilista työmuistikapasiteettia toisiinsa

- a) riippumattomien otosten t-testillä
- b) riippumattomien otosten varianssianalyysillä.

Vertaile a- ja b-kohdissa saamiasi tuloksia keskenään.

10. Tarkastellaan tutkimusta, jossa on selvitetty neljäs- viides- ja kuudesluokkalaisten lasten itsetunnon eroja. Mittaukset on tehty asteikolla 1–10, ja luokakeskiarvot ovat $\bar{x}_4 = 5.6$, $\bar{x}_5 = 5.8$ ja $\bar{x}_6 = 6.6$. Monivertailujen p-arvot ovat seuraavien suuruiset:

- a) neljäsluokkalaiset vs. viidesluokkalaiset = .03
- b) neljäsluokkalaiset vs. kuudesluokkalaiset = .01
- c) viidesluokkalaiset vs. kuudesluokkalaiset = .01.

Suorita p-arvoille Bonferroni-korjaukset ja tulkitse sen jälkeen tulokset.

11. Tutkimuksessa on selvitetty riippumattomien otosten asetelmalla, miten huulten liikkeiden näkyminen vaikuttaa puheen ymmärtämiseen. Tutkittavat (n = 30) jaettiin kolmeen ryhmään, joista kontrolliryhmä näki puhujan kasvot normaalisti. Ensimmäinen koeryhmä ei nähnyt puhujan huulien liikkeitä ollenkaan, ja toiselle koeryhmälle esitetyn puhujan suun liikkeet eivät olleet synkronoituja kuultuun puheeseen. Tutkittavien tehtävinä oli tunnistaa puhujan lausumat sanat. Havaittiin, että eri tekijät vaikuttivat sanojen tunnistamisprosenttiin seuraavasti:

- a) käsittelytekijän aiheuttama vaihtelu = 100 yksikköä, satunnaisvirhe = 20 yksikköä, yksilölliset erot 20 yksikköä
- b) käsittelytekijän aiheuttama vaihtelu = 20 yksikköä, satunnaisvirhe = 100 yksikköä, yksilölliset erot 20 yksikköä
- c) käsittelytekijän aiheuttama vaihtelu = 20 yksikköä, satunnaisvirhe = 20 yksikköä, yksilölliset erot 100 yksikköä.

Laske F-suhteet ja määritä, olisivatko ne tilastollisesti merkitseviä.

12. Tarkastellaan lukeminen ja työmuisti -aineistossa (liite 1) hypoteeseja H0 (työmuistikapasiteetti ei vaikuta kokonaispalautukseen) ja H1 (työmuistikapasiteetti vaikuttaa kokonaispalautukseen). Laske aineistosta

- a) yhteisneliösumma ja sen vapausasteet
- b) ryhmien välinen neliösumma ja sen vapausasteet
- c) ryhmien sisäinen neliösumma ja sen vapausasteet
- d) F-suhde ja tulkitse se.

13. Tarkastellaan lukeminen ja työmuisti -aineistossa (liite 1) hypoteeseja H_0 (työmuistikapasiteetti ei vaikuta tekstin lukuaikaan) ja H_1 (työmuistikapasiteetti vaikuttaa tekstin lukuaikaan). Laske aineistosta
 - a) yhteisneliösumma ja sen vapausasteet
 - b) ryhmien välinen neliösumma ja sen vapausasteet
 - c) ryhmien sisäinen neliösumma ja sen vapausasteet
 - d) F-suhde ja tulkitse se.
14. Suorita edellisen tehtävän analyysille kontrastivertailut riippumattomien otosten t-testien avulla.
15. Tarkastellaan työn kuormittavuustutkimusta (liite 5). Testaa, kumpi seuraavista hypoteeseista saa tukea: H_0 (koulutus ei ole yhteydessä työn kuormittavuuden kokemiseen) vai H_1 (koulutus on yhteydessä työn kuormittavuuden kokemiseen). Tee tarvittaessa monivertailut riippumattomien otosten t-testien avulla.
16. Tarkastellaan prosessointikontekstiaineistoa (liite 12). Testaa kumpi seuraavista hypoteeseista saa tukea: H_0 (prosessointikonteksti ei vaikuta tekstin ymmärtämiseen) vai H_1 (prosessointikonteksti vaikuttaa tekstin ymmärtämiseen). Tee tarvittaessa monivertailut riippumattomien otosten t-testien avulla.
17. Tarkastellaan prosessointikontekstiaineistoa (liite 12). Testaa kumpi seuraavista hypoteeseista saa tukea: H_0 (prosessointikonteksti ei vaikuta tekstin kokonaispalautukseen) vai H_1 (prosessointikonteksti vaikuttaa tekstin kokonaispalautukseen). Tee tarvittaessa monivertailut riippumattomien otosten t-testien avulla.

13 Useampisuuntainen varianssianalyysi

1. Miten useampisuuntainen varianssianalyysi eroaa yksisuuntaisesta varianssianalyysistä?
2. Millaisia oletuksia useampisuuntainen varianssianalyysi tekee analysoitavasta aineistosta?
3. Miten vapausasteet määräytyvät useampisuuntaisessa varianssianalyysissä?

4. Mitä tarkoittavat päävaikutus ja yhdysvaikutus?
5. Mitä seuraavat varianssitermit kuvaavat?
 - a) tekijän A aiheuttama varianssi
 - b) tekijän B aiheuttama varianssi
 - c) yhdysvaikutuksen varianssi
 - d) ryhmien sisäinen varianssi
6. Määritä, kuinka monta tekijää ja kuinka monta tasoa kullakin tekijällä on seuraavissa varianssianalyysiasetelmissa
 - a) 2×2
 - b) $3 \times 3 \times 4$
 - c) $2 \times 3 \times 3$
 - d) $4 \times 3 \times 3 \times 4$
7. Tarkastellaan riippumattomien otosten asetelmalla sukupuolen (kaksi tasoa: mies ja nainen) ja tupakoinnin (kaksi tasoa: tupakoi, ei tupakoi) vaikutuksia hajuaistin toimintaan (jatkuva muuttuja). Millaisiin alaryhmiin aineisto jakautuu tekijöiden perusteella?
8. Tarkastellaan riippumattomien otosten asetelmalla kirjasinleikkauksen (kolme tasoa: serif, sans serif, fraktuura) ja kirjasin-koon (kaksi tasoa: normaali, pieni) vaikutusta tekstin lukuaikaan (jatkuva muuttuja). Millaisiin alaryhmiin aineisto jakautuu tekijöiden perusteella?
9. Tarkastellaan koulukiusaamisaineistoa (liite 2). Millaisiin alaryhmiin aineisto jakautuu sukupuoli- ja sisarusten lukumäärä-muuttujien perusteella?
10. Miten varianssianalyysin päävaikutukset tulkitaan, jos analyysissa havaitaan myös yhdysvaikutus?
11. Tarkastellaan lukemistutkimuksen aineistoa (liite 1) ja analysoidaan, miten työmuistin kapasiteetti ja sukupuoli vaikuttavat tekstin kokonaislukuaikaan. Jaa aineisto alaryhmiin
 - a) työmuistikapasiteetin
 - b) sukupuolen
 - c) työmuistikapasiteetin ja sukupuolen perusteella.

Laske alaryhmien keskiarvot. Hahmottele a-, b- ja c- kohtien keskiarvot erillisin kuvaajiin. Mitä voit arvioida tutkimuksen todennäköisistä tuloksista keskiarvojen perusteella?

- 12. Tarkastellaan lukemistutkimuksen aineistoa (liite 1) ja analysoidaan, miten työmuistin kapasiteetti ja sukupuoli vaikuttavat kokonaispalautukseen. Jaa aineisto alaryhmiin**

- a) työmuistikapasiteetin
- b) sukupuolen
- c) työmuistikapasiteetin ja sukupuolen perusteella.

Laske alaryhmien keskiarvot. Hahmottele a-, b- ja c- kohtien keskiarvot erillisiin kuvaajiin. Mitä voit arvioida tutkimuksen todennäköisistä tuloksista keskiarvojen perusteella?

- 13. Tarkastellaan edelleen lukemistutkimuksen aineistoa (liite 1). Käsitellään työmuistin kapasiteettia rivitekijänä (A) ja sukupuolta saraketekijänä (B) ja tarkastellaan niiden vaikutuksia tekstin kokonaisluku-aikaan. Laske**

- a) yhteisneliösumma ja sen vapausasteet
- b) rivineliösumma ja sen vapausasteet
- c) sarakeneliösumma ja sen vapausasteet
- d) yhdysvaikutuksen neliösumma ja sen vapausasteet.

- 14. Laske edellisen tehtävän tulosten perusteella**

- a) F-suhde työmuistikapasiteetin päävaikutukselle
 - b) F-suhde sukupuolen päävaikutukselle
 - c) F-suhde työmuistikapasiteetin ja sukupuolen yhdysvaikutukselle.
- Tulkitse saamasi tulokset.

- 15. Tarkastellaan edelleen lukemistutkimusaineistoa (liite 1). Käsitellään työmuistin kapasiteettia rivitekijänä (A) ja sukupuolta saraketekijänä (B) ja tarkastellaan niiden vaikutuksia kokonaispalautukseen. Laske**

- a) yhteisneliösumma ja sen vapausasteet
- b) rivineliösumma ja sen vapausasteet
- c) sarakeneliösumma ja sen vapausasteet
- d) yhdysvaikutuksen neliösumma ja sen vapausasteet.

- 16. Laske edellisen tehtävän tulosten perusteella**

- a) F-suhde työmuistikapasiteetin päävaikutukselle
 - b) F-suhde sukupuolen päävaikutukselle
 - c) F-suhde työmuistikapasiteetin ja sukupuolen yhdysvaikutukselle.
- Tulkitse saamasi tulokset

17. Tarkastele prosessointikontekstitutkimusta (liite 12). Laske, miten koulutus ja prosessointikonteksti vaikuttavat tekstin ymmärtämiseen. Piirrä tuloksista sopiva kuvaaja.
18. Tarkastele prosessointikontekstitutkimusta (liite 12). Laske, miten koulutus ja prosessointikonteksti vaikuttavat tekstin palautukseen. Piirrä tuloksista sopiva kuvaaja.

14 Toistettujen mittausten varianssianalyysi

1. Miksi toistettujen mittausten varianssianalyysin avulla voidaan havaita pienempiä efektejä kuin riippumattomien otosten varianssianalyysillä?
2. Miten vapausasteet määräytyvät toistettujen mittausten varianssianalyysissa?
3. Mitä seuraavat varianssitermit kuvaavat toistettujen mittausten varianssianalyysissa?
 - a) yhteisvarienssi
 - b) käsittelyjen sisäinen varienssi
 - c) käsittelyjen välinen varienssi
 - d) jäännösvarienssi
4. Mieti esimerkkejä sellaisista tutkimusasetelmista, joissa kerättyjä aineistoja voitaisiin analysoida toistettujen mittausten varianssianalyysin avulla.
5. Tutkimuksessa on selvitetty sanojen yleisyyden (harvinainen, melko yleinen, yleinen) vaikutusta niiden tunnistamisnopeuteen (jatkuva muuttuja). Tutkimukseen osallistuu 30 koehenkilöä, ja se suoritetaan toistomittausasetelmalla. Oletetaan, että eri tekijät aiheuttavat mittaustuloksiin vaihtelua seuraavasti:
 - a) käsittelytekijän aiheuttama vaihtelu = 20 yksikköä, satunnaisvirhe = 20 yksikköä, yksilölliset erot 80 yksikköä
 - b) käsittelytekijän aiheuttama vaihtelu = 80 yksikköä, satunnaisvirhe = 20 yksikköä, yksilölliset erot 20 yksikköä
 - c) käsittelytekijän aiheuttama vaihtelu = 20 yksikköä, satunnaisvirhe = 80 yksikköä, yksilölliset erot 80 yksikköä.

Laske F-suhteet ja määritä, olisivatko ne tilastollisesti merkitseviä.

6. Oletetaan vielä, että sama tutkimus suoritettaisiin riippumattomien otosten asetelmalla. Tutkittavat ($n = 30$) jaetaan siten kolmeen kymmenen hengen ryhmään, ja yksi ryhmä saa aina tunnistettavakseen ainoastaan yhden yleisyysryhmän sanoja. Oletetaan, että tekijät aiheuttavat mittaustuloksiin variaatiota samalla tavalla kuin edellisessä tehtävässä.
- Laske F-suhteet, määritä ovatko ne tilastollisesti merkitseviä ja vertaa tuloksia edellisessä kohdassa saatuihin.
7. Tarkastellaan Stroop-tutkimuksen aineistoa (liite 7). Testaa, miten data tukee hypoteeseja H_0 : sanan ja värin kongruenssi ei vaikuta sanan prosessointinopeuteen ja H_1 : sanan ja värin kongruenssi vaikuttaa sanan prosessointinopeuteen.
8. Tarkastellaan emootiot ja tarkkaavaisuus -aineistoa (liite 6). Laske keskiarvot todennäköisyyksistä sille, että tarkkaavaisuus on siirtynyt ensin
- a) epämiellyttävään
 - b) neutraaliin
 - c) miellyttävään kuvaan.
- Hahmottele tuloksista kuvaaja. Mitä voit sanoa tutkimuksen todennäköisistä tuloksista keskiarvojen perusteella?
9. Analysoidaan emootiot ja tarkkaavaisuus -aineisto (liite 6) toistettujen mittausten varianssianalyysin avulla. Tarkastellaan hypoteeseja H_0 : kohdekuvan emotionaalinen sisältö ei vaikuta siihen, mihin tarkkaavaisuus suuntautuu ensimmäiseksi ja H_1 : kohdekuvan emotionaalinen sisältö vaikuttaa siihen, mihin tarkkaavaisuus suuntautuu ensimmäiseksi. Laske F- suhde ja määritä, onko se tilastollisesti merkitsevä.
10. Tarkastellaan visuaalisen tarkkaavaisuuden suuntautuminen katsevihjeiden perusteella -aineistoa (liite 3). Laske keskiarvot reaktioajoista
- a) kongruenteissa
 - b) neutraaleissa
 - c) inkongruenteissa tilanteissa.
- Hahmottele tuloksista kuvaaja. Mitä voit sanoa tutkimuksen todennäköisistä tuloksista keskiarvojen perusteella?

11. Analysoidaan visuaalisen tarkkaavaisuuden suuntautuminen katsevihjeiden perusteella -aineisto (liite 3) toistettujen mittausten varianssianalyysin avulla. Tarkastellaan hypoteeseja
 H_0 : katsevihjeet eivät suuntaa tarkkaavaisuutta refleksiivisesti ja
 H_1 : katsevihjeet suuntaavat tarkkaavaisuutta refleksiivisesti.
 Laske F-suhde ja määritä, onko se tilastollisesti merkitsevä. Tee tarvittaessa jatkovertailut toistettujen mittausten t-testillä.
12. Tarkastellaan kasvoinversioaineistoa (liite 10). Tarkastele toistettujen mittausten varianssianalyysin ja sopivien jatkovertailujen avulla, vaikuttaako kasvoinversio kasvojen tunnistamiseen.

15 Epäparametriset menetelmät

1. Millaisissa tilanteissa tarvitaan epäparametrisia testejä?
2. Millaisissa tilanteissa parametristen testien käyttäminen erityisesti kasvattaa tyypin 1 virheen tekemisen riskiä?
3. Millaisissa tilanteissa parametristen testien käyttäminen erityisesti kasvattaa tyypin 2 virheen tekemisen riskiä?
4. Millaisia oletuksia epäparametriset testit tekevät testattavien muuttujien jakaumasta?
5. Miksi epäparametristen testien avulla ei voi tehdä päätelmiä siitä, eroavatko jakaumien keskiarvot toisistaan?
6. Miksi epäparametrisia testejä ei kannata käyttää parametristen testien sijaan silloin, kun parametrisen testin tekemät oletukset ovat voimassa?
7. Mikä on epäparametrinen vastine
 - a) riippumattomien otosten t-testille
 - b) toistettujen mittausten t-testille
 - c) yksisuuntaiselle varianssianalyysille
 - d) toistettujen mittausten varianssianalyysille?

8. **Tarkastellaan Stroop-tutkimuksen aineistoa (liite 7). Määritä reaktioaikojen järjestysluvut**
 - a) kongruentissa tilanteessa
 - b) inkongruentissa tilanteessa.
9. **Tarkastele depressiotutkimuksen aineistoa (liite 8). Testaa U-testin avulla, eroavatko lääkehoitoryhmä ja psykoterapia-ryhmä toisistaan**
 - a) depression alkumittauksessa
 - b) depression loppumittauksessa.
10. **Tarkastele työmuistitutkimuksen aineistoa (liite 9). Testaa U-testin avulla, eroavatko sokeat ja näkevät toisistaan**
 - a) taktiilisen työmuistin kapasiteetissa
 - b) fonologisen työmuistin kapasiteetissa.
11. **Tarkastele lukemistutkimuksen aineistoa (liite 1). Testaa U-testin avulla, eroavatko miehet ja naiset toisistaan**
 - a) tekstin kokonaislukuajan
 - b) kokonaispalautuksen suhteen.
12. **Tarkastellaan kasvoinversioaineistoa (liite 10). Testaa Wilcoxonin testin avulla, onko kasvojen tunnistamisessa eroja**
 - a) 90 asteen ja 180 asteen tilanteiden välillä
 - b) 0 asteen ja 90 asteen tilanteiden välillä.
13. **Tarkastele depressiotutkimuksen aineistoa (liite 8). Testaa Wilcoxonin testin avulla, onko tutkittavien depression kokemisessa eroja mittauskerroilla 1 ja 2, kun hoitoryhmää ei oteta huomioon.**
14. **Tarkastele Stroop-tutkimuksen aineistoa (liite 7). Testaa Wilcoxonin testin avulla, vaikuttaako sanan ja värin kongruenssi sanojen prosessointiin.**
15. **Tarkastele lukemistutkimuksen aineistoa (liite 1). Testaa Kruskal–Wallis-testin avulla, vaikuttaako työmuistin kapasiteetti**
 - a) tekstin lukuaikaan
 - b) kokonaispalautukseen.

16. **Tarkastele työn kuormittavuustutkimuksen aineistoa (liite 5). Testaa Kruskall–Wallis-testin avulla, onko koulutusryhmien välillä eroa**
 - a) työssä viihtymisessä
 - b) työn kuormittavuuden kokemisessa.
17. **Tarkastele prosessointikontekstitutkimuksen aineistoa (liite 12). Testaa Kruskall–Wallis-testin avulla, vaikuttiko prosessointi-kontekstin esittäminen**
 - a) tekstin ymmärtämiseen
 - b) kokonaispalautukseen.

Tee tarvittaessa jatkovertailut U-testillä.
18. **Tarkastele kasvoinversioaineistoa (liite 10). Testaa Friedmanin testin avulla, vaikuttaako kasvoinversio kasvojen tunnistamiseen.**
19. **Tarkastele tarkkaavaisuuden suuntautumistutkimuksen aineistoa (liite 3). Testaa Friedmanin testin avulla, suuntaavatko katse-vihjeet tarkkaavaisuutta refleksiivisesti.**
20. **Tarkastele emotiot ja tarkkaavaisuus -tutkimuksen aineistoa (liite 6). Testaa Friedmanin testin avulla, vaikuttaako kuvien emotionaalinen sisältö tarkkaavaisuuden suuntautumiseen.**

Yhteisvaihtelun tutkiminen

IV

16 Kovarianssi ja korrelaatio

1. Millaisia erilaisia tulkintoja kahden muuttujan yhteisvaihtelulle voidaan antaa?
2. Millaisissa tilanteissa kahden muuttujan yhteisvaihtelun voidaan ajatella merkitsevän muuttujien välistä kausaalista yhteyttä?
3. Piirrä sirontakuviioon kahden viidestä havainnosta koostuvan muuttujan välinen
 - a) positiivinen lineaarinen yhteys
 - b) negatiivinen lineaarinen yhteys
 - c) positiivinen monotoninen yhteys
 - d) epälineaarinen yhteys.
4. Mikä on kovarianssikertoimen vaihteluväli?
5. Mitä tulkinnallisia vaikeuksia liittyy
 - a) korrelaatioon ja kausaalipäätelmiin
 - b) korrelaatiokertoimen suuruuteen ja tilastolliseen merkitsevyyteen?Miten tulkinnalliset vaikeudet voidaan ottaa huomioon?
6. Miksi kovarianssi on tulkinnallisesti hankala yhteisvaihtelua kuvaava tunnusluku?
7. Mitä eroa on Pearsonin tulomomentti- ja Spearmanin järjestyskorrelaatiokertoimella?
8. Millaisissa tilanteissa kannattaa käyttää Spearmanin järjestyskorrelaatiokerrointa, vaikka jakaumaoletusten puolesta voisikin käyttää Pearsonin tulomomenttikorrelaatiokerrointa?
9. Mitä tarkoittaa
 - a) vipuvoima
 - b) diskrepanssi?
10. Miten erilaiset poikkeavat havainnot saattavat vaikuttaa korrelaatiokertoimien suuruuteen?

11. Millaisessa tapauksessa $r \neq 1$, vaikka sirontakuviassa kaikki havaintopisteet sijoittuisivat yhdelle suoralle?
12. Miksi korrelaatiokertoimen tilastollinen merkitsevyys pitää testata?
13. Miten osittaiskorrelaatiokerroin ja multippelikorrelaatiokerroin poikkeavat toisistaan?
14. Miksi tulomomenttikorrelaatiokerroin ei välttämättä anna luotettavaa estimaattia muuttujien välisen yhteyden voimakkuudesta, jos muuttujien varianssit ovat vähäisiä?
15. Tuota sellaiset 2×10 datamatriisit, joissa
 - a) $r_{xy} = r_{s(xy)}$
 - b) $r_{xy} < r_{s(xy)}$
16. Tuota sellaiset 3×10 datamatriisit, joissa
 - a) $r_{xy,z} > r_{xy}$
 - b) $r_{xy,z} < r_{xy}$
 - c) $r_{xy,z} \approx r_{xy}$
17. Tutkimuksessa havaitaan, että esikoululaisten fonologisen prosessoinnin tason (0–100) ja kuullun ymmärtämisen (0–100) välinen järjestyskorrelaatiokerroin on 0.32. Jos käytetään .05:n havaittua merkitsevyystasoa, miten tulkitset tuloksen, jos tutkimuksen otoskoko on ollut
 - a) 10
 - b) 25
 - c) 50?
18. Minkä suuruinen otos olisi poimittava, jotta seuraavat korrelaatiokertoimet olisivat tilastollisesti merkitseviä?
 - a) $r = .22$ ($p < .05$)
 - b) $r = .33$ ($p < .01$)
 - c) $r = .12$ ($p < .001$)
19. Tarkastele työn kuormittavuusaineistoa (liite 5). Laske seuraavien muuttujien väliset tulomomenttikorrelaatiokertoimet ja testaa niiden tilastollinen merkitsevyys.
 - a) työssä viihtyminen ja työn kuormittavuus
 - b) työn kuormittavuus ja työilmapiiri
 - c) työssä viihtyminen ja työilmapiiri

- 20. Laske lukemistutkimusaineistosta (liite 1) tekstin kokonaislukuan ja kokonaispalautuksen välinen**
- a) kovarianssi
 - b) tulomomenttikorrelaatiokerroin
 - c) järjestyskorrelaatiokerroin.
- 21. Laske koulukiusaamistutkimuksen aineistosta (liite 2) seuraavien muuttujien väliset tulomomenttikorrelaatiokertoimet ja testaa niiden tilastollinen merkitsevyys:**
- a) työmuistikapasiteetti ja itsetunto
 - b) kiusatuksi tulemisen kokemus ja fyysinen aggressio
 - c) fyysinen aggressio ja työmuistikapasiteetti.
- 22. Laske koulukiusaamistutkimuksen aineistosta (liite 2) seuraavien muuttujien väliset järjestyskorrelaatiokertoimet ja testaa niiden tilastollinen merkitsevyys:**
- a) itsetunto ja työmuistikapasiteetti
 - b) kiusatuksi tulemisen kokemus ja itsetunto
 - c) itsetunto ja fyysinen aggressio.
- 23. Laske koulukiusaamistutkimuksen aineistosta (liite 2)**
- a) itsetunnon ja kiusatuksi tulemisen multippelikorrelaatio aggression kanssa
 - b) työmuistikapasiteetin ja itsetunnon multippelikorrelaatio kiusatuksi tulemisen kanssa.
- 24. Laske työn kuormittavuusaineistosta (liite 5)**
- a) työssä viihtymisen ja työn kuormittavuuden välinen korrelaatio, kun työilmapiiirin vaikutus on poistettu
 - b) työssä viihtymisen ja työn kuormittavuuden välinen korrelaatio, kun itsetunnon vaikutus on poistettu.

17 Kategoristen muuttujien välisten yhteyksien tarkasteleminen

1. Mitä eroa on χ^2 -yhteensopivuus- ja χ^2 -riippumattomuustesteillä?
2. Millainen jakauma on χ^2 -jakauma?
3. Mikä on χ^2 -jakauman odotusarvo?
4. Mieti esimerkkejä sellaisista tutkimuskysymyksistä, joiden vastaamiseen tarvittaisiin
 - a) χ^2 -yhteensopivuustestiä
 - b) χ^2 -riippumattomuustestiä.
5. Tarkastele, ovatko seuraavat χ^2 -testisuureen arvot tilastollisesti merkitseviä ($p=.05$).
 - a) $\chi^2(2)=2.24$
 - b) $\chi^2(2)=10$
 - c) $\chi^2(8)=10$
 - d) $\chi^2(4)=15$
6. Tutkimuksen ($n = 100$) otos koostuu 56 miehestä ja 44 naisesta. Testaa χ^2 -yhteensopivuustestin avulla, vastaako otos odotettua ”yhtä paljon naisia ja miehiä” -jakaumaa?
7. Noin yksi prosentti suomalaisista sairastaa skitsofreniaa. Otoksessa ($n = 100$) havaitaan viiden henkilön sairastavan skitsofreniaa. Voidaanko otosta pitää edustavana skitsofrenian suhteen?
8. Tiedetään, että kymmenen prosenttia ihmisistä on vasenkätisiä. Otokseen poimittujen henkilöiden kätisyydet ovat *oikea, oikea, vasen, oikea, oikea, oikea, vasen, vasen, oikea, vasen, oikea, oikea, oikea, oikea, vasen, oikea, oikea, oikea, oikea, oikea*. Testaa, poikkeako otoksen kätisyysjakauma tilastollisesti merkitsevästi populaation kätisyysjakaumasta.
9. Tutkimuksessa on selvitetty depression esiintymistä miehillä ja naisilla. Tutkimukseen osallistuu 40 miestä ja 65 naista. Miehistä viidellä ja naisilla viidellätoista havaitaan depressio. Testaa χ^2 -yhteensopivuustestin avulla, saako hypoteesi ”depressio on yhtä yleistä miehillä ja naisilla” tukea aineistosta.

10. Tutkimuksessa on tarkasteltu tupakoinnin yhteyttä ylipainoisuuteen. Tutkimusaineisto on seuraavanlainen:

	Tupakoi	Ei tupakoi
Ylipainoinen	18	24
Normaalipainoinen	44	26

Testaa χ^2 -yhteensopivuustestin avulla, onko tupakointi yhteydessä ylipainoisuuteen.

11. Tutkimuksessa on tarkasteltu suuronnettomuuden jälkeen suoritettujen debriefingin (jälkipuinnin) tehoa posttraumaattisen stressireaktion (PTSD) laukeamiseen. Puolet koehenkilöistä osallistui debriefingiin, puolet ei. Kuukauden kuluttua selvitetiin, kuinka moni tutkittavista oli saanut PTSD-oireita. Aineisto oli seuraavankaltainen:

	Debriefing	Kontrolli
Oireita	25	14
Ei oireita	25	36

Testaa χ^2 -yhteensopivuustestin avulla, miten debriefing vaikutti PTSD-oireiden ilmenemiseen.

12. Tutkimuksessa selvitettiin, miten Rorschachin mustetahrastein (Ro) ja Beckin depressioinventaarin (BDI-II) perusteella voidaan diagnosoida masentuneisuus (depressio). Puolet tutkimukseen osallistuneista psykologeista perusti diagnoosinsa Ro-testiin, puolet taas BDI-II:een. Alla olevassa taulukossa on esitetty, kuinka hyvin nämä kaksi ryhmää onnistuivat koepotilaiden depression diagnosoimisessa.

	Ro	BDI-II
Oikea diagnoosi	10	29
Väärä diagnoosi	33	14

Testaa χ^2 -yhteensopivuustestin avulla, ovatko menetelmät yhtä tehokkaita depressiodiagnostiikassa.

13. Tutkimuksessa on selvitetty, miten amygdalan eli manteliumakkeen leesio (vaurioittaminen) vaikuttaa rottien pelkoehdollistumiseen. Kokeessa rotat jaetaan kolmeen ryhmään. Kontrolliryhmän rottien amygdalaa ei vaurioiteta, ensimmäisen koeryhmän rotille tehdään unilateraalinen (yksipuolinen) amygdalaleesio ja toisen koeryhmän rotille bilateraalinen (molemminpuoleinen) amygdalaleesio. Tämän jälkeen testataan, kykenevätkö eri ryhmissä olevat rotat oppimaan (ehdollistamaan) yhteyden kovan äänen ja epämiellyttävän ärsykkeen (jalkaan annettavan heikon sähköiskun) välille.

	ehdollistuu	ei ehdollistuu
Kontrolliryhmä	24	6
Unilateraalinen leesio	13	17
Bilateraalinen leesio	5	25

Testaa χ^2 -yhteensopivuustestin avulla, vaikuttaako amygdalaleesio pelon ehdollistumiseen.

14. Tarkastellaan työuupumusaineistoa (liite 11). Testaa, onko työuupumukseen sairastumisella ja sukupuolella yhteyttä.
15. Tarkastele, havaitaanko prosessointikontekstiaineistossa (liite 12) yhteyttä tutkittavien sukupuolen ja koulutuksen välillä.

18 Lineaarinen regressio

1. Millaisia oletuksia lineaarinen regressioanalyysi tekee tutkimusaineistosta?
2. Mitä tarkoittavat seuraavat käsitteet ja millainen tulkinta niillä on?
 - a) vakiotermi
 - b) regressiokerroin
 - c) standardoitu regressiokerroin
 - d) jäännöstermi
3. Miten lineaarisessa regressioanalyysissä testataan
 - a) mallin sopivuus aineistoon
 - b) mallin selitysaste?

4. Miksi selittäjien multikolineaarisuus on ongelmallista lineaarissa regressioanalyysissa?
5. Hahmottele jäännöstermien jakaumakuvaajat tilanteissa, joissa
 - a) malli sopii hyvin kaikilla y-muuttujan arvoilla
 - b) malli ennustaa systemaattisesti liian suuria y-muuttujan arvoja
 - c) jäännöstermin suuruuden ja y-muuttujan arvojen välillä on negatiivinen yhteys
 - d) malli sopii aineistoon pienillä, mutta ei suurilla y-muuttujan arvoilla.
6. Tutkimuksessa ennustetaan oppilaiden koulumenestystä matematiikassa (4–10) yleisellä älykkyydellä (0–160). Regressiomalliksi saadaan $\hat{y} = .05 \times \text{älykkyyys} + 4$. Millaiseksi malli ennustaa koehenkilön koulumenestyksen, jos hänen yleinen älykkyytensä on

a) 100	c) 120
b) 60	d) 160?
7. Mitä voit sanoa edellisen kohdan regressiomallin toimivuudesta?
8. Tutkimuksessa on selvitetty, miten aikuisopiskelijoiden tietokoneen käyttökokemus (0–10) ja ikä selittävät opiskelijoiden suhtautumista www-ympäristössä suoritettaviin opintoihin (0 = erittäin kielteinen – 10 = erittäin myönteinen). Regressiomalliksi saadaan $\hat{y} = .5 \times \text{käyttökokemus} - .1 \times \text{ikä} + 6$.
Miten käyttökokemus ja ikä vaikuttavat www-ympäristössä suoritettavien opintojen koettuun mielekkyyteen? .
9. Millaiseksi edellä mainittu malli arvioi www-ympäristössä suoritettavien opintojen koetun mielekkyyden tutkittavalle, joka on
 - a) 30-vuotias ja jonka tietokoneen käyttökokemus on 8
 - b) 40-vuotias ja jonka tietokoneen käyttökokemus on 2
 - c) 60-vuotias ja jonka tietokoneen käyttökokemus on 0
 - d) 48-vuotias ja jonka tietokoneen käyttökokemus on 5?

10. Oletetaan, että koululaisten kaverien lukumäärä (0–8) on lineaarisessa yhteydessä kiusatuksi tulemiseen todennäköisyyteen luokassa (0–100 %). Muodostetaan lineaarinen regressiomalli, jossa ennustetaan kiusatuksi tulemistä kaverien lukumäärällä. Tiedetään, että jos kaverien lukumäärä on 0 yksikköä, kiusatuksi tulemisen todennäköisyys on 70 % ja että kavereiden lukumäärä-muuttujan regressiokerroin on –10. Millaiseksi regressiomalli ennustaa kiusatuksi tulemisen todennäköisyyden, jos tutkittavan kavereiden lukumäärä on
a) 3
b) 4
c) 5
d) 7?
11. Millainen on edellisen tehtävän regressiomallin yhtälö?
12. Tarkastellaan, miten fyysinen aggressio ja työmuistin kapasiteetti ovat yhteydessä toisiinsa koulukiusaamistutkimuksen aineistossa (liite 2).
a) Laske muuttujien *fyysinen aggressio* ja *työmuisti* välisen tulomomenttikorrelaatiokertoimen suuruus.
b) Piirrä muuttujien välinen sirontakuviio.
c) Laadi regressiomalli, jossa selität fyysistä aggressiota työmuistilla. Määritä regressiokerroin työmuistille (β_1) ja vakiotermille (β_0).
d) Määritä mallin selityssaste.
e) Tulkitse regressiomalli ja arvioi sen toimivuutta aineistossa.
13. Tarkastellaan, miten kiusatuksi tulemisen kokemus selittää itsetuntoa koulukiusaamisaineistossa (liite 2).
a) Laske muuttujien *kiusatuksi tulemisen kokemus* ja *itsetunto* välisen tulomomenttikorrelaatiokertoimen suuruus ja testaa sen tilastollinen merkitsevyys.
b) Piirrä muuttujien välinen sirontakuviio.
c) Laadi regressiomalli, jossa selität itsetuntoa kiusatuksi tulemisen kokemuksella. Määritä regressiokerroin kiusatuksi tulemisen kokemukselle (β_1) ja vakiotermille (β_0).
d) Määritä mallin selityssaste.
e) Tulkitse regressiomalli.
14. Tarkastellaan työuupumusaineistoa (liite 11). Määritä regressiomalli, jossa ennustat
a) kuormituksen kokemusta väsymyksellä
b) kuormituksen kokemusta unen puutteella.
Määrittele myös mallien selityssasteet.

15. Tarkastellaan työuupumusaineistoa (liite 11). Määritä regressiomalli, jossa ennustat
- a) kuormituksen kokemusta väsymyksellä
 - b) kuormituksen kokemusta unen puutteella.
- Määrittele myös mallien selitysasteet.

19 Logistinen regressio

1. Miten lineaarinen ja logistinen regressioanalyysi eroavat toisistaan?
2. Mitä eroa on binäärisillä ja multinomisilla logistisilla malleilla?
3. Miksi logistisessa regressiomallissa käytetään logitia sen sijaan että mallinnettaisiin suoraan esimerkiksi riskisuhdetta tai referenssikategoriaan kuulumisen todennäköisyyksiä?
4. Miten riski ja logit määritellään?
5. Tiedetään, että $P(A)$ ja $P(B)$ ovat toistensa komplementti-tapahtumia. Määrittele riskisuhteet ja logitit, kun
 - a) $P(A) = .4$ ja referenssikategoria on A
 - b) $P(B) = .2$ ja referenssikategoriana on A
 - c) $P(A) = .7$ ja referenssikategoriana on B.
6. Todennäköisyys sille, että koehenkilö saa epileptisiä oireita transkraniaalisen magneettistimulaation aikana on $P(A) = .01$. Määritä riskit ja logit, jos referenssikategoriana on
 - a) henkilö saa epileptisiä oireita
 - b) henkilö ei saa epileptisiä oireita.
7. Millaisia oletuksia logistinen regressioanalyysi tekee malliin tulevien muuttujien jakaumista?
8. Milloin logistista regressiomallia kannattaa käyttää lineaarisen regressiomallin sijaan?

9. Miten eri tavat käyttää χ^2 -testisuuretta arvioivat logistisen regressiomallin sopivuutta, ja miksi niitä molempia tulee käyttää mallin arvioimiseen?
10. Miten logistisen regressiomallin ennustustarkkuutta voidaan arvioida?
11. Tutkimuksessa ($n = 200$) on selvitetty logistisen regressio-analyysin avulla tekijöitä, jotka ennustavat skitsofrenian esiintymistä sellaisilla lapsilla, joiden vanhemmista vähintään toinen on sairastanut skitsofreniaa. Tutkimukseen valituilla aikuisilla henkilöillä ei ole diagnosoitu skitsofreniaa tutkimuksen aloittamishetkellä. Heitä seurattiin kymmenen vuoden ajan, ja tutkimuksen päättyessä 20 henkilöä oli sairastunut skitsofreniaa. Mitä voit sanoa seuraavien mallien luokittelutarkkuuksista?
 - a) Malli luokittelee kaikista havainnoista oikein 170, sairastumattomista 160 ja sairastuneista 10.
 - b) Malli luokittelee kaikista havainnoista oikein 165, sairastumattomista 150 ja sairastuneista 15.
 - c) Malli luokittelee kaikista havainnoista oikein 98, sairastumattomista 80 ja sairastuneista 18.
12. Mieti esimerkkejä tutkimusasetelmista, joiden analysoimiseen voitaisiin käyttää logistista regressioanalyysia.

Latentit muuttujat

V

20 Mittaaminen, reliabiliteetti ja validiteetti

1. Tarkastellaan yhden henkilön kuvitteellisia mittaustuloksia älykkyyttä mittaavassa testissä. Määritä mittaukseen liittyvä
 - a) havaittu pistemäärä, kun $T = 75$ ja $e = 25$
 - b) mittarin todellinen pistemäärä, kun $X = 100$ ja $e = 20$
 - c) mittarin todellinen pistemäärä, kun $X = 85$ ja $e = -10$
 - d) mittausvirhe, kun $X = 100$ ja $T = 120$
2. Mikä on
 - a) reliabiliteettikertoimen vaihteluväli
 - b) validiteettikertoimen vaihteluväli?
3. Määritä mittauksen reliabiliteetti, kun
 - a) $\sigma_T^2 = 80$ ja $\sigma_X^2 = 95$
 - b) $\sigma_T^2 = 40$ ja $\sigma_X^2 = 100$
 - c) $\sigma_T^2 = 95$ ja $\sigma_X^2 = 95$
 - d) $r_{XT} = .75$
 - e) $r_{XT} = .50$
4. Miten havaitut pistemäärät käyttäytyisivät jos
 - a) mittausvirhe korreloisi todellisen pistemäärän kanssa positiivisesti
 - b) mittausvirheen keskiarvo olisi nollaa suurempi
 - c) mittausvirheen keskiarvo olisi nollaa pienempi?
5. Kyselylomakemittari koostuu kymmenestä väittämästä. Määritä mittarin reliabiliteetti, kun väittämien interkorrelaatioiden keskiarvo on
 - a) .85
 - b) .60
 - c) .50

6. **Tarkastellaan Tunteiden säätelykysely -aineistoa (liite 13). Määritä**
- a) uudelleenarviointia mittaavan lomakkeen osan sisäinen konsistenssi (Cronbachin α)
 - b) ekspressiivistä suppressiota mittaavan lomakkeen osan sisäinen konsistenssi (Cronbachin α)
 - c) uudelleenarviointia mittaavan lomakkeen osan split half -reliabiliteetti, jossa käsittelet ensimmäistä kolmea väittämää (U1-U3) ensimmäisenä ja kolmea viimeistä väittämää (U4-U6) toisena mittarin puoliskona
 - d) uudelleenarviointia mittaavan lomakkeen osan split half -reliabiliteetti, jossa käsittelet parittomia väittämiä (U1, U3, U5) ensimmäisenä ja parillisia väittämiä (U2, U4, U6) toisena mittarin puoliskona.
7. **Mittarin alkuperäinen reliabiliteetti on .5. Mikä on mittarin uusi reliabiliteetti, jos sitä pidennetään homogeenisiä itemeitä lisäämällä**
- a) 1.5-kertaiseksi?
 - b) kaksinkertaiseksi?
 - c) kolminkertaiseksi?
8. **Kannattaako kyselylomakemittarin reliabiliteettia kasvattaa väittämiä (itemeitä) lisäämällä, jos**
- a) alkuperäisen testin reliabiliteetti on .6 ja siinä on viisi itemiä
 - b) alkuperäisen testin reliabiliteetti on .94 ja siinä on viisitoista itemiä
 - c) alkuperäisen testin reliabiliteetti on .6 ja siinä on kymmenen itemiä
 - d) alkuperäisen testin reliabiliteetti on .6 ja siinä on sata itemiä?
9. **Mittarin alkuperäinen reliabiliteettikerroin on .80. Määritä mittarin uusi validiteettikerroin, jos**
- a) mittarin alkuperäinen validiteettikerroin on .5 ja mittari pidennetään kaksinkertaiseksi
 - b) mittarin alkuperäinen validiteettikerroin on .9 ja mittarin pituus puolitetaan
 - c) mittarin alkuperäinen validiteettikerroin on .5 ja mittarin pituus kymmenkertaistetaan.

10. Määritä validiteettikertoimen teoreettinen yläraja, kun mittauksen validiteetti on .30 ja reliabiliteetti on
- .50
 - .80
 - .95
11. Tarkastellaan Tunteiden säätelykysely -aineistoa (liite 13). Uudelleenarviointia mittaavan lomakkeen osan validiteetti on .60 ja ekspressiivistä suppressiota mittaavan lomakkeen osan validiteetti on .50. Määritä validiteettikertoimen teoreettinen yläraja
- uudelleenarviointia mittaavalle lomakkeen osalle .69
 - ekspressiivistä suppressiota mittaavalle lomakkeen osalle. Käytä reliabiliteetin estimaattina Cronbachin α :aa.
12. Voiko
- validiteettikerroin olla reliabiliteettikerrointa suurempi
 - reliabiliteettikerroin olla validiteettikerrointa suurempi
 - validiteettikerroin olla tasan 1?

21 Reliabiliteetin vaikutus mittaustuloksiin

1. Määritä mittauksen keskivirhe, kun
- $\sigma = 10$ ja $r_{xx} = .90$
 - $\sigma = 20$ ja $r_{xx} = .90$
 - $\sigma = 20$ ja $r_{xx} = .50$.
2. Määritä
- todellisten pistemäärien keskiarvo, kun havaittujen pistemäärien keskiarvo on 100
 - havaittujen pistemäärien keskiarvo, kun todellisten pistemäärien keskiarvo on 125.
3. Milloin $\sigma_T^2 = \sigma_X^2$?

4. Miten havaittujen pistemäärien varianssi käyttäytyy, jos mittauksen reliabiliteetti
 - a) pienenee
 - b) kasvaa?
5. Määritä attenuaatiokorjatut korrelaatiokertoimet jos mittaukset x ja y oletettaisiin virheettömiksi ja
 - a) $r_{xy} = .5$, $r_{xx} = .8$ ja $r_{yy} = .8$
 - b) $r_{xy} = .5$, $r_{xx} = .6$ ja $r_{yy} = .6$
6. Määritä attenuaatiokorjatut korrelaatiokertoimet jos mittauksien x ja y uusiksi reliabiliteetiksi oletettaisiin $.7$, ja alkuperäisten mittausten
 - a) $r_{xy} = .5$, $r_{xx} = .5$ ja $r_{yy} = .6$
 - b) $r_{xy} = .7$, $r_{xx} = .5$ ja $r_{yy} = .7$
7. Tarkastellaan Tunteiden säätelykysely -aineistoa (liite 13). Laske uudelleenarviointia mittaavien väittämien ja ekspressiivistä suppressiota mittaavien väittämien summapistemäärien välinen attenuaatiokorjattu korrelaatiokerroin jos
 - a) mittaukset oletetaan täysin virheettömiksi
 - b) molempien mittausten reliabiliteetiksi oletetaan $.8$.
8. Mikä selittää sen, että poikkeuksellisen *hyvän* suorituksen jälkeen urheilija todennäköisesti onnistuu seuraavassa suorituksessaan *huonommin*, ja poikkeuksellisen *huonon* suorituksen jälkeen taas *paremmin*?
9. Tutkimuksessa on selvitetty miten säännölliset ryhmäohjaustapaamiset vaikuttavat työntekijöiden stressiin. Miksi on todennäköistä, että ohjaustapaamisten havaittu vaikutus on suurempaa työntekijöillä jotka kokevat voimakasta kuin keskinkertaista stressiä?
10. Miten mittaustulosten palautumiselle kohti keskiarvoa kävisi, jos $\sigma_e^2 = 0$?

11. Tutkimuksessa on tarkasteltu, miten neuropsykologinen kuntoutus vaikuttaa ADHD:sta kärsivien aikuisten tarkkaavaisuustoimintoihin. Tutkimustulosten analysoimista varten muodostetaan erotusmuuttuja tutkittavien tarkkaavaisuustoimintojen tasosta ennen kuntoutusta ja kuntoutuksen jälkeen.. Laske erotusmuuttujan reliabiliteetti, kun

a) $r_{xx} = .8$, $r_{yy} = .7$ ja $r_{xy} = .7$

b) $r_{xx} = .8$, $r_{yy} = .7$ ja $r_{xy} = .4$

12. Tutkimuksessa on mitattu potilaiden masentuneisuutta kahdelle eri mittarilla: Beckin ja Hamiltonin depressioinventaa-reilla. Tutkimustulosten analysoimista varten mittaustuloksista lasketaan summamuuttuja. Määritä summamuuttujan reliabiliteetti, kun

a) $r_{xx} = .8$, $r_{yy} = .8$ ja $r_{xy} = .7$

b) $r_{xx} = .7$, $r_{yy} = .6$ ja $r_{xy} = .5$

22 Efektikoko

1. Millaisessa tilanteessa

- a) efektikoko voi olla pieni vaikka p-arvo on tilastollisesti erittäin merkitsevä
- b) efektikoko voi olla suuri vaikka p-arvo ei olekaan tilastollisesti merkitsevä?

2. Miten *d*-tyyppinen efektikoko käyttäytyy, jos

- a) havaittu keskiarvojen erotus pysyy vakiona, ja ryhmien hajonnat kasvavat
- b) havaittu keskiarvojen erotus pysyy vakiona, ja ryhmien hajonnat pienenevät
- c) havaittu keskiarvojen erotus kasvaa, ja ryhmien hajonnat pysyvät vakiona
- d) havaittu keskiarvojen erotus kasvaa ja ryhmien hajonta kasvaa
- e) kaikki muut tekijät pysyvät vakiona ja otoskoko kasvaa?

3. Tutkimuksessa on selvitetty kuinka hyvin amerikkalaiset ja japanilaiset koehenkilöt tunnistavat amerikkalaisten näyttelijöiden esittämiä kasvonilmeitä. Laske Cohenin d kun tiedetään, että
- amerikkalaiset tutkittavat tunnistavat ilmeitä keskimäärin 95%:n tarkkuudella, japanilaiset 80%:n tarkkuudella ja ryhmien yhdistetty keskihajonta on 10
 - amerikkalaiset tutkittavat tunnistavat ilmeitä keskimäärin 90%:n tarkkuudella, japanilaiset 88%:n tarkkuudella ja ryhmien yhdistetty keskihajonta on 2
 - amerikkalaiset tutkittavat tunnistavat ilmeitä keskimäärin 90%:n tarkkuudella, japanilaiset 90%:n tarkkuudella ja ryhmien yhdistetty keskihajonta on 4.
 - Määritä ovatko edellä laskemasi efektitkoot pieniä, keskinkertaisia vai suuria.
4. Tarkastellaan sokeiden työmuisti -tutkimusta (Liite 9). Vertailaan sokeiden ja näkevien taktiilista työmuistia. Laske tälle vertailulle
- Cohenin d -efektikoko
 - $Glassin \Delta$ -efektikoko käyttäen näkevien ryhmää kontrolliryhmänä
 - $Glassin \Delta$ -efektikoko käyttäen sokeiden ryhmää kontrolliryhmänä.
 - Määritä ovatko edellä laskemasi efektitkoot pieniä, keskinkertaisia vai suuria
5. Tarkastellaan Koulukiusaaminen-aineistoa (Liite 2). Laske r -efektikoko
- kiusatuksi tulemisen ja itsetunnon väliselle yhteydelle
 - itsetunnon ja aggression väliselle yhteydelle.
6. Muunna seuraavat r -efektikoot Cohenin d :ksi:
- .50
 - .78
 - .55
7. Muunna seuraavat Cohenin d -efektikoot r -efektikoksi
- 1
 - 1.5
 - 0.5

8. Tutkimuksessa on selvitetty uuden ahdistuneisuutta lievittävän lääkkeen tehoa. Ahdistuneisuuden muutosta on vertailtu lääkeainetta ja plaseboa käyttävissä ryhmissä. Miten tutkimuksen kanssa kannattaa edetä, jos vertaillaan ahdistuneisuuden muutosta lääkeaine- ja plasebo -ryhmissä, ja havaitaan että
- $t(9) = 2.10, p = .63, d = 1.28?$
 - $t(499) = 1.5, p = .23, d = .11?$
 - $t(9) = 4.60, p = .001, d = 2.06?$
 - $t(499) = 5.00, p < .001, d = 0.5?$
9. Tutkimuksessa on selvitetty tekijöitä, jotka selittävät tutkittavien kyselylomakkeella mitattua tyyppillistä ruoanhimoa. Tutkimukseen osallistui 150 miestä ja 150 naista, joista kolmannes oli normaalipainoisia, kolmannes lievästi ylipainoisia ja kolmannes huomattavasti ylipainoisia. Tutkimusasetelma oli siis 2 (sukupuoli) $\times 3$ (paino). Seuraavat neliösummat ovat tiedossa

Tekijä	Neliösumma
Total	200
Sukupuoli	20
Paino	120
Sukupuoli $\times$ paino	50
Error	10

Laske sekä $\hat{\eta}^2$ - että $\hat{\eta}_p^2$ -efektikoko

- sukupuolen päävaikutukselle
 - painon päävaikutukselle
 - sukupuolen ja painon yhdysvaikutukselle.
10. Kolme toisistaan riippumatonta tutkimusta selvitti, kuinka aikuisten suomalaisten sosioekonomista asemaa (1–10) voidaan ennustaa heidän vanhempiansa aikuisiän sosioekonomisella asemalla. Miten tulkitsisit tulokset, jos
- tutkimusten otoskoot olisivat 100, 100 ja 100 henkilöä, ja vastaavat r -tyyppiset efektikoot .67, .52 ja .53
 - tutkimusten otoskoot olisivat 20, 20 ja 1000 henkilöä, ja vastaavat r -tyyppiset efektikoot .10, .42 ja .53
 - tutkimusten otoskoot olisivat 20, 20 ja 20 henkilöä, ja vastaavat r -tyyppiset efektikoot .02, .56 ja .15?

23 Faktorianalyysi

1. Mitä tarkoitetaan faktorilla?
2. Mitä faktorien avulla pyritään kuvaamaan aineistosta?
3. Miksi faktorianalyysi helpottaa suurten korrelaatiomatriisien tulkitsemista?
4. Millaisia oletuksia faktorianalyysi tekee analysoitavasta aineistosta?
5. Kuinka suuri aineisto suunnilleen tarvitaan, jos halutaan muodostaa
 - a) 100 muuttujasta 50 faktoria
 - b) 50 muuttujasta 10 faktoria
 - c) 20 muuttujasta kolme faktoria?
6. Miksi faktorianalyysin tulosten tulkitsemisessä on noudatettava erittäin suurta huolellisuutta?
7. Mieti, miten faktorianalyysin tulosten tulkita muuttuisi, jos havaittujen muuttujien korrelaatiomatriisin sijaan analysoitaisiinkin muuttujien kovarianssimatriisia.
8. Miten
 - a) ryhmäfaktorit
 - b) yleiset faktorit
 - c) spesifit faktorit selittävät aineistossa havaittujen muuttujien vaihtelua?
9. Miksi faktoriratkaisu kannattaa joskus rotatoida?
10. Mitä eroa on suorakulmaisella ja vinorotaatiolla? Millaisissa tilanteissa eri rotaatioratkaisuja kannattaa käyttää?
11. Millaista informaatiota faktoriratkaisusta saadaan tulkitsemalla
 - a) latauksia
 - b) faktorien ominaisarvoja
 - c) kommunaliteetteja?

12. Millä välillä vaihtelevat

- a) faktorien lataukset
- b) faktorien ominaisarvot
- c) muuttujien kommunaliteetit?

13. Millaisissa tilanteissa soveltuu käytettäväksi

- a) maximum likelihood -ekstraktointi
- b) yleistettyjen neliösummien ekstraktointimenetelmä?

14. Millainen on hyvä faktoriratkaisu?

15. Tutkimuksessa selvitettiin sosiaalisen älykkyyden eri osatekijöiden välisiä yhteyksiä. Tutkimukseen osallistui 500 suomalaista aikuista (256 miestä, 244 naista). Tutkittavilta mitattiin kyselylomakkeiden avulla *empatiakyky (V1)*, *emootioiden tunnistuskyky (V2)*, *toisten ihmisten toiminnan attribuointikyky (V3)*, *sosiaalinen ongelmanratkaisutaito (V4)* ja *sosiaalinen joustavuus (V5)*. Kaikki muuttujat olivat normaalisti jakautuneita. Muuttujien välinen korrelaatiomatriisi oli seuraavanlainen:

	V1	V2	V3	V4	V5
V1	1	0.74*	0.82*	0.21	0.23
V2	0.74*	1	0.79*	0.16	0.14
V3	0.82*	0.79*	1	0.12	0.12
V4	0.21	0.16	0.12	1	0.81*
V5	0.23	0.14	0.12	0.81*	1

Tilastollisesti merkitsevät korrelaatiot on merkitty matriisiin asteriskilla (*). Tutkimuksen tarkoituksena oli selvittää faktorianalyysin avulla, miten eri sosiaalisen älykkyyden muodot ovat yhteydessä toisiinsa. Vastaa korrelaatiomatriisiin perusteella seuraaviin kysymyksiin.

- a) Mitä ekstraktointi- ja rotatointimenetelmää analyysissa pitäisi käyttää? Miksi?
- b) Kuinka monta faktoria aineistosta todennäköisesti ekstraktoituu? Miksi?
- c) Arvioi näiden faktoreiden lataukset eri muuttujille.
- d) Arvioi myös summittaisesti faktorien ominaisarvot ja muuttujien kommunaliteetit.

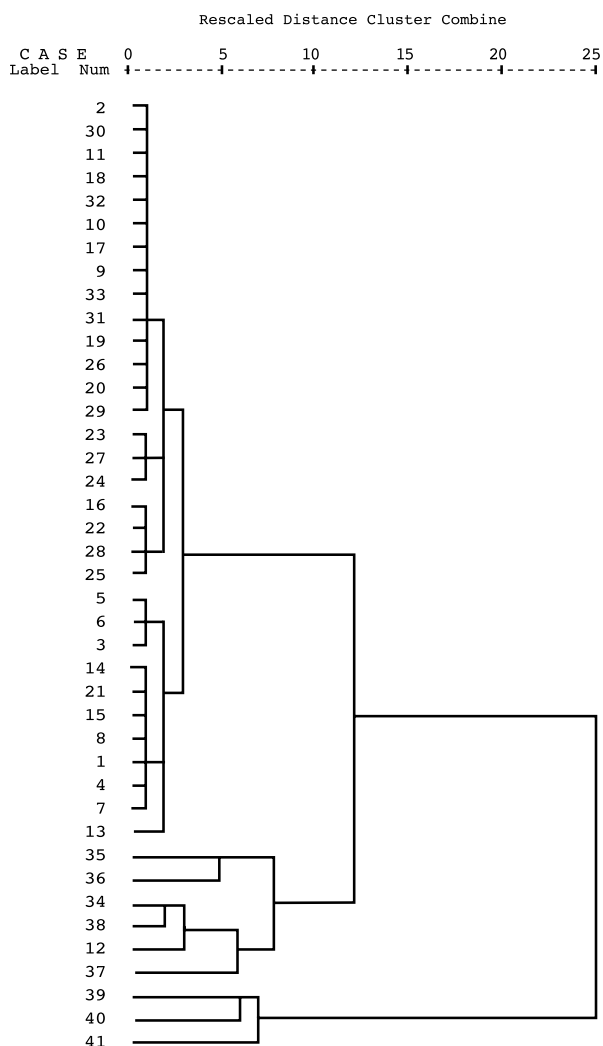
24 Moniulotteinen asteikointi

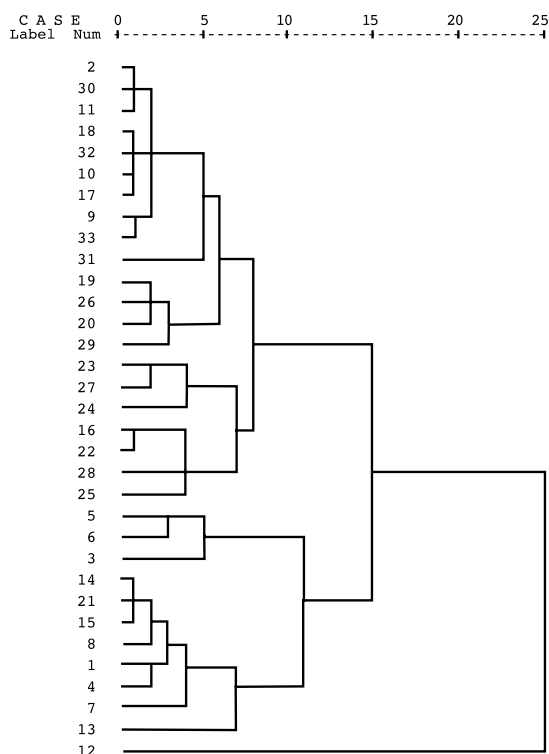
1. Miten faktorianalyysin, moniulotteinen asteikoinnin ja klusterianalyysin teoriat eroavat toisistaan?
2. Millaisiin tutkimuskysymyksiin edellä mainitut menetelmät soveltuvat?
3. Millaisilla menetelmillä havaintojen väliset etäisyydet voidaan määrittää?
4. Miten moniulotteisen asteikointiratkaisun toimivuus voidaan arvioida?
5. Miten stressiarvot S_0 ja S_1 eroavat toisistaan?
6. Miksi moniulotteisen asteikoinnin avulla ei voida tarkastella muuttujien välistä yhteisvaihtelua?
7. Tarkastellaan tutkimusta, jossa halutaan selvittää kuissa esitettyjen vartalon asentojen havaittua samankaltaisuuksia. Jos asentojen samankaltaisuudet määritellään suoraa arviointia käyttäen, kuinka monta ärsykeparia tutkittavien tulee arvioida, jos asentokuvia on
 - a) 10
 - b) 25
 - c) 100 kappaletta?

25 Klusterianalyysi

1. Millaisissa tutkimuksissa klusterianalyysi voisi sopia aineiston analysoimiseen?
2. Mieti esimerkkejä tutkimuksista, joiden aineistojen analysoimiseen klusterianalyysi voisi soveltua.
3. Miksi muuttujille on suoritettava z-muunnos ennen klusterianalyysin suorittamista?

4. Mitä ongelmia liittyy klusterianalyysin käyttöön tutkimusaineiston analysoimisessa?
5. Miten k-keskiarvoklusterointi eroaa hierarkkisesta klusteroinnista?
6. Montako klusteria näyttäisi dendogrammin perusteella muodostuvan seuraavissa aineistoissa?





26 Rakenneyhtälömallit

1. Mitä faktorianalyysin ongelmia voidaan ratkaista siirtymällä käyttämään rakenneyhtälömalleja?
2. Mitä eroa on havaituilla ja latenteilla muuttujilla?
3. Mitä eroa rakenneyhtälömalleissa on endogeenisillä ja eksogeenisillä muuttujilla?
4. Miten testataan rakenneyhtälömallin sopivuus aineistoon?
5. Mitä kuvaa
 - a) havaittujen muuttujien polkumalli
 - b) mittamalli
 - c) rakennemalli?

6. Millä perusteella rakenneyhtälömalleja muodostettaessa määritellään muuttujien (sekä havaittujen että latenttien) väliset yhteydet?
7. Mieti, millaisilla edellytyksillä rakenneyhtälömalleja voidaan käyttää mitattuja muuttujia koskevien kausaalipäätelmien tekemiseen.
8. Tarkastellaan tutkimusta, jossa on mitattu yläasteikäisten lasten
 - 1) *empatiakyky*, 2) *emootioiden tunnistuskyky*, 3) *toisten ihmisten toiminnan attribuointikyky*, 4) *sosiaalinen ongelmanratkaisutaito*, 5) *sosiaalinen älykkyys* ja 6) *sosiaalinen joustavuus*. Hahmottele seuraavat rakenneyhtälömallit kuvaajien avulla:
 - a) Havaittujen muuttujien polkumalli, jossa empatiakyky ja sosiaalinen älykkyys selittävät sosiaalista ongelmanratkaisutaitoa, joka selittää edelleen sosiaalista joustavuutta. Lisää malliin vielä emootioiden tunnistamiskyky selittämään sosiaalista joustavuutta.
 - b) havaittujen muuttujien polkumalli, jossa emootioiden tunnistuskyky selittävät toisten ihmisten toiminnan attribuointikykyä, joka selittää edelleen sosiaalista ongelmanratkaisutaitoa, sosiaalista älykkyyttä ja sosiaalista joustavuutta
 - c) mittamalli, jossa kaksi latenttia muuttujaa (η_1 , η_2) selittää kaikkia havaittuja muuttujia
 - d) rakennemalli, jossa on kolme latenttia muuttujaa (η_1 , η_2 , η_3) joista η_1 selittää η_2 :ta ja η_3 :a ja η_2 selittää η_3 :a
 - e) mittamalli, jossa yksi latentti muuttuja (η_1) selittää empatiakyvyn ja emootioiden tunnistamiskyvyn, toinen latentti muuttuja (η_2) toisten ihmisten toiminnan attribuointikyvyn ja sosiaalisen ongelmanratkaisutaidon ja kolmas latentti muuttuja (η_3) s sosiaalisen älykkyyden ja sosiaalisen joustavuuden
 - f) yhdistä edelliseen mittamalliin rakennemalli, jossa latentti muuttuja (η_1) selittää latenttia muuttujaa (η_2), joka selittää edelleen latenttia muuttujaa (η_3).
9. Määrittele, mitkä muuttujat ovat endogeenisiä ja mitkä eksogeenisiä edellisen tehtävän kohdissa d–f.

Vastaukset

Mitä tilastotiede on?

1. Tutkimusaineiston analysoimisessa käytettävien menetelmien kehittäminen
2. Muodostaa yleisiä lainalaisuuksia meitä ympäröivän maailman toiminnasta
3. Suurten lukujoukkojen mahdollisesti kuvaamien ilmiöiden ymmärtäminen on ihmisille vaikeaa, joten lukujoukkojen sisältämää informaatiota täytyy tiivistää. Lisäksi mittaustuloksista tarvitsee tehdä päätelmiä, jotka yleistyvät suurempaan mittausjoukkoon.
5. Aineistoa kuvaavat, tiivistävät ja mallintavat

2 Tutkimuksen suunnitteleminen

1. Data → informaatio → tieto
2. Hyvin perusteltu, tosi uskomus
4. Koko populaation mittaaminen on useissa tapauksissa mahdotonta tai vähintäänkin kallista ja aikaa vievää.
5. Jos otos muistuttaa ominaisuuksiltaan riittävän paljon populaatiota, voidaan olettaa, että otoksessa havaitut ilmiöt toistuvat myös populaatiossa.
6. a) 0.1 %
b) 0.2 %
c) 2 %
7. a) poimintaväli = 3, otos = 1, 4, 7, 10, 13, 16, 19, 22, 25, 28
b) poimintaväli = 6, otos = 1, 7, 13, 19, 25
8. Yleisesti ottaen voidaan todeta, että mitä suurempia yksilöllisiä eroja tarkasteltavassa ominaisuudessa oletetaan havaittavan, sitä suurempi aineisto tarvitaan.
9. a) korrelatiivinen tutkimus, riippuvat muuttujat taktiilinen ja fonologinen työmuisti sekä koehenkilön näkökyky
b) korrelatiivinen tutkimus, riippuvat muuttujat tekstin lukuaika ja kokonaispalautus sekä koehenkilön työmuistin kapasiteetti
c) korrelatiivinen tutkimus, riippuvat muuttujat sisarusten lukumäärä, kiusatuksi tuleminen kokemus, itsetunto, aggressiivisuus ja työmuistin kapasiteetti
d) kokeellinen tutkimus, riippuva muuttuja reaktioaika, riippumaton muuttuja vihje- ja kohdeärsykkeiden kongruenssi

3 Mittaaminen ja aineisto

1. Symbolin liittämistä mittauskohteeseen jonkin säännön mukaan
2. Kvalitatiiviset muuttujat kuvaavat mittauskohteiden laatua tai tyyppiä luokkien avulla. Kvantitatiiviset muuttujat taas kuvaavat mittauskohteen suuruutta, lukumäärää tai järjestystä numeroiden avulla.
3. Kvalitatiiviset muuttujat: sukupuoli. Kvantitatiiviset muuttujat: reaktio-aika, koulutus, oppilaan luokka-aste
4. a) sukupuoli = laatueroasteikko, ikä = suhdeasteikko, sisarusten lukumäärä = järjestysasteikko, kiusatuksi tulemisen kokemus = järjestysasteikko, itsetunto = järjestysasteikko, aggressio = järjestysasteikko, työmuistin kapasiteetti = järjestysasteikko
b) sukupuoli = laatueroasteikko, ikä = suhdeasteikko, koulutus = järjestysasteikko, itsetunto = suhdeasteikko, työssä viihtyminen = suhdeasteikko, tutkittavan kokemus työilmapiiristä = suhdeasteikko, tutkittavan kokemus työn kuormittavuudesta = suhdeasteikko
c) sukupuoli = laatueroasteikko, ikä = suhdeasteikko, kokeellinen ryhmä, johon tutkittava kuului = laatueroasteikko, tutkittavan depressio tutkimuksen alussa mitattuna BDI-II-lomakkeella = järjestysasteikko, tutkittavan depressio tutkimuksen lopussa mitattuna BDI-II-lomakkeella = järjestysasteikko
5. a) laatueroasteikolla
b) järjestysasteikolla
c) järjestysasteikolla tai välimatka-asteikolla
d) järjestysasteikolla, välimatka-asteikolla, suhdeasteikolla ja absoluuttisella asteikolla
6. a) luokkien nimiöt voidaan vaihtaa, mutta invarianssin käsite ei ole mielekäs laatueroasteikolla
b) monotonisten muunnosten (jos $a < b$, niin $f(a) < f(b)$)
c) $X' = cX + b$
d) $X' = cX$
7. Jokainen muuttuja syötetään yhtenä sarakkeena ja yhdeltä tilastoyksiköltä mitatut havainnot syötetään aina yhdelle riville.
8. a) 30 tilastoyksikköä, 6 muuttujaa ja 180 havaintoa
b) 20 tilastoyksikköä, 8 muuttujaa ja 160 havaintoa
c) 20 tilastoyksikköä, 10 muuttujaa ja 200 havaintoa
d) 15 tilastoyksikköä, 5 muuttujaa ja 75 havaintoa

5 Aineiston kuvaaminen numeerisesti

1. Jotta voitaisiin tiivistää mittaus-
tulokset ja kuvailla muuttujien
arvojen tyyppillinen suuruus tai
havaintojen suuruuksien vaihtelu
yhden lukuarvon avulla
2. Parametrit ovat populaation
ominaisuuksia, joiden suuruus
pyritään arvioimaan otoksen
perusteella laskettujen estimaattien
avulla
3. a) ei mitään
b) vaihteluväli
c) vaihteluväli, varianssi, keski-
hajonta ja variaatiokerroin
d) vaihteluväli, varianssi, keski-
hajonta ja variaatiokerroin
4. a) moodi
b) moodi, mediaani, kaikki
fraktiilit ja fraktiilivälien
pituudet
c) moodi, mediaani, kaikki
fraktiilit ja fraktiilivälien
pituudet, keskiarvo
d) moodi, mediaani, kaikki
fraktiilit ja fraktiilivälien
pituudet, keskiarvo
5. a) 55
b) 140
c) 210
d) 70
e) 45
6. Jos jakauma ei ole symmetrinen, ts.
jos jakauma on vino tai jos jakauma
on useampihioppinen
7. a) $\bar{x} = MO$
b) $MO < \bar{x}$
c) $MO > \bar{x}$
8. a) 58
b) 29.29
c) 5.13
d) 57.8
e) 9 %
f) min = 49, max = 67
9. a) 20
b) 51.04
c) 67
d) 2 341
e) 150
10. a) 10 tyttöä, 10 poikaa
b) vaihteluväli = 2, MD = 1
c) tyttö ja poika (10 havaintoa)
d) $MO = 1, \bar{x} = 1.10$
e) $\bar{x} = 4.55, s = 2.50$
f) $\bar{x} = 4.75, s^2 = 6.04$
11. a) $MO_{\text{sukupuoli}} = \text{nainen}, MO_{\text{kätisyys}} = \text{oikea}$
b) MD = 21, v = 12.59 %
c) $s^2 = 2.93, s = 1.71, v = 24.08 \%$
d) mielihyvä: $\bar{x} = 7.10, MD = 8,$
viha: $\bar{x} = 1.50, MD = 1,$
inho: $\bar{x} = 1.55, MD = 1,$
pelko: $\bar{x} = 2.95, MD = 2.50,$
suru: $\bar{x} = 2.30, MD = 2,$
hämmästys: $\bar{x} = 2.90,$
MD = 2.50

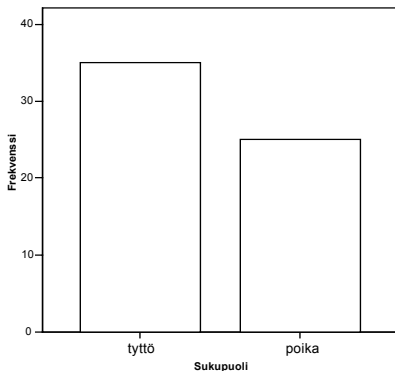
12. a) keskiaste 15, korkea-aste 15
 b) MO = mies (18 havaintoa)
 c) min = 2, max = 10
 d) MD = 4.00
 e) MO = 2, 4, 5 ja 7

13. a) $MD_0 = 98$, $MD_{90} = 92$, $MD_{180} = 69.50$
 b) $MO_0 = 98$ ja 99 , $MO_{90} = 93$, $MO_{180} = 71$ ja 72
 c) $\bar{x}_0 = 97.60$, $\bar{x}_{90} = 91.60$, $\bar{x}_{180} = 69$

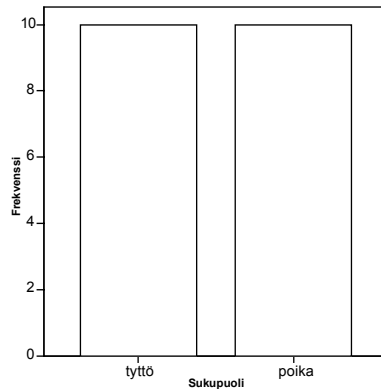
6 Aineiston kuvaaminen graafisesti

1. a) Sektoridiagrammi ei yleisesti ottaen sovellu aineiston graafiseen esittämiseen.
 b) Histogrammi sopii jatkuvien muuttujien alustavaan tarkasteluun, mutta yleensä sitä ei käytetä tutkimustulosten esittämisessä.
2. Varjostukset ja kolmiulotteiset kuvaajien osat, epäselvä väritys sekä x- ja y-akselien skaalojen huolimaton valinta
3. a) Pylväsdiaagrammia, viivadiaagrammia ja sektoridiaagrammia
 b) Histogrammia, laatikko-janakuviota, pylväsdiaagrammia ja viivadiaagrammia

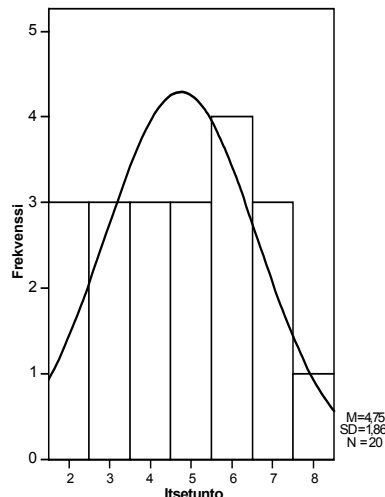
4.

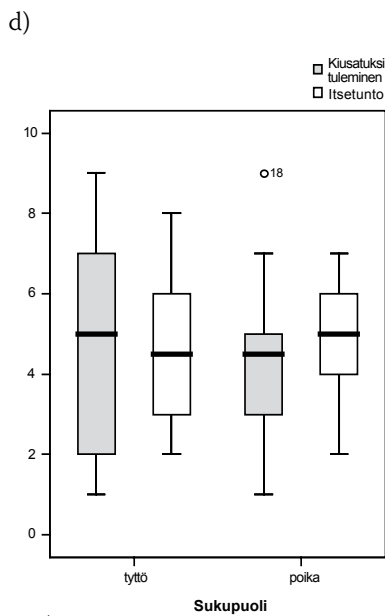
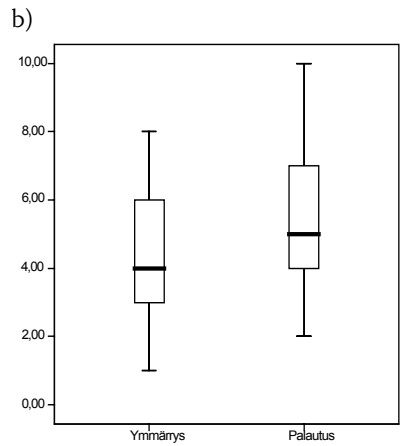
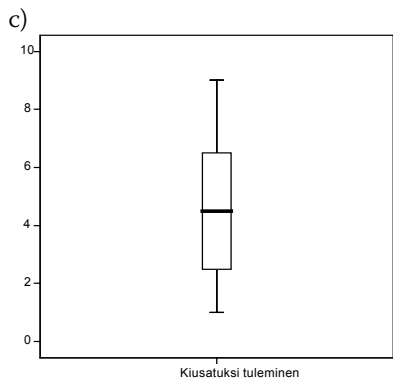


5. a)

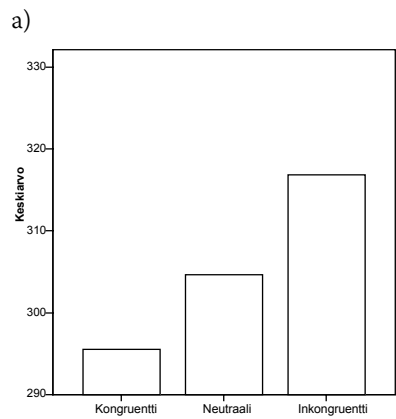


b)

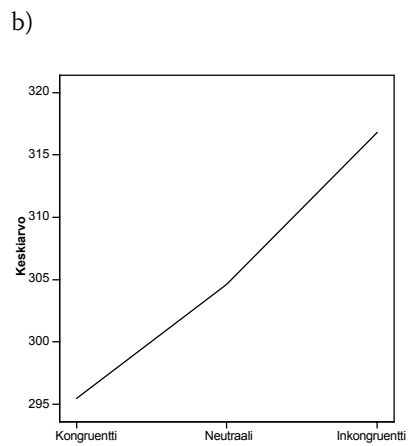
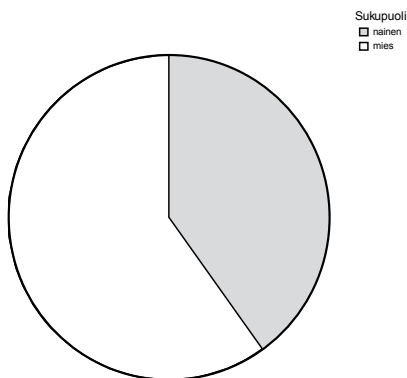




7.



6. a)



7 Todennäköisyyslaskennan perusteet

1. a) Stokastinen
b) Deterministinen
c) Stokastinen
d) Deterministinen
2. a) On.
b) Ei ole.
c) On.
3. a) Kyllä
b) Ei
c) Ei
d) Kyllä
4. a) Ei ole.
b) Ei ole.
c) On.
d) Ei ole.
5. a) $\bar{A} = \{1,3,5,7,9\}$
b) $A \cap B = \{2,8,10\}$
c) $A \cup D = \{2,4,6,7,8,9,10\}$
d) $E \cap B = \{1,3,4,5,6,7,9\}$
6. $P(\text{väärä diagnoosi})=0.7$
7. a) $P(\text{pata}) = 1/4$
b) $P(\text{herttaässä}) = 1/52$
c) $P(\text{ruutu tai nelonen}) = 4/13$
d) $P(\text{ei herttakymppi tai ruutu}) = 19/26$
e) $P(\text{ei pata tai vitonen})= 9/13$
8. a) $P(\text{yksi kutonen}) = 1/12$
b) $P(\text{kaksi kuutosta}) = 1/36$
c) $P(\text{kolme vitosta}) = 0$
d) $P(\text{summa vähintään } 5) = 5/6$
e) $P(\text{summa vähintään } 2) = 1$
9. a) $P(\text{tehoa kolmeen}) = 0.51$
b) $P(\text{tehoa neljään}) = 0.41$
c) $P(\text{tehoa viiteen}) = 0.33$
10. a) $P(\text{kahdelta}) = 0.94$
b) $P(\text{kuudelta}) = 0.83$
c) $P(\text{kymmeneltä}) = 0.74$
11. a) $P(\text{kahdelta}) = 0.88$
b) $P(\text{kolmelta}) = 0.83$
c) $P(\text{viideltäkymmeneltä}) = 0.05$
12. a) 6 järjestyksessä
b) 120 järjestyksessä
c) 40 320 järjestyksessä
d) 3 628 800 järjestyksessä
13. a) 720
b) 30 240
c) 1 814 400
d) 3 628 800
14. a) 210 tavalla
b) 252 tavalla
c) 120 tavalla

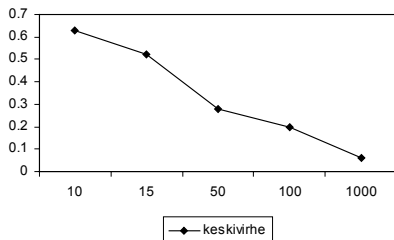
8 Todennäköisyysjakaumat

1. a) $P(\text{enintään } 10) = 11/12$
 b) $P(\text{vähintään } 12) = 1/36$
 c) $P(\text{tulos väliltä } 6-8) = 4/9$
 d) $P(\text{enintään } 2) = 1/36$
2. a) Useimmin esiintyvä tulos, jos satunnaisilmiö toistuu ∞ kertaa
 b) Satunnaisilmiön lopputuloksen suuruuden tyypillisin poikkeama odotusarvosta, jos satunnaisilmiö toistuu ∞ kertaa
3. a) $EX = 7$
 b) $D^2X = 5.83$
 c) $DX = 2.41$
4. a) $EX = -2\,000\text{ €}$
 b) $EX = 5\,000\text{ €}$
 c) $EX = 5\,500\text{ €}$
 d) $EX = 3\,000\text{ €}$
 Tulosten perusteella Kelan kannattaisi suosia kognitiivista lyhytterapiaa.
5. a) $EX = 3$
 b) $D^2X = 2.96$
 c) $DX = 1.72$
6. a) $P(50 \text{ naista}) = 0.08$
 b) $P(45 \text{ miestä}) = 0.05$
 c) $P(10 \text{ miestä}) = 1.37 \cdot 10^{-17}$
 d) $P(\text{enintään } 3 \text{ naista}) = 3.22 \cdot 10^{-24}$
7. a) $P(\text{viisi diagnoosia}) = 0.038$
 b) $P(\text{kolme diagnoosia}) = 0.0076$
 c) $P(\text{kaksitoista diagnoosia}) = 0.094$
8. a) $P(\text{kymmenen elektrodia hajooa}) = 0.10$
 b) $P(\text{viisitoista elektrodia hajooa}) = 0.07$
 c) $P(\text{sata elektrodia hajooa}) = 2.41 \cdot 10^{-50}$
9. a) $P(\text{kaksi}) = 1.86 \cdot 10^{-40}$
 b) $P(\text{viisi}) = 3.10 \cdot 10^{-36}$
 c) $P(\text{kymmenen}) = 1.03 \cdot 10^{-30}$
10. a) 2.28 %
 b) 15.87 %
 c) 4.56 %
 d) 31.74 %
11. a) Keskiarvojen jakaumaa
 b) Keskiarvojen jakaumaa
 c) Varianssien jakaumaa
 d) Kahden varianssin osamäärän jakaumaa
12. a) $DX(t) = 0$
 b) $DX(F) = 1$

9 Tilastolliset testit ja hypoteesien testaaminen

1. Populaation parametrin ja otoksen perusteella lasketun estimaatin välistä eroavaisuutta
2. Sen avulla voidaan osoittaa, että otoskeskiarvojen jakauma noudattaa normaalijakaumaa alkuperäisestä populaatiojakaumasta riippumatta. Täten normaali-jakauma on hyvin käyttökelpoinen kun tehdään otos- ja populaatio-keskiarvoja koskevia päätelmiä.

3. a) $\sigma_{\bar{x}} = 0.63$
b) $\sigma_{\bar{x}} = 0.52$
c) $\sigma_{\bar{x}} = 0.28$
d) $\sigma_{\bar{x}} = 0.2$
e) $\sigma_{\bar{x}} = 0.06$



4. a) $\bar{x} = 1366$
b) {1550, 1320, 1400, 1450}; {1110, 1320, 1400, 1450}; {1110, 1550, 1400, 1450}; {1110, 1550, 1320, 1450}; {1110; 1550; 1320; 1400}
c) $\bar{x}_1 = 1430$
 $\bar{x}_2 = 1320$
 $\bar{x}_3 = 1377.5$
 $\bar{x}_4 = 1357.5$
 $\bar{x}_5 = 1345$
d) $\sigma_{\bar{x}} = 82.81$

5. Keskiarvojen keskivirhe populaatiolle ja otokselle A on 3.16 ja populaatiolle ja otokselle B 2.83. Täten otoksen B perusteella voidaan tehdä todennäköisesti tarkempia estimaatteja.
6. a) $P(z > 1) = 0.1587$
b) $P(z > 2.57) = 0.0051$
c) $P(z > 0.18) = 0.4286$
d) $P(z > 2.12) = 0.0170$
e) $P(z > 1.65) = 0.0495$
7. $z = 4$, $p < .001$ joten aineisto ei tue hypoteesia "suomalaiset lapset katselevat televisiota keskimäärin kaksi tuntia vuorokaudessa".
8. a) $z = 1.05$, $p = 0.1469$. Otos tukee hypoteesia "suomalaisten aikuisten älykkyyssosamäärien keskiarvo on 100".
b) $z = 3.33$, $p = 0.0005$. Otos ei tue hypoteesia "suomalaisten aikuisten älykkyyssosamäärien keskiarvo on 100".
9. $z = 0.32$, $p = 0.3745$. Aineisto tukee hypoteesia "ihmiset pystyvät seuraamaan kerrallaan keskimäärin seitsemää liikkuvaa esinettä".

10. a) $z = 1.68$, $p = .0465$. Aineisto ei tue hypoteesia ”suomalaisten aikuisten fonologisen työmuistin kapasiteetti on keskimäärin 5 yksikköä”.
- b) $z = 0.79$, $p = 0.2148$. Aineisto tukee hypoteesia ”suomalaisten aikuisten fonologisen työmuistin kapasiteetti on keskimäärin 5 yksikköä”.
- c) $z = 1.58$, $p = 0.0571$. Aineisto tukee hypoteesia ”suomalaisten aikuisten fonologisen työmuistin kapasiteetti on keskimäärin 5 yksikköä”.
11. Tyypin 1 virhe on väärä positiivinen ratkaisu (vaihtoehtoinen hypoteesi jää voimaan, vaikka se olisi pitänyt hylätä) ja tyypin 2 virhe puolestaan väärä negatiivinen ratkaisu (nollahypoteesi jää voimaan vaikka se olisi pitänyt hylätä). Tyypin 1 virheen tekemisen todennäköisyyttä voidaan vähentää valitsemalla pienempi kriittinen p-arvon raja ja tyypin 2 virheen tekemisen todennäköisyyttä voidaan puolestaan vähentää valitsemalla suurempi kriittinen raja. Molempien virhetyyppien tekemisen todennäköisyyttä voidaan pienentää käyttämällä voimakkaampaa tutkimusasetelmaa ja poimimalla suurempi otos.
12. Parametriset testit tekevät analysoitavan aineiston jakaumista oletuksia, joiden rikkominen heikentää testien tulosten luotettavuutta. Epäparametriset testit puolestaan eivät tee tällaisia oletuksia.

10 Aineiston valmistelu

1. Jotta voidaan varmistaa että aineisto on syötetty oikein, välttyä siltä että poikkeavat havainnot vaikuttaisivat tunnuslukuihin ja siltä että puuttuvat havainnot pienentäisivät tarpeettomasti otoskokoja.
2. 1) Havainto on syötetty väärin, 2) mittalaitteisto on toiminut virheellisesti, 3) tutkittava on ymmärtänyt tehtävän väärin / ei ole suorittanut tehtävää tarkoitettulla tavalla tai 4) havainto on todellinen.
3. Koska keskiarvon suuruisen havainnon lisääminen jakaumaan ei muuta jakauman keskiarvoa; tällöin menettely on perusteltu, jos tarkastellaan pelkästään muuttujan keskiarvoja.
4. Verrataan toisiinsa joidenkin muiden muuttujien (B, C, D...) arvoja henkilöillä, joilta 1) on havainto muuttujalta A, 2) ei ole havaintoa muuttujalta A.
5. a) $\sqrt{x}$
b) $\ln(k-x)$
c) $\sqrt{k-x}$
d) $\frac{1}{\bar{x}}$

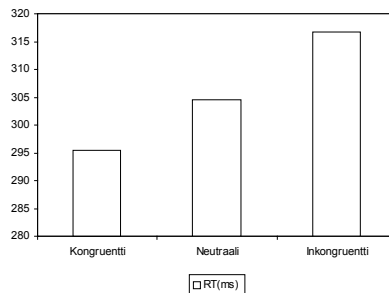
11 T-testit

1. Yhden otoksen t-testin laskemiseksi ei tarvitse tuntea populaation hajontaa.
2. Symmetrinen, keskiarvojen jakaumaa mallintava jakauma jonka odotusarvo on 1. Jakauman muoto riippuu vapausasteista ja t-jakauma alkaa muistuttaa [0,1]-normaalijakaumaa kun $df \rightarrow \infty$
3. Noin 15 havaintoa / testattava muuttuja; testattava muuttuja on jatkuva, mitattu vähintään välimatka-asteikolla ja noudattaa normaalijakaumaa.
4. a) Kun keskiarvon keskivirhe kasvaa, t-suureen arvo pienenee. Kun keskiarvon keskivirhe pienenee, t-suureen arvo kasvaa.
b) Kun otoskoko kasvaa, t-suureen arvo kasvaa. Kun otoskoko pienenee, t-suureen arvo pienenee.
c) Kun otoksen keskihajonta kasvaa, t-suureen arvo pienenee. Kun otoksen keskihajonta pienenee, t-suureen arvo kasvaa.

5. Yhden otoksen t-testin avulla voidaan verrata otoksen havaittua keskiarvoa mielivaltaiseen lukuun. Riippumattomien otosten t-testillä voidaan verrata toisiinsa kahden riippumattoman otoksen keskiarvoja. Toistettujen mittausten t-testillä voidaan verrata samojen tilastoyksiköiden mittaustuloksia toisiinsa kahdella eri muuttujalla tai mittauskerralla.
7. a) $t = 2.5824$
b) $t = 2.8073$
c) $t = 2.6259$
d) $t = 2.6603$
8. a) $t = 4.3168$
b) $t = 2.8779$
c) $t = 3.4966$
d) $t = 2.8385$
9. $-\infty < t < \infty$
10. Lähellä nollaa olevat t-suureen arvot ovat yleisiä, nollasta poikkeavat taas harvinaisia.
11. $t(49) = -2.89$, $p = <.01$, joten otoksen henkilöt ovat kuluttaneet alkoholia keskimääräistä vähemmän.
12. $t(9) = 1.43$, $p > .1$, joten tutkittavat ovat olleet mittaushetkellä samantasoisessa mielialassa kuin ihmiset keskimäärin.
13. $t(9) = 0.732$, $p > .1$, joten tutkittavat ovat älykkyyden suhteen edustava otos suomalaisista.
14. $t(19) = 1.32$, $p > .1$ joten hypoteesi ”suomalaisten peruskoulun alaluokkalaisten itsetunto asteikolla 1–10 on keskimäärin 4.2” saa tukea aineistolta.
15. Molemmissa otoksessa on eri tilastoyksiköt, ja siten keskiarvojen keskivirheidenkään ei voida olettaa olevan samanlaisia eri otoksilla.
16. a) Ei ole, koska $t(18) = -0.12$, $p > .1$
b) Ei ole, koska $t(18) = -0.09$, $p > .1$
17. a) Ei ole, koska $t(18) = -0.46$, $p > .1$
b) Ei ole, koska $t(18) = 0.63$, $p > .1$
c) Kyllä on, koska $t(18) = 3.17$, $p < .01$. Miehet ovat olleet tutkimushetkellä hämmästyneempiä kuin naiset.
18. a) Ei ole, koska $t(18) = -0.19$, $p > .1$
b) Kyllä on, koska $t(18) = 7.53$, $p < .001$. Sokeiden taktiilinen työmuistikapasiteetti on suurempi kuin näkevillä.
19. a) Ei ole, koska $t(18) = .77$, $p > .1$
b) Kyllä on, koska $t(18) = -4.45$, $p < .001$. Lääkehoito on ollut tehokkaampaa, ts. tutkimuksen lopussa lääkehoitoa saanut ryhmä on ollut vähemmän depressiivinen.
20. a) Tutkimusasetelmaa, jossa verrataan muuttujan X arvoja kahden tai useamman riippumattoman otoksen (ryhmän) välillä.
b) Tutkimusasetelmaa, jossa suoritetaan tutkittavien sisäisiä vertailuja siten, että jokaisen tutkittavan suoritusta muuttujalla (tai mittauskerralla) A verrataan hänen omaan suoritukseensa muuttujalla (tai mittauskerralla) B.

21. Tällaisessa asetelmassa ei tarvita kahta erillistä otosta, koska tutkimus perustuu koehenkilöiden sisäiseen vertailuun. Tällöin esimerkiksi koehenkilöiden jakaminen koe- ja kontrolliryhmiin ei voi vaikuttaa tutkimuksen tuloksiin.
22. Reaktioajat ovat olleet nopeampia kongruentissa tilanteessa, $t(14) = -9.90$, $p < .001$.
23. Depressio on vähentynyt tutkimuksen kuluessa, koska $t(19) = 6.90$, $p < .001$.

24. a) Reaktioajat ovat olleet nopeampia kongruenteissa tilanteissa, koska $t(9) = -6.04$, $p < .001$.
- b) Reaktioajat ovat olleet nopeampia neutraaleissa tilanteissa, koska $t(9) = -4.50$, $p < .005$.
- c) Reaktioajat ovat olleet nopeampia kongruenteissa tilanteissa, koska $t(9) = -3.58$, $p < .01$.



12 Yksisuuntainen varianssianalyysi

1. Yli kahden keskiarvon välistä erotusta ei voi määritellä yksikäsitteisesti.
2. Siinä missä t-testi vertaa kahden keskiarvon erotusta keskiarvon keskivirheeseen, varianssianalyysi vertaa käsittelytekijän aiheuttamaa varianssia virhevarianssiin.
3. a) Käsittelytekijä, yksilölliset erot, mittausvirhe
b) Yksilölliset erot, mittausvirhe
4. F-suhde määrittää riippumattoman muuttujan aiheuttaman varianssin suhteena virhevarianssiin. Jos $F < 1$, mittausvirhe on vaikuttanut tuloksiin enemmän kuin käsittelytekijä. Jos $F = 1$, käsittelytekijä ja mittausvirhe ovat vaikuttaneet tuloksiin yhtä paljon. Jos $F > 1$, käsittelytekijä on vaikuttanut mittaus tuloksiin enemmän kuin mittausvirhe.
5. a) $\sigma^2 = 20$
b) $\sigma^2 = 40$
c) $\sigma^2 = 100$

6. a) $F = 3.94$
 b) $F = 2.79$
 c) $F = 2.12$
 d) $F = 3.49$
 e) $F = 6.61$
7. $0 \leq F \leq \infty$
8. Lähellä ykköstä olevat ovat yleisiä, ykköstä suuremmat taas harvinaisia.
9. a) $t(18) = 7.53, p < .001$. Sokeiden taktiilinen työmuistikapasiteetti on suurempi kuin näkevillä.
 b) $F(1,80) = 56.65, p < .001$. Sokeiden taktiilinen työmuistikapasiteetti on suurempi kuin näkevillä. A- ja b-kohtien tulokset ovat samansuuntaisia.
10. a) $p = .09$
 b) $p = .03$
 c) $p = .03$
 Bonferroni-korjausten jälkeen ainoastaan neljäs- ja kuudesluokkalaiset sekä viides- ja kuudesluokkalaiset poikkeavat itsetunnon suhteen toisistaan.
11. a) $F(2,27) = 3.50, p < .05$. F-suhde on tilastollisesti merkitsevä.
 b) $F(2,27) = 1.17, p > .05$. F-suhde ei ole tilastollisesti merkitsevä.
 c) $F(2,27) = 1.17, p > .05$. F-suhde ei ole tilastollisesti merkitsevä.
12. a) $SS_{\text{total}} = 1688.17, df_{\text{total}} = 29$
 b) $SS_{\text{BG}} = 1329.27, df_{\text{BG}} = 2$
 c) $SS_{\text{WG}} = 358.90, df_{\text{WG}} = 27$
 d) $F(2,27) = 50.00, p < .01$. Työmuistikapasiteetti vaikuttaa siihen, kuinka paljon tekstistä muistetaan.
13. a) $SS_{\text{total}} = 1599.20, df_{\text{total}} = 29$
 b) $SS_{\text{BG}} = 1360.80, df_{\text{BG}} = 2$
 c) $SS_{\text{WG}} = 238.40, df_{\text{WG}} = 27$
 d) $F(2,27) = 77.06, p < .01$. Työmuistikapasiteetti vaikuttaa siihen, kuinka kauan koehenkilöt lukevat tekstiä.
14. Verrataan kaikkia työmuistikapasiteettiryhmiä toisiinsa. Bonferroni-korjausten kanssa saadaan tulokseksi
 a) matala vs. keskinkertainen $t(18) = 3.52, p < .015$
 b) matala vs. korkea $t(18) = 13.08, p < .003$
 c) keskinkertainen vs. korkea $t(18) = 9.10, p < .003$
15. $F(2,27) = 1.091, p > .05$. Koulutus ei ole yhteydessä työn kuormittavuuden kokemiseen, joten monivertailuja ei tarvitse suorittaa.
16. $F(2,27) = 27.41, p < .01$. Prosessointikonteksti vaikuttaa tekstin ymmärtämiseen. Verrataan kaikkia kokeellisia ryhmiä toisiinsa ja suoritetaan Bonferroni-korjaukset. Saadaan
 a) konteksti ennen vs. ei kontekstia $t(18) = 6.01, p < .003$
 b) konteksti ennen vs. konteksti jälkeen $t(18) = 7.69, p < .003$
 c) ei kontekstia vs. konteksti jälkeen $t(18) = -.87, p > .05$.
 Kontekstin antaminen ennen tekstiä parantaa siis tekstin ymmärtämistä verrattuna siihen, että kontekstia ei anneta ollenkaan tai että se annetaan tekstin lukemisen jälkeen.

17. $F(2,27) = 20.26$, $p < .01$.

Prosessointikonteksti vaikuttaa tekstin kokonaispalautukseen.

Verrataan kaikkia kokeellisia ryhmiä toisiinsa ja suoritetaan Bonferroni-korjaukset. Saadaan

- a) konteksti ennen vs. ei kontekstia $t(18) = 5.71$, $p < .003$
- b) konteksti ennen vs. konteksti jälkeen $t(18) = 5.24$, $p < .003$
- c) ei kontekstia vs. konteksti jälkeen $t(18) = -1.10$, $p > .05$.

Kontekstin antaminen ennen tekstiä parantaa siis tekstin muistamista verrattuna siihen, että kontekstia ei anneta ollenkaan tai että se annetaan tekstin lukemisen jälkeen.

13 Useampisuuntainen varianssianalyysi

1. Useampisuuntaisessa varianssianalyysissä voidaan tarkastella samanaikaisesti kahden tai useamman lohkoketekijän vaikutuksia riippuvaan muuttujaan.
2. Välimatka-asteikollinen mittaus, normaalijakaumaoletus, varianssien homogeneisuusoletus sekä yli 20 havaintoa jokaisessa tarkasteltavassa ryhmässä
3. $df_{total} = N - 1$
 $df_{WG} = N - a \times b$
 $df_{BG_a} = a - 1$
 $df_{BG_b} = b - 1$
 $df_{axb} = (a - 1) \times (b - 1)$
4. Päävaikutus tarkoittaa sellaista riippumattoman muuttujan vaihtelua, joka selittyy pelkästään yhden lohkoketekijän avulla. Yhdysvaikutus puolestaan tarkoittaa sellaista vaihtelua, joka ei selity pelkästään yhden, vaan useamman lohkoketekijän avulla.
5. a) Tekijän A vaikutuksia riippuvaan muuttujaan
 b) Tekijän B vaikutuksia riippuvaan muuttujaan
 c) Tekijöiden A ja B yhteistä vaikutusta riippuvaan muuttujaan
 d) Yksilöllisistä eroista aiheutuvaa variaatiota

6. a) Kaksi kaksitasoista tekijää
 b) Kolme tekijää, joista kaksi on kolme- ja yksi nelitasoinen
 c) Kolme tekijää, joista yksi on kaksitasoinen ja kaksi kolmetasoista
 d) Neljä tekijää, joista kaksi on neli- ja kaksi kolmetasoista

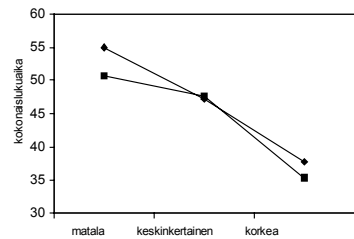
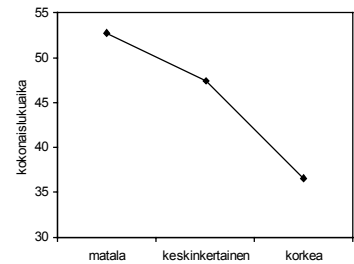
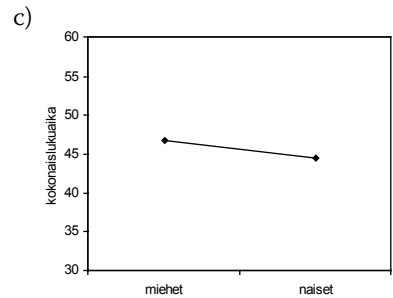
7. a) tupakoivat miehet
 b) tupakoimattomat miehet
 c) tupakoivat naiset
 d) tupakoimattomat naiset

8. a) normaali serif
 b) normaali sans serif
 c) **normaali fraktuura**
 d) pieni serif
 e) pieni sans serif
 f) **pieni fraktuura**

9. a) pojat, joilla on 0 sisarusta
 b) tytöt, joilla on 0 sisarusta
 c) pojat, joilla on 1 sisarus
 d) tytöt, joilla on 1 sisarus
 e) pojat, joilla on 2 sisarusta
 f) tytöt, joilla on 2 sisarusta

10. Jos varianssianalyysissä havaitaan yhdysvaikutus, päävaikutuksia ei yleensä analysoida koska ne aiheutuvat tyypillisesti yhdysvaikutuksesta.

11. a) $\bar{x}_{matala} = 52.80$, $\bar{x}_{keskinkertainen} = 47.40$, $\bar{x}_{korkea} = 36.60$
 b) $\bar{x}_{miehet} = 46.67$, $\bar{x}_{naiset} = 44.53$
 $\bar{x}_{miehet, matala} = 55.00$
 $\bar{x}_{miehet, keskinkertainen} = 47.20$
 $\bar{x}_{miehet, korkea} = 37.80$
 $\bar{x}_{naiset, matala} = 56.60$
 $\bar{x}_{naiset, keskinkertainen} = 47.60$
 $\bar{x}_{naiset, korkea} = 35.40$



—♦— mies —■— nainen

Työmuistikapasiteetilla näyttäisi olevan päävaikutus, mutta sukupuolen päävaikutusta sekä sukupuolen ja työmuistikapasiteetin yhdysvaikutusta ei todennäköisesti havaita.

12. a) $\bar{x}_{matala} = 74.90$, $\bar{x}_{keskinkertainen} = 83.40$, $\bar{x}_{korkea} = 91.20$

b) $\bar{x}_{miehet} = 83.67$, $\bar{x}_{naiset} = 82.67$

$\bar{x}_{miehet, matala} = 75.20$

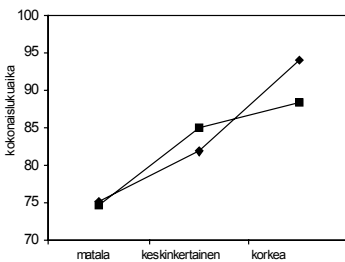
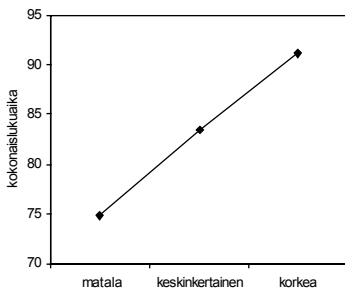
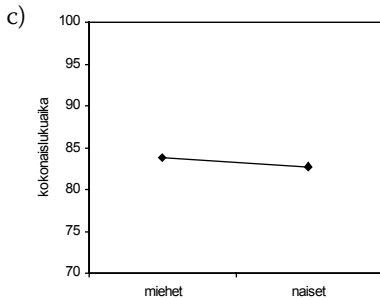
$\bar{x}_{miehet, keskinkertainen} = 81.80$

$\bar{x}_{miehet, korkea} = 94.00$

$\bar{x}_{naiset, matala} = 74.60$

$\bar{x}_{naiset, keskinkertainen} = 85.00$

$\bar{x}_{naiset, korkea} = 88.40$



—♦— mies —■— nainen

Työmuistikapasiteetilla näyttäisi olevan päävaikutus, mutta sukupuolen päävaikutusta tai sukupuolen ja työmuistikapasiteetin yhdysvaikutusta ei todennäköisesti havaita.

13. a) $SS_{total} = 1599.20$, $df_{total} = 29$

b) $SS_A = 1360.80$, $df_A = 2$

c) $SS_B = 34.13$, $df_B = 2$

d) $SS_{A \times B} = 29.07$, $df_{A \times B} = 2$

14. a) $F(2,24) = 93.21$, $p < .01$. Suuri työmuistikapasiteetti näyttäisi lyhentävän tekstin luku-aikoja.

b) $F(1,24) = 4.68$, $p < .05$. Naiset ovat lukeneet tekstin miehiä nopeammin.

c) $F(2,24) = 1.99$, $p > .05$. Työmuistikapasiteetti ja sukupuoli eivät yhdessä vaikuta tekstin kokonaislukuaikaan, vaan niillä molemmilla on erilliset vaikutukset.

15. a) $SS_{total} = 1688.17$, $df_{total} = 29$

b) $SS_A = 1329.27$, $df_A = 2$

c) $SS_B = 7.50$, $df_B = 1$

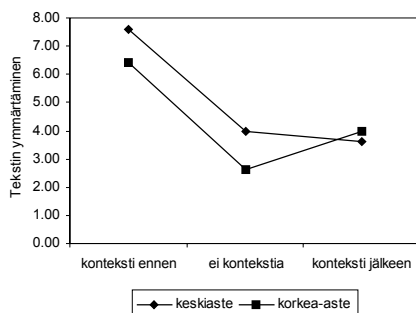
d) $SS_{A \times B} = 97.40$, $df_{A \times B} = 2$

16. a) $F(2,24) = 62.80$, $p < .05$. Korkea työmuistikapasiteetti parantaa kokonaispalautusta.

b) $F(1,24) = 0.71$, $p > .05$. Sukupuoli ei vaikuta tekstin kokonaispalautukseen.

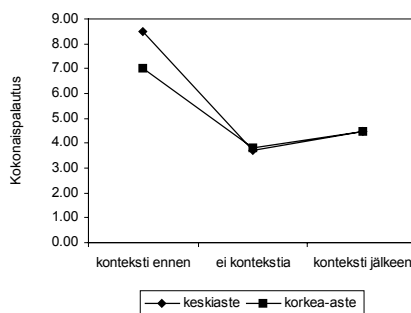
c) $F(2,24) = 4.60$, $p < .05$. Työmuistin kapasiteetti ja sukupuoli näyttäisivät vaikuttavan yhdessä kokonaispalautukseen.

17.



Prosessointikonteksti vaikuttaa tekstin ymmärtämiseen, $F(2,24) = 31.40$, $p < .05$, mutta koulutuksen päävaikutus, $F(1,24) = 3.14$, $p > .05$, ja prosessointikontekstin ja koulutuksen yhdysvaikutus, $F(2,24) = 1.90$, $p > .05$ eivät ole tilastollisesti merkitseviä.

18.



Prosessointikonteksti vaikuttaa tekstin kokonaispalautukseen, $F(2,24) = 19.87$, $p < .05$, mutta koulutuksen päävaikutus, $F(1,24) = .72$, $p > .05$, ja prosessointikontekstin ja koulutuksen yhdysvaikutus, $F(2,24) = .88$, $p > .05$, eivät ole tilastollisesti merkitseviä.

14 Toistettujen mittausten varianssianalyysi

- Koska samat tutkittavat on mitattu kaikilla käsittelytekijän tasolla, yksilölliset erot ja otantavirhe eivät vaikuta tuloksiin läheskään niin paljon kuin riippumattomien otosten varianssianalyysissa.
- $$df_{\text{total}} = N - 1$$

$$df_{\text{within-treatments}} = df_{\text{total}} - df_{\text{between-treatments}}$$

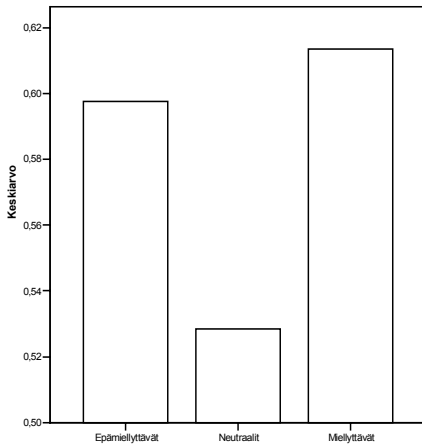
$$df_{\text{between-treatments}} = k - 1$$

$$df_{\text{between-subjects}} = n - 1$$

$$df_{\text{error}} = (n - 1)(k - 1)$$
- Populaatiovarianssin estimaatti koko aineiston varianssin perusteella laskettuna
 - Populaatiovarianssin estimaatti eri käsittelykertojen sisällä havaittavan vaihtelun perusteella laskettuna
 - Populaatiovarianssin estimaatti eri käsittelykertojen keskiarvojen varianssin perusteella laskettuna
 - Populaatiovarianssin estimaatti laskettuna mittausvirheen varianssin perusteella

5. a) $F(2, 58) = 2, p > .05$
 b) $F(2, 58) = 5, p < .05$
 c) $F(2, 58) = 1.25, p > .05$
6. a) $F(2, 27) = 1.2, p > .05$
 b) $F(2, 27) = 3, p > .05$
 c) $F(2, 27) = 1.13, p > .05$
7. Sanan ja värin kongruenssi vaikuttaa prosessointinopeuteen, koska $F(1, 14) = 97.97, p < .01$.

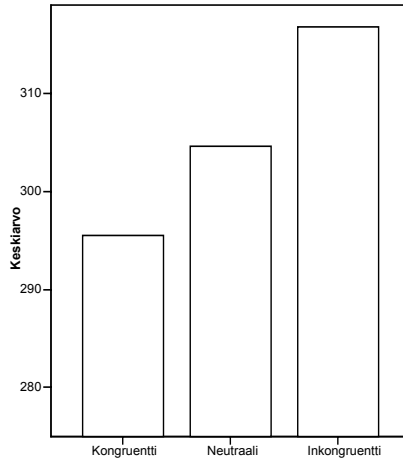
8. $\bar{x}_{\text{epämiellyttävät}} = .60$
 $\bar{x}_{\text{neutraalit}} = .53$
 $\bar{x}_{\text{miellyttävät}} = .61$



Tarkkaavaisuus näyttäisi suuntautuvan sekä epämiellyttäviin että miellyttäviin kuviin todennäköisemmin kuin neutraaleihin kuviin.

9. Kuvan emotionaalinen sisältö näyttäisi vaikuttavan tarkkaavaisuuden suuntautumiseen, koska $F(2, 38) = 12.93, p < .05$.

10. $\bar{x}_{\text{kongruentti}} = 295.50$
 $\bar{x}_{\text{neutraalit}} = 304.60$
 $\bar{x}_{\text{inkongruentti}} = 316.80$



Reaktioajat näyttäisivät olevan nopeimpia kongruentissa ja hitaimpia inkongruentissa tilanteissa. Neutraalin tilanteen reaktioajat jäävät näiden kahden tilanteen väliin.

11. Katsevihjeet näyttäisivät suuntaavan tarkkaavaisuutta refleksiivisesti, sillä $F(2, 18) = 26.13, p > .05$. Reaktioajat ovat kongruentissa tilanteessa nopeampia kuin neutraalissa, $t(9) = -3.58, p < .01$, ja inkongruentissa, $t(9) = -6.04, p < .001$. Reaktioajat neutraalissa tilanteessa ovat myös nopeampia kuin inkongruentissa tilanteessa, $t(9) = -4.50, p < .001$.

12. Kasvojen kääntäminen hankaloittaa kasvojen tunnistamista, koska $F(2, 18) = 329.47$, $p < .01$. Kasvojen tunnistaminen 0 asteen tilanteessa on tarkempaa kuin 90 asteen, $t(9) = 8.22$, $p < .001$, ja 180 asteen, $t(9) = 23.58$, $p < .001$, tilanteessa. Lisäksi kasvojen tunnistaminen on tarkempaa 90 asteen kuin 180 asteen tilanteessa, $t(9) = 15.46$, $p < .001$.

15 Epäparametriset menetelmät

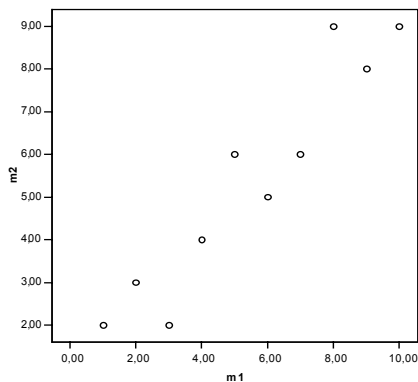
1. Jos mittaukset on suoritettu järjestysasteikolla tai jos parametristen testien oletukset (tyypillisesti normaalijakaumaoletus) eivät toteudu.
2. Jos verrataan sellaisia jakaumia, jotka 1) eivät ole normaalisti jakautuneita, 2) joissa molemmissa on suunnilleen sama moodi (ja jakauman painopiste) ja 3) jotka ovat eri suuntaan vinoja siten, että jakaumien vinouden ansioista jakaumien keskiarvot ovat kaukana toisistaan ja jakauman painopisteestä.
3. Jos verrataan sellaisia jakaumia, jotka 1) eivät ole normaalisti jakautuneita, 2) joiden moodit ovat erisuuruiset ja 3) ovat eri suuntaan vinoja siten, että jakaumien vinouden ansioista jakaumien keskiarvot ovat kaukana toisistaan ja jakauman painopisteestä. Lisäksi riski tyyppin 2 virheen tekemiseen kasvaa, jos epäparametrisia menetelmiä käytetään parametristen sijaan silloin, kun parametristen testien oletukset ovat voimassa, koska epäparametriset testit eivät ole yhtä voimakkaita kuin niiden parametriset vastineet.
4. Eivät mitään. Riittää, että muuttujat on mitattu järjestysasteikolla.
5. Koska 1) epäparametrisia testejä käytetään yleensä, kun normaali-jakaumaoletus ei ole voimassa, ja tällöin keskiarvo ei ole hyvä jakauman sijaintia kuvaava tunnusluku ja 2) epäparametriset testit perustuvat havaintoaineiston muuttamiseen järjestysasteikolle (tai niiden avulla voidaan analysoida valmiiksi järjestysasteikollista dataa), ja tällä mitta-asteikolla keskiarvoa ei ole määritetty.
6. Koska epäparametriset testit eivät ole yhtä voimakkaita kuin niiden parametriset vastineet, ja riski tyyppin 2 virheen tekemiseen kasvaa.
7. a) Mann–Whitneyn U-testi
b) Wilcoxonin merkittyjen järjestyslukujen testi
c) Kruskal–Wallis-testi
d) Friedmanin testi

8. a) 7.5, 3, 9.5, 15, 7.5, 12.5, 1, 11, 12.5, 14, 3, 5, 6, 3, 9.5
b) 14, 9, 15, 13, 10, 11, 4.5, 8, 12, 2, 3, 4.5, 6, 1, 7
9. a) Eivät, sillä $U = 39.50$, $p > .01$
b) Kyllä, sillä $U = 7.00$, $p < .01$
10. a) Kyllä, sillä $U = 1$, $p < .01$
b) Eivät, sillä $U = 47.50$, $p > .01$
11. a) Eivät, sillä $U = 88.50$, $p > .01$
b) Eivät, sillä $U = 106.00$, $p > .01$
12. a) Kyllä, sillä $T = 0$, $p < .05$
b) Kyllä, sillä $T = 0$, $p < .05$
13. Kyllä, sillä $T = 2.5$, $p < .05$.
14. Kyllä, sillä $T = 0$, $p < .05$
15. a) Kyllä, sillä $\chi^2(2) = 23.15$, $p < .001$
b) Kyllä, sillä $\chi^2(2) = 22.36$, $p < .001$
16. a) Ei, sillä $\chi^2(2) = 3.15$, $p < .05$
b) Ei, sillä $\chi^2(2) = 1.19$, $p < .05$
17. a) Prosessointikonteksti vaikuttaa tekstin ymmärtämiseen, $\chi^2(2) = 18.56$, $p < .05$. Konteksti ennen -ryhmä poikkeaa sekä ei kontekstia, $U = 4$, $p < .01$, että konteksti jälkeen, $U = 1$, $p < .01$ -ryhmistä. Ei kontekstia ja konteksti jälkeen -ryhmät eivät eronneet toisistaan, $U = 34$, $p > .05$.
b) Prosessointikonteksti vaikuttaa tekstin kokonaispalautukseen, $\chi^2(2) = 17.68$, $p < .05$. Konteksti ennen -ryhmä poikkeaa sekä ei kontekstia, $U = 3.50$, $p < .01$, että konteksti jälkeen, $U = 3.00$, $p < .01$ -ryhmistä. Ei kontekstia ja konteksti jälkeen -ryhmät eivät eronneet toisistaan, $U = 35$, $p > .05$.
18. Kasvoinversio vaikuttaa kasvojen tunnistamiseen, sillä $\chi^2(2) = 20.00$, $p < .001$.
19. Katsevihjeet suuntaavat tarkkaavaisuutta refleksiivisesti, sillä $\chi^2(2) = 17.59$, $p < .001$.
20. Emotionaalinen sisältö vaikuttaa tarkkaavaisuuden suuntautumiseen, sillä $\chi^2(2) = 17.44$, $p < .001$. Ensimmäinen fiksaatio osuu neutraalia kuvaa todennäköisemmin sekä epämiellyttäviin, $T = 8$, $p < .05$, että miellyttäviin, $T = 3$, $p < .05$, kuviin.

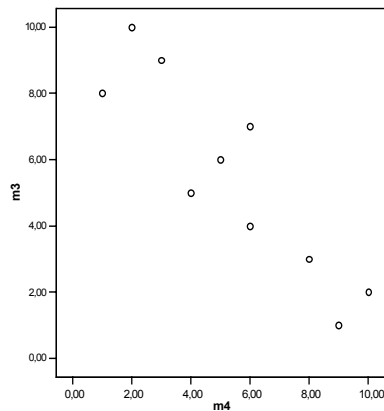
16 Kovarianssi ja korrelaatio

1. Yhteisvaihtelu voi tarkoittaa 1) kahden ilmiön esiintymistä samanaikaisesti, 2) kahden ilmiön liittymistä toisiinsa tai 3) kahden ilmiön välistä syy-seuraussuhdetta.
2. Jos 1) muuttujien x ja y välillä havaitaan yhteisvaihtelua, 2) syy edeltää seurausta ajallisesti ja 3) mikään kolmas tutkittavia muuttujia ajallisesti edeltävä muuttuja ei aiheuta tutkittavien muuttujien yhteisvaihtelua.

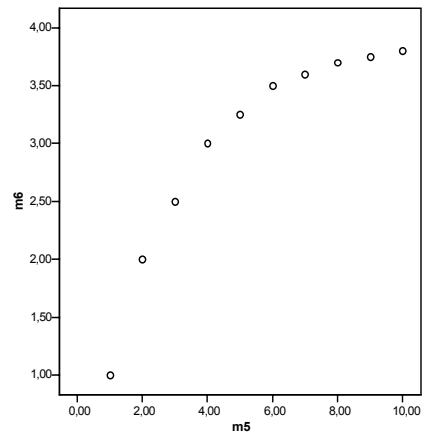
3. a)



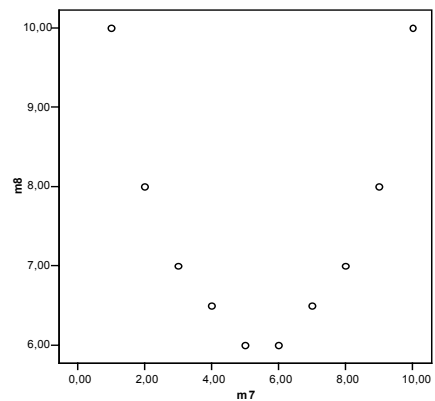
b)



c)



d)



4) $-\infty < s_{xy} < \infty$

5. a) Kahden muuttujan välillä havaittu korrelaatio ei sinällään kerro mitään muuttujien välisestä kausaalisuhteesta. Korrelaatio voi kuitenkin tarkoittaa muuttujien välistä kausaalisuhdetta, jos se havaitaan esimerkiksi kokeellisessa asetelmassa riippuvan ja riippumattoman muuttujan välillä. Käytettyä tutkimusasetelmaa tarkastelemalla voidaankin päätellä, tarkoittaako kahden muuttujan välinen korrelaatio kausaalista yhteyttä vai ei.
- b) Suurissa otoksissa tulkinnan kannalta merkityksettömän pienetkin korrelaatiot saattavat olla tilastollisesti merkitseviä, koska korrelaatiokertoimen tilastollinen testaaminen testaa ainoastaan sitä, poikkeako korrelaatiokerroin nolasta. Täten korrelaatiokertoimesta on aina tarkasteltava sen tilastollisen merkitsevyyden lisäksi myös itse kertoimen suuruutta.
6. Kovarianssi ei ole standardoitu $[-\infty < s_{xy} < \infty]$, joten kahden muuttujaparin väliset kovarianssikertoimet eivät ole keskenään vertailukelpoisia.
7. Pearsonin tulomomenttikorrelaatiokerroin on parametrinen korrelaatiokerroin, joka voidaan laskea kahden normaalisti jakautuneen muuttujan välille. Se estimoii kahden muuttujan välistä lineaarista yhteisvaihtelua. Spearmanin järjestyskorrelaatiokerroin puolestaan on epäparametrinen korrelaatiokerroin, joka voidaan laskea kahden järjestys-, välimatka- tai suhdeasteikollisen muuttujan välille. Se estimoii kahden muuttujan välistä monotonista yhteyttä, joka voi olla myös lineaarista.
8. Jos epäillään, että muuttujien välinen yhteys ei olekaan lineaarista vaan monotonista.
9. a) Poikkeavan havainnon etäisyyttä siitä sirontakuvion pisteparvesta, jonka tyypillisen suuruiset havainnot muodostavat.
- b) Poikkeavan havainnon epäjohdonmukaista sijaintia suhteessa siihen sirontakuvion pisteparveen, jonka tyypillisen suuruiset havainnot muodostavat.
10. Sekä havainnon vipuvoima että diskrepanssi vaikuttavat korrelaatiokertoimen suuruuteen, ja suurin vaikutus on sellaisilla havainnoilla, joiden vipuvoima ja diskrepanssi ovat suuria. Suuri vipuvoima ei kuitenkaan välttämättä vaikuta korrelaatiokertoimen suuruuteen kovin paljon, jos havainnon diskrepanssi on pieni.
12. Koska melko suuressakin havaintoaineistossa voidaan havaita suuria korrelaatiokertoimia sattumalta.
13. Osittaiskorrelaatiokerroin estimoii kahden muuttujan todellisen yhteisvaihtelun määrän, josta on poistettu kolmannen muuttujan vaikutus. Multippelikorrelaatiossa puolestaan lasketaan kahden muuttujan yhteinen vaihtelu jonkin kolmannen muuttujan kanssa.
14. Jos muuttujilla ei ole juuri ollenkaan varianssia, niillä ei voi myöskään olla paljoa kovarianssia. Koska korrelaatiokerroin on muuttujien välinen standardoitu kovarianssikerroin, ei tällaisessa tilanteessa havaita muuttujien välistä korrelaatiota.

17. a) Fonologisen prosessoinnin ja kuullun ymmärtämisen välillä ei ole yhteyttä.
b) Fonologisen prosessoinnin ja kuullun ymmärtämisen välillä ei ole yhteyttä.
c) Fonologisen prosessoinnin ja kuullun ymmärtämisen välillä on yhteys.
18. a) $r = .22$ ($p < .05$), $n \geq 80$
b) $r = .33$ ($p < .01$), $n \geq 50$
c) $r = .12$ ($p < .001$), $n \geq 750$
19. a) $r = .90$, $p > .1$
b) $r = -.54$, $p < .01$
c) $r = .51$, $p < .01$
20. a) $s_{xy} = -40.87$
b) $r = -.75$
c) $r_s = -.69$
21. a) $r = -.13$, $p > .1$
b) $r = -.18$, $p > .1$
c) $r = -.33$, $p > .1$
22. a) $r_s = -.17$, $p > .1$
b) $r_s = -.93$, $p < .01$
c) $r_s = .01$, $p > .1$
23. a) $r_{z,xy} = .28$
b) $r_{z,xy} = .92$
24. a) $r_{xy,z} = .50$
b) $r_{xy,z} = .05$

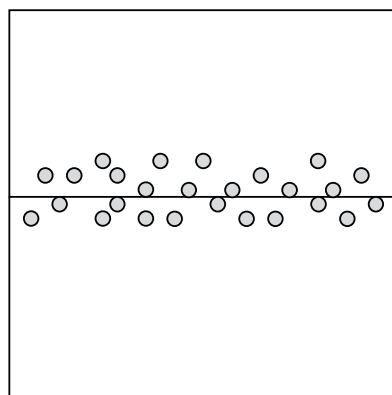
17 Kategoristen muuttujien välisten yhteyksien tarkasteleminen

1. χ^2 -yhteensopivuustestillä voidaan testata, poikkeako havaittu frekvenssijakauma jostain mieltävaltaisesta frekvenssijakaumasta. χ^2 -riippumattomuustestillä voidaan puolestaan testata, onko jokin havaittu frekvenssijakauma riippuvainen jonkin toisen kategorisen muuttujan jakaumasta.
2. χ^2 -jakauma on epäsymmetrinen, oikealle vino jakauma, joka mallintaa variansseja. χ^2 -jakauman muoto riippuu vapausasteista (df).
3. χ^2 -jakauman odotusarvo on $df + 1$.
5. a) Ei, koska kriittinen raja on 5.9915
b) Kyllä, koska kriittinen raja on 5.9915
c) Ei, koska kriittinen raja on 18.3070
d) Kyllä, koska kriittinen raja on 9.4877
6. Vastaa, sillä $\chi^2(1) = 1.44$, $p > .05$
7. Ei, sillä $\chi^2(1) = 16.16$, $p < .001$. Otoksessa on enemmän skitsofreenikkoja kuin populaation perusteella voisi olettaa.
8. Kyllä poikkeaa, sillä $\chi^2(1) = 5$, $p < .05$. Otoksessa näyttäisi olevan suhteessa enemmän vasenkätisiä kuin populaatiossa.
9. Aineiston perusteella depressio on yhtä yleistä sekä miehillä että naisilla, sillä $\chi^2(1) = 1.80$, $p > .05$.
10. Tupakointi ja ylipainoisuus näyttäsivät olevan yhteydessä toisiinsa, sillä $\chi^2(1) = 5.04$, $p < .05$. Ylipaino on suhteellisesti harvinaisempaa tupakoivien ryhmässä.
11. PTSD-oireiden esiintyvyydessä on eroa ryhmien välillä, koska $\chi^2(1) = 5.09$, $p < .05$. Kontrolliryhmässä oireilua ilmenee suhteellisesti vähemmän kuin debriefing-ryhmässä, joten tämän tutkimuksen perusteella debriefing näyttäisi olevan haitallista.
12. Menetelmät eivät aineiston perusteella ole yhtä tehokkaita, sillä $\chi^2(1) = 16.94$, $p < .001$. Psykologit tekevät suhteessa enemmän oikeita diagnooseja BDI-II:n perusteella.
13. Amygdalaleesio näyttäisi vaikuttavan pelon ehdollistumiseen, sillä $\chi^2(2) = 24.38$, $p < .001$. Kaikkein vähiten ehdollistumattomia rottia on bilateraalisen leesio-ryhmässä ja eniten ehdollistuneita rottia kontrolliryhmässä.
14. Sekä miehet että naiset ovat sairastuneet yhtä usein työuupumukseen, sillä $\chi^2(1) = .60$, $p > .05$.
15. Koulutuksen jakauma on samanlainen sekä miehillä että naisilla, sillä $\chi^2(1) = .56$, $p > .05$.

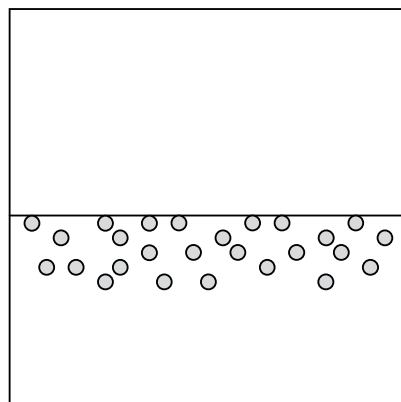
18 Lineaarinen regressio

1. Vähintään 50, mielellään yli 100 havaintoa, selittäjät ja selitettävä muuttuja normaalisti jakautuneita, selittäjät eivät ole keskenään multikolineaarisia
2. a) Kohta, jossa regressiosuora leikkaa y-akselin. Vakiotermi ilmoittaa, mikä on selitettävän muuttujan arvo, jos kaikkien selittävien muuttujien arvot ovat tasan 0.
b) Yhden selittäjän regressiomallissa regressiosuoran kulma-kerroin. Regressiokerroin ilmoittaa, kuinka paljon selitettävän muuttujan arvot muuttuvat, kun selittävän muuttujan arvo kasvaa yhden yksikön.
c) Standardoitu regressiokerroin saadaan, jos sekä selitettävä muuttuja että selittävät muuttujat standardoidaan (suoritetaan z-muunnos). Täten eri regressiomallien standardoidut regressiokertoimet ovat keskenään vertailukelpoisia.
d) Selitettävän muuttujan todellisen arvon ja regressiomallin ennustaman arvon välinen erotus. Jäännöstermi kuvaa sitä, kuinka hyvin yksittäisen selitettävän muuttujan arvon ennustaminen onnistuu regressiomallin avulla.
3. a) Vertaamalla F-testillä mallin selittämää vaihtelua selittämättä jäämään vaihteluun.
b) Laskemalla selittävien muuttujien neliöity multippelikorrelaatio selitettävän muuttujan kanssa.
4. Koska tällöin muuten selittäjää hyvin ennustavista muuttujista kaikki eivät "mahdu" mukaan malliin, sillä ne selittävät samaa vaihtelua selitettävästä muuttujasta.

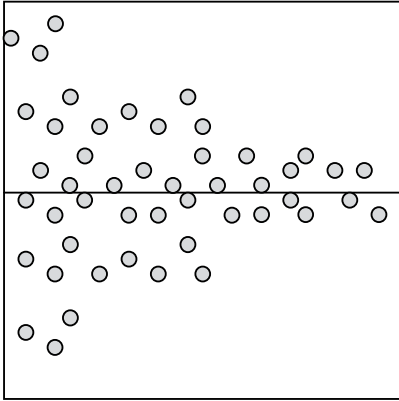
5. a)



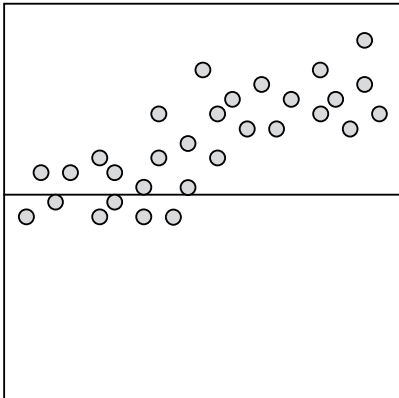
b)



c)



d)



6. a) $\hat{y} = 9$
 b) $\hat{y} = 7$
 c) $\hat{y} = 10$
 d) $\hat{y} = 12$

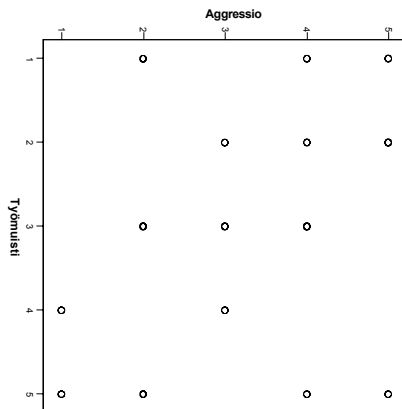
7. Malli ei selvästikään voi toimia ainakaan silloin kun tutkitun lapsen älykkyydosamäärä ylittää arvon 120. Tätä suuremmilla älykkyydosamäärän arvoilla malli ennustaisi lapsen koulumenestykseksi matematiikassa yli 10, mikä on mahdollonta koska koulumenestys matematiikassa vaihtelee välillä 4–10.

8. Käyttökokemus kasvattaa opintojen koettua mielekkyyttä siten, että yhden pisteen lisäys käyttökokeuksessa lisää koettua mielekkyyttä puolella pisteellä. Ikä puolestaan vähentää koettua mielekkyyttä siten, että yhden vuoden lisäys iässä vähentää koettua mielekkyyttä pisteen kymmenesosalla.

9. a) $\hat{y} = 7$
 b) $\hat{y} = 3$
 c) $\hat{y} = 0$
 d) $\hat{y} = 3.7$
10. a) $\hat{y} = 40\%$
 b) $\hat{y} = 30\%$
 c) $\hat{y} = 20\%$
 d) $\hat{y} = 0\%$

11. $\hat{y} = -10x + 70$

12. a) $r = -.33$
 b)

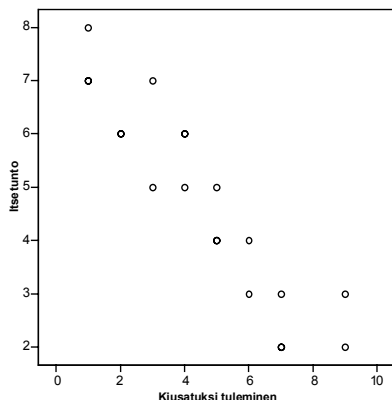


- c) $\beta_1 = -.32, \beta_0 = 4.27$
 d) $r^2 = 0.11$

- e) Työmuistin kapasiteetti selittää fyysisestä aggressiosta 11 %. Jos työmuistin kapasiteetti olisi 0 (kuvitteellinen tilanne), fyysisen aggression määrä olisi 4.27. Yhden yksikön lisäys työmuistikapasiteetissa vähentää aggressiivisuutta .32 yksikköä.

13. a) $r = -.92, p < .001$

b)



c) $\beta_1 = -.68, \beta_0 = 7.86$

d) $r^2 = .84$

- e) Kiusatuksi tuleminen kokemus selittää tutkittavien itsetunnosta 84 %. Jos oppilasta ei kiusata ollenkaan, hänen itsetuntonsa on mallin mukaan 7.86. Kiusaamisen kokemisen kasvaminen yhdellä yksiköllä vähentää itsetuntoa aina .68 yksikön verran.

14. a) $\beta_1 = .42, \beta_0 = 1.12, r^2 = .51$

b) $\beta_1 = .04, \beta_0 = .95, r^2 = .57$

15. a) $\beta_1 = .42, \beta_0 = 1.12, r^2 = .51$

b) $\beta_1 = .04, \beta_0 = .95, r^2 = .57$

19 Logistinen regressio

- Lineaarisessa regressioanalyysissä ennustetaan, miten jatkuvan selitetävän muuttujan vaihtelu voidaan ennustaa selittävien muuttujien avulla. Logistisessa regressioanalyysissä ennustetaan, miten kategorisen selitetävän muuttujan havaintojen jakautuminen luokkiin voidaan ennustaa selittävien muuttujien avulla.
- Binäärisissä malleissa selitetään kaksi-, multinomisissa useampi-luokkaisten riippuvien muuttujien havaintojen jakautumista luokkiin eri selittävien muuttujien avulla.
- Jos logitia ei käytettäisi, mallit eivät olisi symmetrisiä jos mallinnetaan todennäköisyyksiä P ja $1-P$. Logitia ennustamalla tämä ongelma voidaan ratkaista, ja tämän jälkeen regressioyhtälöstä voidaan ratkaista referenssiakategoriaan kuulumisen todennäköisyys P ja mallintaa sitä.
- $$OR = \frac{P}{1-P}, \logit(P) = \ln\left(\frac{P}{1-P}\right)$$
- $OR = 0.67, \logit(A) = -0.41$
 - $OR(A) = 4, \logit(A) = 1.39$
 - $OR(B) = 0.43, \logit(B) = -0.85$

6. a) $OR(A) = 0.01$, $\logit(A) = -4.60$
 b) $OR(B) = 0.99$ $\logit(B) = 4.60$
7. Selitettävän muuttujan tulee olla kategorinen.
8. Silloin, jos jakaumaoletukset ja selittäjien kolinearisuusoletus eivät ole voimassa, voidaan selitettävä muuttuja muuntaa järjestyks-asteikolle ja käyttää logistista regressioanalyysia.
9. Vertaamalla lopullisen mallin tekemää ennustetta pelkän vakio-termimallin avulla saatavaan ennusteeseen χ^2 -testin avulla voidaan arvioida, ovatko selittävät muuttujat ylipäättään merkityksellisiä. Jos taas verrataan mallin ennustamaa selitettävän muuttujan jakauma havaittuun jakaumaan χ^2 -testin avulla, voidaan arvioida, kykeneekö malli tuottamaan havaitun kaltaisen datan riittävän tarkasti.
10. Tarkastelemalla, kuinka hyvin malli luokittelee 1) kaikki havainnot ja 2) eri luokkiin kuuluvat havainnot.
11. a) Mallin luokittelutarkkuus on koko aineistolle (94 %) ja sairastumattomille (89 %) korkea, mutta sairastuneille puolestaan matala (50 %). Kokonaisuutena malli siis toimii korkeintaan keskinkertaisesti, ja sitä voidaan käyttää vain, jos halutaan arvioida skitsofreniaan sairastumiselta suojaavia tekijöitä.
- b) Mallin luokittelutarkkuus on korkea sekä koko aineistolle (83 %), sairastumattomille että sairastuneille. Malli toimii siis kokonaisuutena hyvin ja sen avulla voidaan arvioida sekä skitsofrenialta suojaavia että sille altistavia tekijöitä.
- c) Mallin luokittelutarkkuus on matala koko aineistolle (49 %) sekä sairastumattomille (50 %) mutta korkea sairastuneille (90 %). Kokonaisuutena malli toimii siis korkeintaan keskinkertaisesti, ja sitä voidaan käyttää vain jos halutaan arvioida skitsofrenialle altistavia tekijöitä.

20 Mittaaminen, reliabiliteetti ja validiteetti

1. a) $X = 100$
 b) $T = 80$
 c) $T = 95$
 d) $e = -20$
2. a) $[0, 1]$
 b) $[-1, 1]$
3. a) $r_{xx} = .89$
 b) $r_{xx} = .40$
 c) $r_{xx} = 1$
 d) $r_{xx} = .56$
 e) $r_{xx} = .35$

4. a) Suuret havaitut pistemäärät olisivat keskimäärin kauempana todellisista pistemääristä kuin pienet
 b) Havaitut pistemäärät olisivat systemaattisesti suurempia kuin todelliset pistemäärät
 c) Havaitut pistemäärät olisivat systemaattisesti pienempiä kuin todelliset pistemäärät
5. a) $\alpha = .98$
 b) $\alpha = .94$
 c) $\alpha = .91$
6. a) $\alpha = .76$
 b) $\alpha = .81$
 c) $r_{xx} = .55$
 d) $r_{xx} = .76$
7. a) uusi $r_{xx} = .6$
 b) uusi $r_{xx} = .67$
 c) uusi $r_{xx} = .75$
8. a) Kannattaa. Spearman-Brownin kaavan perusteella voidaan havaita, että esimerkiksi viidentoista itemin lisääminen mittariin kasvattaa sen reliabiliteetin .86:ksi, ja jo viidellä lisäitemillä reliabiliteetiksi saadaan .75.
 b) Ei kannata. Alkuperäinen reliabiliteetti on jo poikkeuksellisen korkea.
 c) Ei välttämättä. Reliabiliteetin kasvattaminen edellyttäisi melko paljon uusia itemeitä (noin 20 jotta reliabiliteetiksi tulisi yli .8). Aluksi kannattaisi koittaa muokata alkuperäisiä välttämättä ja koittaa saada ne toimimaan homogeenisemmin. Jos tämä ei auta, mittaria voi koittaa pidentää.
- d) Ei kannata. Reliabiliteetin kasvattaminen esimerkiksi .8:ksi edellyttäisi lähes kahdensadan ylimääräisen itemin kehittämistä. Tässä tapauksessa mittari on todennäköisesti huonosti laadittu, ja voi olla helpompaa laatia kokonaan uusi mittari
9. a) uusi $r_{xy} = .53$
 b) uusi $r_{xy} = .82$
 c) uusi $r_{xy} = .55$
10. a) $r_{xy} = .42$
 b) $r_{xy} = .36$
 c) $r_{xy} = .31$ Huomaa että mitä korkeampi alkuperäisen mittauksen reliabiliteetti on ollut, sitä vähemmän validiteettia voidaan kasvattaa mittaria pidentämällä
11. a) $r_{xy} = .69$
 b) $r_{xy} = .56$
12. a) Ei, koska validiteettikertoimen yläraja on reliabiliteetin alkuperäisen mittarin validiteetti jaettuna sen reliabiliteetilla.
 b) Voi. Reliabiliteettikerroin on odotusarvoisesti aina validiteettikerrointa suurempi.
 c) Voi, mutta tämä edellyttää sitä että $r_{xx} = 1$ ja että mittari ennustaa kriteerisuoritusta täydellisesti.

21 Reliabiliteetin vaikutus mittaustuloksiin

1. a) $SEM = 3.16$
b) $SEM = 6.32$
c) $SEM = 14.14$
2. a) $\overline{x_T} = 100$
b) $\overline{x_X} = 100$
3. Silloin, kun $\sigma_e^2 = 0$.
4. a) σ_X^2 kasvaa
b) σ_X^2 pienenee
5. a) $r'_{xy} = .63$
b) $r'_{xy} = .83$
6. a) $r'_{xy} = .64$
b) $r'_{xy} = .83$
7. a) $r'_{xy} = .33$
b) $r'_{xy} = .27$
8. Palautuminen keskiarvoon. Poikkeuksellisen hyvä tai huono suoritus johtuu todennäköisesti sattumasta (tai mittausvirheestä),

joten seuraavalla kerralla sekä hyvät että huonot suoritukset palautuvat kohti keskiarvoa

9. Keskiarvoon palautumisen johdosta poikkeuksellisen suuret stressimittaustulokset todennäköisesti pienenisivät eniten, kun taas keskiarvoa lähellä olevat mittaustulokset pysyisivät lähes ennallaan.
10. Mittausvirheestä johtuvaa keskiarvoon palautumista ei havaittaisi, koska mittausvirheen vaikutus olisi samanlaista molemmilla mittauskerroilla. Palautuminen kohti keskiarvoa ei kuitenkaan häviäisi kokonaan, sillä monet muutkin tekijät kuten tutkittavan suoritus-tason hetkittäinen tai päivittäinen vaihtelu vaikuttaa palautumiseen.
11. a) $r_{DD} = .7$
b) $r_{DD} = .58$
12. a) $r_{ss} = .88$
b) $r_{ss} = .77$

22 Efektikoko

1. a) Jos otoskoko on suuri
b) Jos otoskoko on pieni
2. a) Efektikoko pienenee
b) Efektikoko kasvaa
- c) Efektikoko saattaa pienentyä, pysyä ennallaan tai kasvaa, riippuen siitä kuinka paljon keskiarvojen erotus kasvaa suhteessa hajontaan

- d) Efektikoko ei muutu
3. a) $d = 1.5$ (suuri efekti)
 b) $d = .5$ (keskinkertainen efekti)
 c) $d = 0$ (ei efektiä)
4. a) $d = 3.55$ (suuri efekti)
 b) $\Delta = 3.54$ (suuri efekti)
 c) $\Delta = 3.34$ (suuri efekti)
5. a) $r = -.92$
 b) $r = .08$
6. a) $d = 1.15$
 b) $d = 2.49$
 c) $d = 1.32$
7. a) $r = 0.45$
 b) $r = 0.60$
 c) $r = 0.24$
8. a) Tulos ei ole tilastollisesti merkitsevä, mutta otoskoko on hyvin pieni mutta efektikoko on suuri. Tutkimukseen kannattaa kerätä lisää dataa, jotta voidaan varmistua siitä ettei havaittu efekti johdu sattumasta.
 b) Tulos ei ole tilastollisesti merkitsevä, otoskoko on hyvin suuri ja efektikoko hyvin pieni. Lääkeaine ei selvästikään vaikuta ahdistuneisuuteen juurikaan. Lisää aineistoa ei siis kannata kerätä.
 c) Tulos on tilastollisesti merkitsevä ja efektikoko on suuri, mutta otoskoko on hyvin pieni. Koska kyseessä on kliininen tutkimus, tutkimukseen kannattaa kerätä lisää dataa, jotta voidaan varmistua siitä ettei havaittu efekti johdu sattumasta ja että

se toistuu myös suuremmalla otoksella.

- d) Tulos on tilastollisesti merkitsevä ja efektikoko on keskinkertainen, ja otoskoko on suuri. Näyttäisi siis siltä, että lääkitys vähentää ahdistuneisuutta jonkun verran. Tulosta voidaan pitää luotettavana, eikä tutkimukseen täydy kerätä lisää aineistoa.
9. a) $\hat{\eta}^2 = 0.10$, $\hat{\eta}_p^2 = 0.67$
 b) $\hat{\eta}^2 = 0.60$, $\hat{\eta}_p^2 = 0.92$
 c) $\hat{\eta}^2 = 0.25$, $\hat{\eta}_p^2 = 0.83$
10. a) Kaikkien tutkimusten otoskoot ja efektikoot ovat melko suuria. Vanhempien ja lasten sosioekonomisen aseman välillä on todennäköisesti systemaattinen yhteys.
 b) Ensimmäisen tutkimuksen efektikoko on pieni mutta myös tutkimuksen otoskoko on pieni. Kahdessa jälkimmäisessä tutkimuksessa efektikoot ovat suuria, ja toisen tutkimuksen otoskoko on pieni ja toisen hyvin suuri. Koska kaksi kolmesta tutkimuksesta (yhteensä 1020 tutkittavaa) tukee voimakasta yhteyttä ja yksi tutkimus (jonka otoskoko on hyvin pieni) heikkoa yhteyttä, on melko todennäköistä että lasten sosioekonomisen aseman välillä on systemaattinen yhteys.
 c) Tutkimusten otoskoot ovat pieniä ja efektikoot hyvin vaihtelevia. Tutkimusten perusteella ei voida tehdä luotettavia päätelmiä vanhempien ja lasten sosioekonomisen aseman yhteydestä.

23

Faktorianalyysi

1. Faktori on latentti yhdistelmämuuttuja, jonka olemassaolo voidaan päätellä havaittujen muuttujien yhteisvaihtelun perusteella.
2. Faktorien avulla pyritään kuvaamaan havaittujen muuttujien yhteisvaihtelua.
3. Faktorit tiivistävät havaitun yhteisvaihtelun pienempään lukujoukkoon, faktorien latauksiin, joita lasketaan yksi kappale jokaiselta faktorilta jokaista havaittua muuttujaa kohden.
4. Aineiston tulisi olla kooltaan suuri, mielellään yli 500 havaintoa. Lisäksi tutkittavia tulisi olla vähintään 20 kertaa niin paljon kuin faktoreita – esimerkiksi 500 hengen aineistosta voitaisiin ekstraktoida luotettavasti noin 25 faktoria.
5. a) Vähintään 1 000 tutkittavaa
b) Vähintään 1 000 tutkittavaa
c) Laskutoimituksen perusteella vähintään 60 tutkittavaa, mutta yleensä faktoroinnin suorittamiseen tarvitaan aina yli 300 tutkittavaa, mitä tässä tapauksessa voidaan pitää alarajana.
6. Koska yhdelle aineistolle voidaan periaatteessa löytää äärettömän monta faktoriratkaisua, on riski tyypin 1 virheen tekemiseen hyvin korkea, ts. aineistosta voidaan faktorianalyysin avulla "löytää" ilmiöitä, jotka ovat teoreettisesti täysin mielettömiä.
7. Kovarianssit eivät ole standardoituja kuten korrelaatiot. Tällöin faktorianalyysin tulokseen vaikuttaisivat eniten sellaiset muuttujat, joiden vaihteluväli on suuri ja joiden kovarianssitkin ovat siten suuria.
8. a) Ryhmäfaktorit selittävät yli yhden, mutta eivät kuitenkaan kaikkien havaittujen muuttujien vaihtelua.
b) Yleiset faktorit selittävät kaikkien havaittujen muuttujien vaihtelua.
c) Spesifi faktori selittää vain yhden havaitun muuttujan vaihtelua.
9. Rotaation avulla voidaan muuttaa faktoriratkaisu teoreettisesti mielekkäämmäksi, ts. maksimoida jokaisen muuttujan lataukset yhteen faktoriin muuttamatta kuitenkaan faktorien yhteenlaskettua selitysosuutta.
10. Suorakulmainen rotaatio säilyttää faktorit korreloimattomina, vinorotaatio puolestaan antaa faktoreiden korreloida keskenään. Yleensä kannattaa käyttää vinorotaatiota, koska suurimman osan faktorien kuvaamista käsitteistä oletetaan korreloivan ainakin jonkin verran toistensa kanssa. Suorakulmainen rotaatio soveltuu käytettäväksi ainoastaan sellaisissa tilanteissa, joissa faktorien kuvaamien käsitteiden ei oleteta korreloivan keskenään ollenkaan.

11. a) Lataukset ilmaisevat, kuinka paljon faktori selittää muuttujan vaihtelusta.
- b) Faktorien ominaisarvot kuvaavat, kuinka paljon kunkin faktorin avulla kyetään selittämään aineiston kokonaisvaihtelusta.
- c) Kommunaliteetti ilmaisee, kuinka paljon yhden muuttujan vaihtelusta kyetään yhteensä selittämään kaikkien faktorien avulla.
12. a) [0,1]
- b) [0,1]
- c) [0,1]
13. a) Jos aineisto on suuri ja normaali-jakaumaoletus on voimassa, kannattaa yleensä käyttää maximum likelihood -ekstraktointimenetelmää.
- b) Jos normaalijakaumaoletus ei täysin toteudu, kannattaa maximum likelihood -ekstraktointin sijaan käyttää yleistettyjen neliösummien menetelmää.
14. Hyvä faktoriratkaisu on teoreettisesti perusteltu, se selittää aineiston yhteisvaihtelusta mahdollisimman paljon ja siinä pitää olla mahdollisimman vähän faktoreita. Lisäksi malliin pitäisi tulla mahdollisimman paljon itseisarvoltaan pieniä ja suuria latauksia sekä vähän itseisarvoltaan keskinkertaisia latauksia.
15. a) Koska aineisto on normaalisti jakautunut ja suuri, niin käytetään maximum likelihood -ekstraktointia.
- b) Aineistosta ekstraktoituu todennäköisesti kaksi faktoria, koska empatiakyky, emootioiden tunnistuskyky ja toisten ihmisten toiminnan attribuointikyky korreloivat keskenään mutta eivät sosiaalisen ongelmanratkaisutaidon ja sosiaalisen joustavuuden kanssa. Näiden muuttujien selittämiseen riittää yksi ryhmäfaktori η_1 . Vastaavasti sosiaalinen ongelmanratkaisu-taito ja sosiaalinen joustavuus korreloivat keskenään mutta eivät empatiakyvyn, emootioiden tunnistamiskyvyn ja toisten ihmisten toiminnan attribuointikyvyn kanssa. Näiden muuttujien selittämiseen riittää myös yksi faktori, η_2 .
- c) Koska empatiakyvyn, emootioiden tunnistuskyvyn ja toisten ihmisten toiminnan attribuointikyvyn keskinäiset korrelaatiot ovat korkeita ja positiivisia, tulee niiden latauksista faktorille η_1 todennäköisesti suuria ($>.7$). Koska nämä muuttujat eivät korreloi toiseen faktoriin (η_2) liittyvien muuttujien (sosiaalinen ongelmanratkaisu-taito ja sosiaalinen joustavuus) kanssa, tulee muuttujien latauksista faktorille η_2 hyvin pieniä. Vastaavasti, koska sosiaalisen ongelmanratkaisutaidon ja sosiaalisen joustavuuden keskinäinen korrelaatio on korkea, tulee näiden muuttujien latauksista faktorille η_2 todennäköisesti suuria. Koska muuttujat eivät korreloi toiseen faktoriin (η_1) liittyvien muuttujien (empatiakyky, emootioiden tunnistuskyky ja toisten ihmisten toiminnan attribuointikyky) kanssa, tulee muuttujien latauksista faktorille η_1 todennäköisesti hyvin pieniä.
- d) Koska molempiin faktoreihin liittyvien muuttujien keskinäiset korrelaatiot ovat suuria ja samansuuntaisia, tulee sekä faktorien ominaisarvoista että kaikkien muuttujien kommunaliteeteista todennäköisesti suuria (> 0.7)

24 Moniulotteinen asteikointi

1. Faktorianalyysi pyrkii tiivistämään muuttujien yhteisvaihtelun latenteihin muuttujiin, moniulotteinen asteikointi puolestaan pyrkii esittämään havaintoarvojen väliset etäisyydet ulottuvuuksien avulla ja klusterianalyysi ryhmittelee muuttujilla samankaltaisia arvoja saaneet havainnot klustereiksi.
2. Faktorianalyysi soveltuu näistä menetelmistä yleensä parhaiten käyttäytymistieteelliseen tutkimukseen, sillä sen avulla voidaan tarkastella muuttujien välisiä yhteyksiä, mikä on eräs käyttäytymistieteellisen tutkimuksen peruskysymyksistä. Moniulotteinen asteikointi sopii lähinnä erilaisten objektien (ilmeet, eleet, sanat...) samankaltaisuuksien tiivistämiseen, ja klusterianalyysi puolestaan koehenkilöjoukkojen ryhmittelemiseen.
3. Joko suoralla arvioinnilla jossa arvioidaan suoraan jokaisen mahdollisen ärsykeparin samankaltaisuus tai epäsuoralla arvioinnilla, jossa havainnot luokitellaan ryhmiin niiden samankaltaisuuden perusteella.
4. Asteikointiratkaisun toimivuus arvioidaan stressiarvojen perusteella. Ne ilmoittavat, kuinka tarkasti analyysissa muodostetut ulottuvuudet kykenevät kuvaamaan havaintojen väliset todelliset etäisyydet.
5. Koska S_0 perustuu havaittujen ja mallinnettujen etäisyyksien erotuksiin, se ei ole kovin herkkä pienille parannuksille asteikointiratkaisun toimivuudessa. S_1 :ssa tämä ongelma on ratkaistu suhteuttamalla jokainen havaitun ja mallinnetun etäisyyden välinen erotus havaittuun etäisyyteen.
6. Koska moniulotteinen asteikointi ei tarkastele muuttujien välistä kovarianssia, vaan havaintojen välisiä etäisyyksiä.
7. a) $\binom{10}{2} = 45$
 b) $\binom{25}{2} = 300$
 c) $\binom{100}{2} = 4950$

25 Klusterianalyysi

1. Sellaisissa tutkimuksissa, joissa halutaan luokitella tutkittavat henkilöt tai objektit keskenään samankaltaisiin ryhmiin.
3. Jotta vaihteluväliltään ja keskiarvoltaan erilaiset muuttujat olisivat yhtä merkityksellisiä, kun klusterointia suoritetaan.
4. Menetelmä on täysin eksploraatiivinen eikä se siten sovellu kovinkaan hyvin tutkimusaineistoa koskevien hypoteesien testaamiseen.
5. K-keskiarvoklusteroinnissa havainnot pyritään luokittelemaan ennalta määriteltyyn määrään ryhmiä, kun taas hierarkkisessa klusteroinnissa klustereita yhdistellään suuremmiksi klustereiksi niin kauan, kunnes kaikki havainnot kuuluvat samaan klusteriin.
6. Molemmissa aineistoissa näyttäisi muodostuvan kaksi tai kolme klusteria.

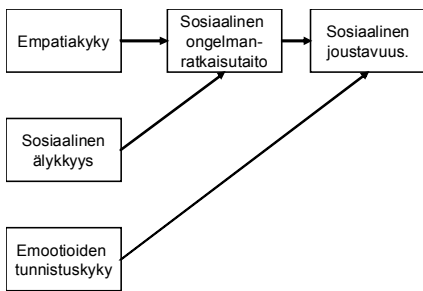
26 Rakenneyhtälömallit

1. Faktoriratkaisu voidaan määritellä tarkasti etukäteen ja sen jälkeen testata, miten se sopii aineistoon.
2. Havaitut muuttujat ovat empiirisiä, mitattuja muuttujia. Latentit muuttujat määritellään havaittujen muuttujien yhteisvaihtelun perusteella.
3. Endogeenisten muuttujien vaihtelua selitetään mallin muilla muuttujilla, eksogeenisten muuttujien vaihtelua puolestaan ei selitetä muilla mallin muuttujilla.
4. Verrataan havaittujen muuttujien korrelaatiomatriisia mallin ennustamaan korrelaatiomatriisiin χ^2 -testin avulla. Jos matriisit eivät poikkea toisistaan, malli sopii aineistoon.
5.
 - a) Havaittujen muuttujien yhteyksiä toisiinsa
 - b) Latenttien muuttujien yhteyksiä toisiinsa
 - c) Latenttien muuttujien välisiä yhteyksiä
6. Ennakkotiedon eli tutkimuskysymykseen liittyvän teoreettisen ja empiirisen tiedon perusteella

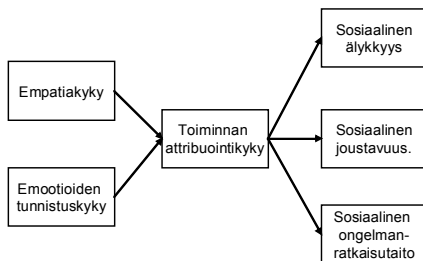
7. Vahvimpia kausaalipäätelmiä voidaan tehdä silloin, kun on käytetty kokeellista asetelmaa ja aineisto analysoidaan rakenneyhtälömallien avulla. Heikompia kausaalipäätelmiä voidaan tehdä eksploratiivisissa tutkimuksissa silloin, jos on kerätty esimerkiksi pitkittäisaineisto – useita toistettuja mittauksia mutta ei kokeellista asetelmaa – ja on pyritty kontrolloimaan mahdollisimman tarkkaan muiden kuin riippuvan muuttujan vaikutukset riippumattomiin muuttujiin.

8.

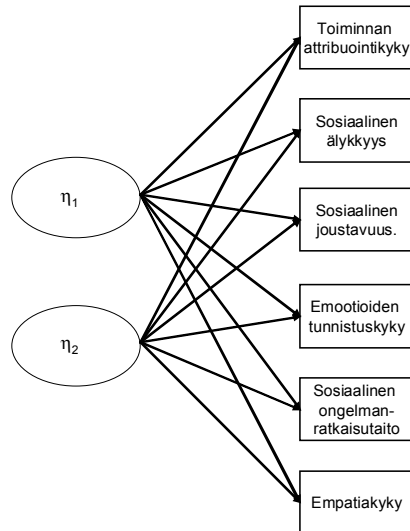
a)



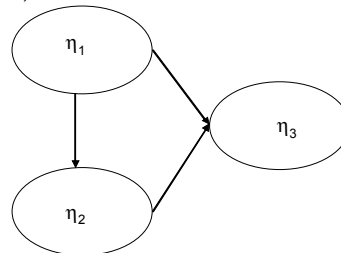
b)



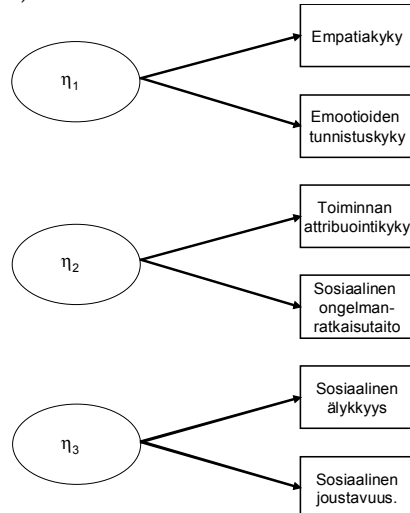
c)



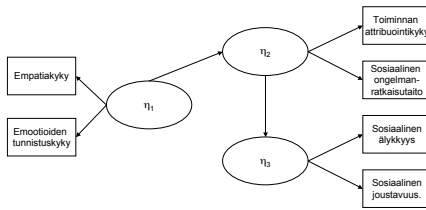
d)



e)



f)



9. eksogeeninen: η_1 , endogeenisia: η_2 ja η_3

eksogeenisiä: η_1, η_2, η_3 . Endogeenisiä kaikki havaitut muuttujat.

eksogeeninen: η_1 . Endogeenisiä η_2 ja η_3 ja kaikki havaitut muuttujat

Lähteet

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory – Second Edition manual*. San Antonio: The Psychological Corporation.
- Bliss, I., Kujala, T., & Hämäläinen, H. (2004). Comparison of blind and sighted participants' performance in a letter recognition working memory task. *Cognitive Brain Research*, 18, 273–277.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 11, 717–726.
- Calvo, M. G., & Lang, P. J. (2004). Gaze patterns when looking at emotional pictures: Motivationally biased attention. *Motivation and Emotion*, 28, 221–243.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99, 550–553.
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it: Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5, 490–493.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141–145.

Aineistot

Liite I: Lukeminen ja työmuisti

Työmuistilla viitataan lyhytkestoiseen prosessointitiltiltioon, jonka avulla ihmiset 1) prosessoivat informaatiota, 2) tallettavat informaatiota pitkäkestoiseen muistiin ja 3) palauttavat informaatiota pitkäkestoisesta muistista. Työmuistin kapasiteetti on kuitenkin rajallinen, ja ihmisten välillä on eroja työmuistin kapasiteetissa eli siinä, kuinka paljon informaatiota he kykenevät käsittelemään työmuistissa samanaikaisesti. On havaittu, että työmuistin kapasiteetti vaikuttaa monentyypisissä tehtävissä, kuten lukemisessa, suoriutumiseen (Daneman & Carpenter, 1980).

Tässä tutkimuksessa on haluttu selvittää, miten työmuistin kapasiteetti vaikuttaa tekstin lukemiseen käytettävään aikaan ja luetun tekstin muistamiseen. Tutkimuksen aluksi tutkittavien työmuistin kapasiteetti mitattiin ja heidät luokiteltiin kapasiteetin perusteella kolmeen ryhmään (matala, keskinkertainen ja korkea kapasiteetti). Tämän jälkeen tutkittavat lukivat tietokoneen ruudulta tekstin, ja tekstin lukemiseen käytetty aika mitattiin. Kaksi tuntia kokeen jälkeen tutkija pyysi koehenkilöitä kirjoittamaan paperille kaiken, mitä he muistivat tekstistä. Jälkeenpäin arvioitiin, kuinka monta prosenttia kukin tutkittavista oli kyennyt palauttamaan tekstin sisällöstä.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Ikä = tutkittavan ikä vuosina
3. Sukupuoli = tutkittavan sukupuoli
4. TM = tutkittavan työmuistin luokiteltu kapasiteetti
(1 = matala, 2 = keskinkertainen, 3 = korkea)
5. Luku-aika = tekstin lukemiseen käytetty aika minuutteina
6. Kokonaispalautus = tekstin sisällön muistaminen prosentteina (0–100%)

KH	Sukupuoli	Ikä	TM	Luku-aika	Kokonaispalautus
1	mies	22	1	56	77
2	mies	23	1	57	76
3	mies	25	1	54	74
4	mies	22	1	53	77
5	mies	23	1	55	72
6	nainen	24	1	45	71
7	nainen	25	1	49	70
8	nainen	28	1	52	80
9	nainen	29	1	53	81
10	nainen	25	1	54	71
11	mies	26	2	47	80
12	mies	25	2	44	82
13	mies	24	2	49	81
14	mies	26	2	46	82
15	mies	28	2	50	84
16	nainen	30	2	40	85
17	nainen	19	2	48	83
18	nainen	20	2	51	87
19	nainen	22	2	49	86
20	nainen	21	2	50	84
21	mies	24	3	37	95
22	mies	25	3	38	96
23	mies	23	3	39	94
24	mies	22	3	37	93
25	mies	27	3	38	92
26	nainen	28	3	36	80
27	nainen	24	3	37	91
28	nainen	25	3	35	90
29	nainen	19	3	33	90
30	nainen	20	3	36	91

Liite 2: Koulukiusaaminen

Tässä tutkimuksessa on selvitetty, miten lasten kiusatuksi tuleminen, itsetunto, aggressiivisuus ja työmuistin kapasiteetti ovat yhteydessä toisiinsa. Tutkittavien kiusatuksi tuleminen, itsetunto ja aggressiivisuus selvitettiin kyselylomakkeen avulla. Työmuistikapasiteetti mitattiin numerosarjatehtävän avulla, jossa tutkittavan tulee toistaa tutkijan hänelle luettelemia numerosarjoja (esimerkiksi 4–6–8–2). Numerosarjat pitenevät koko ajan, ja niiden pituutta kasvatetaan kunnes tutkittava ei pysty enää palauttamaan numerosarjaa.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Sisarusten lkm. = tutkittavan sisarusten lukumäärä
5. Kiusatuksi tuleminen = lapsen kokemus siitä, kuinka paljon häntä kiusataan koulussa (1 [ei ollenkaan] – 10 [erittäin paljon])
6. Itsetunto = lapsen itsetunto (1 [erittäin matala] – 10 [erittäin korkea])
7. Aggressio = tutkittavan aggressiivisuus (1 [erittäin matala] – 5 [erittäin korkea])
8. Työmuisti = tutkittavan työmuistin kapasiteetti (1 [matala] – 5 [korkea])

KH	Sukupuoli	Ikä	Sisarusten lkm	Kiusatuksi-tuleminen	Itsetunto	Aggressio	Työmuisti
1	tyttö	11	2	4	6	1	5
2	tyttö	11	2	3	5	5	1
3	tyttö	12	0	6	4	4	2
4	tyttö	12	1	7	2	2	3
5	tyttö	11	2	6	3	3	4
6	tyttö	12	1	7	3	3	3
7	tyttö	11	2	1	8	2	5
8	tyttö	11	0	9	3	1	4
9	tyttö	12	2	1	7	5	2
10	tyttö	12	2	2	6	4	3
11	poika	12	1	3	7	5	2
12	poika	11	1	4	5	4	1
13	poika	11	0	5	5	3	3
14	poika	11	1	5	4	5	5
15	poika	11	0	7	2	4	5
16	poika	12	0	1	7	2	3
17	poika	11	1	2	6	3	2
18	poika	12	1	9	2	4	3
19	poika	12	2	5	4	2	1
20	poika	11	1	4	6	4	3

Liite 3: Tarkkaavaisuuden suuntautuminen katsevihjeiden perusteella

Visuaalisen tarkkaavaisuuden avulla ihminen kykenee valitsemaan ympäristön ärsykeistä ne, jotka edellyttävät monimutkaisempaa prosessointia. Tarkkaavaisuutta voi suunnata tahdonalaisesti, mutta on myös osoitettu, että tietynlaiset ympäristön tapahtumat suuntaavat tarkkaavaisuutta refleksiivisesti, havaitsijan tahdosta riippumatta. Friesen ja Kingstone (1998) osoittivat, että toisen henkilön sivulle suunnatun katseen havaitseminen suuntaa havaitsijan tarkkaavaisuuden refleksiivisesti katseen suuntaan.

Tässä tutkimuksessa tutkittavien tehtävänä oli painaa reaktionappulaa mahdollisimman nopeasti, kun he havaitsivat ruudun vasempaan tai oikeaan reunaan sijoitetun kohdeärsyksen (asteriskin). Tasan 200 ms ennen asteriskin esittämistä ruudun keskellä esitettiin vihjeärsyke. Vihjeärsykkeenä oli kuva ihmiskasvoista, jonka katse oli suunnattu joko 1) samalle puolelle ruutua, minne asteriski ilmestyy (kongruentti), 2) suoraan eteenpäin (neutraali) tai 3) eri puolelle ruutua (inkongruentti tilanne) kuin missä asteriski esiintyy. Jokainen tutkittava teki 20 kappaletta kutakin tehtävätyyppiä. Erityyppiset tehtävät oli sekoitettu keskenään ja tehtävät esitettiin satunnaisessa järjestyksessä.

Jos katsevihjeet suuntaavat tarkkaavaisuutta automaattisesti, reaktioaikojen pitäisi olla nopeimpia kongruentissa tilanteessa. Tällöinhän tarkkaavaisuus on jo katsevihjeen ansiosta automaattisesti suuntautunut sinne, minne asteriski ilmestyy. Vastaavasti inkongruentissa tilanteessa reaktioaikojen pitäisi olla hitaimpia, koska katsevihje suuntaa tarkkaavaisuuden aluksi eri puolelle kuin asteriski ilmestyy. Reagoidakseen asteriskiin tutkittavan tulee kuitenkin siirtää tarkkaavaisuutensa asteriskiin, minkä pitäisi hidastaa reaktioaikoja.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Kongruentti = koehenkilön reaktioaikojen keskiarvo (ms) kongruenteilla trialeilla
5. Neutraali = koehenkilön reaktioaikojen keskiarvo (ms) neutraaleilla trialeilla
6. Inkongruentti = koehenkilön reaktioaikojen keskiarvo (ms) inkongruenteilla trialeilla

KH	Sukupuoli	Ikä	Kongruentti	Neutraali	Inkongruentti
1	mies	22	358	363	363
2	mies	21	313	320	325
3	mies	29	280	292	294
4	mies	32	266	282	293
5	mies	33	303	306	316
6	nainen	28	221	245	260
7	nainen	27	284	290	315
8	nainen	24	277	274	290
9	nainen	25	342	359	372
10	nainen	23	311	315	340

Liite 4: Emootioiden kokeminen

Emootiot eli tunteet viittaavat prosesseihin, joiden avulla yksilöt voivat muuttaa toiminta- valmiuttaan siten, että kyky reagoida adaptiivisesti ympäristön uhkaaviin ja mahdollisuuksia tarjoaviin tilanteisiin lisääntyy. Tutkimuksissa on osoitettu, että ns. perusemootiot – mielihyvä, viha, inho, pelko, suru ja hämmästyminen – esiintyvät samanlaisina kaikissa kulttuureissa (Ekman, 1992). Tässä tutkimuksessa on selvitetty, millaisia emootioita opiskelijat kokevat aloittaessaan uuden tilastotieteen kurssin. Tutkittavat täyttivät kurssin ensimmäisellä luen- nolla lomakkeen, jossa heidän tuli arvioida asteikolla 1 (ei lainkaan) – 9 (erittäin paljon), kuinka paljon he kyseisellä hetkellä kokivat kutakin kuudesta perusemootiosta.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Kätsisyys = tutkittavan kätsisyys (vasen/oikea)
5. Mielihyvä = mielihyvän kokemus mittaushetkellä
6. Viha = vihan kokemus mittaushetkellä
7. Inho = inhon kokemus mittaushetkellä
8. Pelko = pelon kokemus mittaushetkellä
9. Suru = surun kokemus mittaushetkellä
10. Hämmästyminen = hämmästyksen kokemus mittaushetkellä

KH	Sukupuoli	Ikä	Kätsisyys	Mielihyvä	Viha	Inho	Pelko	Suru	Hämmästyminen
1	nainen	23	1	9	1	1	1	1	2
2	nainen	20	1	8	2	1	2	3	1
3	nainen	20	2	2	1	1	1	1	1
4	mies	21	2	6	2	1	1	1	3
5	nainen	20	1	8	1	2	2	3	2
6	mies	21	2	8	2	1	3	2	2
7	mies	30	2	6	1	1	3	3	5
8	nainen	19	2	8	1	1	5	2	3
9	nainen	19	2	8	1	2	7	3	4
10	nainen	21	2	5	2	3	4	4	2
11	nainen	19	2	7	1	1	2	1	2
12	nainen	21	2	9	1	1	2	2	2
13	mies	19	2	7	2	1	1	2	8
14	nainen	25	2	8	1	1	2	2	3
15	mies	25	2	8	2	4	6	4	5
16	nainen	20	1	5	2	2	3	4	1
17	nainen	21	2	9	1	1	3	2	4
18	nainen	20	2	8	1	1	2	1	1
19	mies	23	1	7	3	2	5	3	4
20	mies	22	2	6	2	3	4	2	3

Liite 5: Työn kuormittavuus

Tässä tutkimuksessa on selvitetty, miten erilaiset tutkittaviin sekä heidän työoloihinsa liittyvät muuttujat vaikuttavat työn kuormittavuuden kokemiseen. Tutkittaviksi valittiin suomalaisista yrityksistä 30 eri-ikäistä miestä ja naista. Tutkittavien itsetunto arvioitiin kyselylomakkeen avulla. Lisäksi tutkittavilta mitattiin heidän kokemuksensa työssä viihtymisestä, työilmapiiristä ja työn kuormittavuudesta. Edellä mainittujen ominaisuuksien arviointi tapahtui siten, että tutkittavalle esitettiin jana, jonka päissä oli esitetty mitattavan ominaisuuden ääripäät (esim. viihdyn työssäni erittäin huonosti / viihdyn työssäni erittäin hyvin). Mittaus tehtiin siten, että tutkittava piirsi janaan pystysuoran viivan kohtaan, joka vastasi hänen kokemustaan. Vastaus muunnettiin lukuarvoksi siten, että tutkija mittasi viivoittimella mihin kohtaan janaa tutkittava oli viivan piirtänyt.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Koulutus = tutkittavan koulutustaso (perusaste, keskiaste, korkea-aste)
5. Itsetunto = tutkittavan itsetunto (0 [erittäin matala] – 1 000 [erittäin korkea])
6. Viihtyminen = työssä viihtyminen (1 [erittäin huonosti] – 5 [erittäin hyvin])
7. Ilmapiiri = tutkittavan kokemus työpaikan ilmapiiristä (1 [erittäin huono] – 5 [erittäin hyvä])
8. Kuormittavuus = tutkittavan kokemus työn kuormittavuudesta (1 [vähän kuormittavaa] – 5 [paljon kuormittavaa])

KH	Sukupuoli	Ikä	Koulutus	Itsetunto	Viihtyminen	Ilmapiiri	Kuormittavuus
1	nainen	34	perusaste	44	3,7	4,2	1,1
2	mies	35	keskiaste	75	4,3	3,3	4,6
3	nainen	24	korkea-aste	75	4,3	3,5	4,6
4	mies	26	perusaste	80	4,4	2,9	4,6
5	nainen	27	keskiaste	79	4,4	2,7	4,6
6	mies	34	korkea-aste	75	4,3	3,3	4,6
7	nainen	50	perusaste	74	4,3	3,7	4,3
8	mies	44	keskiaste	53	4,0	3,9	3,6
9	nainen	48	korkea-aste	78	4,4	2,9	4,6
10	mies	36	perusaste	76	4,3	2,9	4,6
11	nainen	34	keskiaste	79	4,4	2,8	4,6
12	mies	35	korkea-aste	64	4,2	3,7	4,4
13	nainen	36	perusaste	78	4,4	3,7	4,5
14	mies	37	keskiaste	66	4,2	3,9	4,3
15	nainen	23	korkea-aste	78	4,4	4,2	3,3
16	mies	24	perusaste	56	4,3	3,9	3,2
17	nainen	20	keskiaste	57	3,1	2,9	3,1
18	mies	29	korkea-aste	89	4,2	3,8	2,1
19	nainen	48	perusaste	78	5,0	4,8	1,7
20	mies	47	keskiaste	66	3,6	3,2	2,8
21	nainen	45	korkea-aste	67	3,2	2,7	3,0
22	mies	44	perusaste	68	4,3	4,0	1,2
23	nainen	45	keskiaste	64	3,3	3,0	2,4
24	mies	43	korkea-aste	75	2,4	2,5	4,3
25	nainen	38	perusaste	72	4,0	3,7	1,3
26	mies	34	keskiaste	70	2,3	2,5	4,2
27	nainen	33	korkea-aste	68	3,3	3,2	2,9
28	mies	34	perusaste	70	3,5	3,2	3,0
29	nainen	32	keskiaste	88	3,4	4,0	3,1
30	mies	36	korkea-aste	91	4,0	4,2	1,2

Liite 6: Emootiot ja visuaalinen tarkkaavaisuus

On havaittu, että jos ympäristössä on sekä emotionaalisia (miellyttäviä/epämiellyttäviä) kohteita että neutraaleja kohteita, ihmiset suuntaavat katseensa todennäköisesti ensin emotionaalsiin kohteisiin (Calvo & Lang, 2004). Tämä ilmiö selitetään sillä, että emotionaaliset ärsykkeet viestivät organismin hyvinvoinnille tärkeiden asioiden läsnäolosta, ja adaptiivinen toiminta edellyttää niiden nopeaa rekisteröimistä.

Tässä tutkimuksessa selvitettiin silmänliikerekisteröinnin avulla, miten tutkittavien katse suuntautuu, kun heille esitetään samanaikaisesti emotionaalisia ja neutraaleja kuvia. Tutkittaville näytettiin ruudulla kohdekuvia ihmisistä miellyttävissä, neutraaleissa ja epämiellyttävissä tilanteissa. Kohdekuvien kanssa näytettiin aina neutraali kuva, joka esitti jotain elotonta esinettä. Tutkittavien tehtävänä oli katsella kuvia vapaasti kolmen sekunnin ajan ja arvioida, olivatko ruudulla esitetyt kuvat yhtä miellyttäviä. Tutkimuksessa laskettiin jokaiselle koehenkilölle ensimmäisen fiksaation (silmän pysähtymisen) todennäköisyys kuvatyyppeittäin, ts. todennäköisyys sille, että tutkittava oli katsonut ensimmäiseksi kyseistä kuvatyyppeä. Tämän perusteella voitiin arvioida, millaisiin kuviin tutkittavan tarkkaavaisuus suuntautui ensimmäiseksi.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Epämiellyttävät = todennäköisyys sille, että ensimmäinen fiksaatio on osunut epämiellyttävään kuvaan
5. Neutraalit = todennäköisyys sille, että ensimmäinen fiksaatio on osunut neutraaliin kuvaan
6. Miellyttävät = todennäköisyys sille, että ensimmäinen fiksaatio on osunut miellyttävään kuvaan

KH	Sukupuoli	Ikä	Epämiellyttävät	Neutraalit	Miellyttävät
1	mies	23	0,5	0,56	0,69
2	mies	24	0,6	0,50	0,63
3	mies	24	0,6	0,50	0,69
4	mies	26	0,6	0,56	0,63
5	mies	27	0,5	0,56	0,50
6	mies	23	0,9	0,63	0,69
7	mies	24	0,7	0,56	0,69
8	mies	19	0,5	0,50	0,50
9	mies	20	0,5	0,50	0,63
10	mies	21	0,6	0,44	0,50
11	nainen	22	0,6	0,56	0,56
12	nainen	23	0,7	0,63	0,81
13	nainen	24	0,6	0,50	0,63
14	nainen	26	0,6	0,50	0,50
15	nainen	26	0,6	0,50	0,56
16	nainen	25	0,6	0,50	0,56
17	nainen	23	0,5	0,50	0,50
18	nainen	24	0,6	0,50	0,69
19	nainen	20	0,8	0,63	0,75
20	nainen	19	0,5	0,44	0,56

Liite 7: Stroopin värisanatehtävä

Stroopin värisanatehtävän (Stroop, 1935) avulla tutkitaan niin sanottua valikoivaa tarkkaavaisuutta. Tehtävässä tutkittavalle esitetään ruudulla värien nimiä ("punainen", "sininen", "vihreä"). Sanat on kirjoitettu joko 1) kongruentilla värillä, joka vastaa sanan merkitystä (esimerkiksi sana "punainen" punaisella värillä) tai 2) inkongruentilla värillä, joka ei vastaa sanan merkitystä (esimerkiksi sana "punainen" vihreällä värillä). Tutkittavan tehtävänä on vastata mahdollisimman nopeasti nappulaa painamalla, oliko ruudulla esitetty sana väritään punainen, sininen vai vihreä. Sanan merkitykseen ei tule kiinnittää huomiota.

Jos sanat on kirjoitettu inkongruentilla värillä, havaitaan ns. interferenssiefekti. Tutkittavien reaktioajat ovat hitaampia inkongruentissa kuin kongruentissa tilanteessa, koska sanan semanttinen informaatio ja väri ovat ristiriidassa keskenään. Tutkittavat siis prosessoivat sanan semanttisen informaation automaattisesti, mutta sitä ei tule käyttää vastaamisen perustana ja se tulee jättää huomiotta. Tämä taas tuottaa kognitiivista kuormitusta, joka hidastaa vastauksen tuottamista.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Kongruentti = reaktioajat (ms) kongruentissa tilanteessa (sanat merkitys ja väri vastaavat toisiaan)
5. Inkongruentti = reaktioajat (ms) inkongruentissa tilanteessa (sanat merkitys ja väri eivät vastaa toisiaan)

KH	Sukupuoli	Ikä	Kongruentti	Inkongruentti
1	mies	30	371	452
2	mies	29	362	432
3	mies	28	372	462
4	mies	31	403	448
5	mies	27	371	438
6	mies	27	388	442
7	mies	26	356	412
8	mies	24	379	423
9	nainen	25	388	444
10	nainen	26	391	398
11	nainen	27	362	400
12	nainen	28	366	412
13	nainen	29	369	420
14	nainen	22	362	390
15	nainen	23	372	421

Liite 8: Depression hoitaminen

Noin 6 % väestöstä kärsii depressiosta (masennuksesta). Depression oireisiin kuuluvat tyyppillisesti mm. masentunut mieliala, kyvyttömyys kokea mielihyvää, uupumus sekä itseluottamuksen ja itsearvostuksen menetys. Depression hoitomuotona käytetään sekä erilaisia mielialalääkkeitä että psykoterapiaa. Tässä tutkimuksessa on vertailtu lääkehoidon ja psykoterapian vaikuttavuutta depression hoitomuotoina.

Tutkittavat on jaettu aluksi satunnaisesti lääkehoito- ja psykoterapiahoitoa saaviin ryhmiin, ja heidän depressionsa taso on mitattu BDI-II-kyselyllä (Beck, 1996). Tämän jälkeen on seurannut kahden kuukauden mittainen hoitojakso, jonka jälkeen tutkittavien depression taso on mitattu uudelleen.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Hoito = kokeellinen ryhmä, johon tutkittava kuuluu (lääkitys, psykoterapia)
5. BDI-II alku = tutkittavan depressio tutkimuksen alussa mitattuna BDI-II-lomakkeella
6. BDI-II loppu = tutkittavan depressio tutkimuksen lopussa mitattuna BDI-II-lomakkeella

KH	Sukupuoli	Ikä	Hoito	BDI-IIalku	BDI-IIloppu
1	mies	33	lääkitys	18	11
2	nainen	35	lääkitys	17	12
3	mies	36	lääkitys	18	13
4	nainen	37	lääkitys	19	12
5	mies	32	lääkitys	15	10
6	nainen	29	lääkitys	16	9
7	mies	30	lääkitys	15	8
8	nainen	34	lääkitys	17	7
9	mies	35	lääkitys	18	10
10	nainen	33	lääkitys	19	13
11	mies	32	terapia	17	12
12	nainen	27	terapia	15	14
13	mies	39	terapia	16	15
14	nainen	40	terapia	17	14
15	mies	35	terapia	18	13
16	nainen	37	terapia	19	12
17	mies	36	terapia	15	16
18	nainen	35	terapia	17	17
19	mies	35	terapia	18	15
20	nainen	33	terapia	15	14

Liite 9: Sokeiden työmuisti

Aistipuutoksista kärsivät henkilöt, kuten sokeat, kykenevät korvaamaan puuttuvan aistin toimintoja osittain tehostamalla muiden aistien toimintoja. Esimerkiksi sokeat henkilöt käyttävät käsiään pistekirjoituksen lukemiseen, ja heidän sormiensa tuntoaistin on osoitettu olevan herkempi kuin näkevien ihmisten.

Työmuisti on lyhytkestoinen prosessointitaltio, jonka avulla ihmiset prosessoivat informaatiota, tallettavat informaatiota pitkäkestoiseen muistiin ja palauttavat informaatiota pitkäkestoisesta muistista. Ihmisten välillä on eroja työmuistin kapasiteetissa eli siinä, kuinka paljon informaatiota he kykenevät käsittelemään työmuistissa samanaikaisesti. Koska sokeiden tuntoaisti on näkevien tuntoaistia herkempi, voidaan kenties olettaa, että he kykenevät myös ylläpitämään työmuistissaan enemmän tuntoaistin perusteella hankittua informaatiota kuin näkevät ihmiset.

Tässä tutkimuksessa vertailtiin sokeiden ja näkevien henkilöiden työmuistin prosessointikapasiteettia 1) audiitiiviselle ja 2) taktiiliselle informaatiolle. Työmuistin kapasiteetti mitattiin ns. n-back-tehtävän avulla. Tehtävässä tutkittavat pitivät silmänsä suljettuina, ja heille esitettiin tasaisella tahdilla kirjaimia joko äänteinä kuulokkeiden avulla (audiitiivinen tehtävä) tai kohokuvioina (taktiilinen tehtävä). Tutkittavien tehtävänä oli jokaisen kirjaimen kohdalla palauttaa kirjain, joka oli esitetty n-kirjainta ennen kulloinkin esitettävää kirjainta.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Ryhmä = tutkimusryhmä, johon tutkittava kuuluu (sokea/näkevä)
5. Taktiilinen TM = tutkittavan suoritus tuntoaistin perustuvassa työmuistitehtävässä (0–8)
6. Fonologinen TM = tutkittavan suoritus fonologisessa työmuistitehtävässä (0–8)

KH	Sukupuoli	Ikä	Ryhmä	TaktiilinenTM	FonologinenTM
1	mies	34	sokea	7	5
2	nainen	33	sokea	6	6
3	mies	35	sokea	5	4
4	nainen	37	sokea	6	5
5	mies	38	sokea	6	6
6	nainen	29	sokea	7	7
7	mies	41	sokea	6	3
8	nainen	28	sokea	7	4
9	mies	27	sokea	5	5
10	nainen	30	sokea	4	6
11	mies	34	näkevä	2	5
12	nainen	35	näkevä	3	3
13	mies	36	näkevä	4	6
14	nainen	34	näkevä	3	7
15	mies	34	näkevä	2	6
16	nainen	28	näkevä	1	5
17	mies	29	näkevä	2	4
18	nainen	27	näkevä	3	5
19	mies	40	näkevä	4	6
20	nainen	37	näkevä	2	5

Liite 10: Kasvoinversio

Ihmiset kykenevät muistamaan ja tunnistamaan kasvoja hämmästyttävällä tarkkuudella. Pystymme esimerkiksi tunnistamaan vanhan ystävämme kasvoista, vaikka emme olisi nähneet häntä kymmeniin vuosiin. Tästä erikoistumisesta seuraa kuitenkin mielenkiintoinen ilmiö. Kasvojen tunnistaminen vaikeutuu huomattavasti, jos tunnistettavia kasvoja käännetään normaalista asennosta, jossa silmät ovat nenän ja nenä suun yläpuolella (Yin, 1969). Tämän oletetaan johtuvan siitä, että kasvoja prosessoivat keskushermoston järjestelmät ovat oppineet tunnistamaan kasvoja kokonaisuuksina tietyssä asennossa ja muutokset kasvojen asennossa häiritsevät tätä kokonaisvaltaista prosessointia.

Tässä tutkimuksessa tarkasteltiin, kuinka paljon kasvoja tulee kääntää, jotta kasvoinversioefekti havaittaisiin. Tutkittaville esitettiin aina ensin 300 ms:n ajan kuva kasvoista, minkä jälkeen hänelle esitettiin 1000 ms:n ajan kohdekuva kasvoista, joka saattoi olla sama tai eri kuva kuin edellä oli esitetty. Kohdekuvat esitettiin joko 0°, 45° tai 90° asteen kallistuksella. Tutkittavan tehtävänä oli yksinkertaisesti vastata nappulaa painamalla, esiintyivätkö kuvissa samat kasvot. Tutkimuksessa mitattiin oikeiden vastausten määrää. Jos kasvojen kääntäminen vaikeuttaa kasvojen tunnistamista, pitäisi kasvojen tunnistamisen olla paljon vaikeampaa tilanteessa, jossa kohdekasvot esitetään ylösalaisin.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. inv0 astetta = oikein tunnistettujen kasvojen osuus (%), kun kohdekuvan kallistus oli 0°
5. inv45 astetta = oikein tunnistettujen kasvojen osuus (%), kun kohdekuvan kallistus oli 0°
6. inv90 astetta = oikein tunnistettujen kasvojen osuus (%), kun kohdekuvan kallistus oli 0°

KH	Sukupuoli	Ikä	inv0astetta	inv90astetta	inv180astetta
1	mies	22	99	93	67
2	nainen	23	98	93	68
3	mies	25	97	93	66
4	nainen	24	98	92	64
5	mies	22	99	94	71
6	nainen	23	97	91	72
7	mies	26	96	92	65
8	nainen	27	95	91	74
9	mies	28	98	89	72
10	nainen	21	99	88	71

Liite II: Työuupumus

Työuupumus aiheuttaa vuosittain suuren määrän töistä poissaoloja, joten on ryhdytty entistä tarkemmin selvittämään tekijöitä, jotka ovat yhteydessä työuupumuksen laukeamiseen. Tässä tutkimuksessa on selvitetty, miten koehenkilöiden väsymyksen, unen puutteen ja kuormituksen kokemisella on yhteyttä 1) toisiinsa ja 2) työuupumuksen laukeamiseen seuraavan vuoden aikana. Tutkimuksen alussa kaikilta tutkittavilta mitattiin kyselylomakkeen avulla heidän kokemuksensa väsymyksestä, unen puutteesta ja työn aiheuttamasta kuormituksesta. Tämän jälkeen tutkittavia seurattiin vuoden verran ja kaikki tutkimukseen osallistuneilla henkilöillä havaitut työuupumustapaukset rekisteröitiin.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Väsymys = tutkittavan kokema päivittäinen väsymyksen tunne (1 [ei lainkaan väsynyt] – 10 [erittäin väsynyt])
5. Unen puute = tutkittavan kokema päivittäinen unen puute (0 [riittävästi unta] – 100 [ei lainkaan riittävästi unta])
6. Kuormituksen kokemus = tutkittavan kokema päivittäinen kuormituksen tunne (1 [ei lainkaan kuormitusta] – 100 [erittäin paljon kuormitusta])
7. Burnout = henkilön sairastuminen työuupumukseen vuoden aikana (kyllä/ei)

KH	Sukupuoli	Ikä	Väsymys	Unenpuute	Kuormituksen-kokemus	Burnout
1	mies	34	7	89	5	kyllä
2	mies	35	6	77	4	kyllä
3	mies	33	8	87	5	kyllä
4	mies	29	9	92	5	kyllä
5	mies	45	3	66	2	ei
6	mies	43	4	45	3	ei
7	mies	33	3	30	4	ei
8	mies	36	2	31	1	ei
9	mies	37	3	53	3	ei
10	mies	35	4	43	4	ei
11	mies	33	2	34	3	ei
12	mies	37	1	20	2	ei
13	mies	37	2	25	1	ei
14	mies	36	3	28	2	ei
15	mies	28	4	35	4	ei
16	nainen	29	7	89	5	kyllä
17	nainen	30	8	77	5	kyllä
18	nainen	31	9	87	3	kyllä
19	nainen	31	8	78	4	kyllä
20	nainen	26	7	67	4	kyllä
21	nainen	41	6	70	5	kyllä
22	nainen	42	2	21	2	ei
23	nainen	43	3	26	1	ei
24	nainen	44	4	33	2	ei
25	nainen	43	5	45	1	ei
26	nainen	33	2	10	1	ei
27	nainen	38	1	15	2	ei
28	nainen	37	2	23	2	ei
29	nainen	36	3	25	3	ei
30	nainen	35	4	36	1	ei

Liite I2: Prosessointikontekstin vaikutus luetun ymmärtämiseen

Konteksti, jossa erilaiset asiat painetaan mieleen, vaikuttaa siihen, kuinka asiat kyetään palauttamaan mieleen (Bransford & Johnson, 1972). Ilman kontekstia uuden informaation koodaaminen mieleen on hyvin vaikeaa, koska sitä ei voi yhdistää mihinkään olemassa olevaan tietorakenteeseen. Tämä on osoitettu kokeella, jossa tutkittavien tehtävänä oli lukea lyhyt pyykin pesemistä kuvaava tekstinpätkä. Tekstissä ei kuitenkaan sinällään viitattu pyykinpesuun millään tavalla, ja tapahtumakuvaus oli tarkoituksellisesti epämääräinen.

Tutkimuksessa oli kaksi koeryhmää ja yksi kontrolliryhmä. Toiselle koeryhmistä annettiin prosessointikonteksti ennen tekstin lukemista, ts. heille kerrottiin, että teksti käsittelee pyykin pesemistä. Toiselle ryhmälle prosessointikonteksti annettiin tekstin lukemisen jälkeen, ja kontrolliryhmälle prosessointikontekstia ei annettu ollenkaan. Tutkittavat saivat lukea tekstin omaan tahtiin, ja tekstin lukemisen jälkeen heiltä testattiin luetun tekstin ymmärtäminen sekä luetun tekstin sisällön muistaminen.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. Sukupuoli = tutkittavan sukupuoli
3. Ikä = tutkittavan ikä vuosina
4. Konteksti = milloin prosessointikonteksti on annettu tutkittavalle (ennen tekstin lukemista / ei kontekstia / tekstin lukemisen jälkeen)
5. Ymmärrys = tekstin ymmärtäminen (1–10)
6. Palautus = tutkittavan vapaa palautus tekstistä
7. Koulutus = tutkittavan koulutus (perusaste/keskiaste/korkea-aste)

KH	Sukupuoli	Ikä	Koulutus	Konteksti	Ymmärrys	Palautus
1	nainen	22	korkea-aste	konteksti ennen	8,00	8,00
2	nainen	24	korkea-aste	konteksti ennen	5,00	6,00
3	mies	19	keskiaste	konteksti ennen	8,00	10,00
4	mies	19	keskiaste	konteksti ennen	8,00	10,00
5	nainen	17	keskiaste	konteksti ennen	7,00	7,00
6	nainen	22	korkea-aste	konteksti ennen	6,00	6,00
7	mies	22	korkea-aste	konteksti ennen	7,00	8,00
8	nainen	25	korkea-aste	konteksti ennen	6,00	7,00
9	nainen	24	keskiaste	konteksti ennen	7,00	7,00
10	mies	17	keskiaste	konteksti ennen	8,00	8,50
11	nainen	27	keskiaste	ei kontekstia	6,00	5,50
12	mies	20	keskiaste	ei kontekstia	3,00	2,00
13	mies	19	keskiaste	ei kontekstia	6,00	5,00
14	mies	20	korkea-aste	ei kontekstia	1,00	3,00
15	mies	22	korkea-aste	ei kontekstia	3,00	4,00
16	nainen	19	keskiaste	ei kontekstia	2,00	4,00
17	mies	17	keskiaste	ei kontekstia	3,00	2,00
18	mies	24	korkea-aste	ei kontekstia	4,00	7,00
19	nainen	19	korkea-aste	ei kontekstia	3,00	3,00
20	mies	20	korkea-aste	ei kontekstia	2,00	2,00
21	mies	20	keskiaste	konteksti jälkeen	5,00	6,50
22	nainen	22	keskiaste	konteksti jälkeen	4,00	5,00
23	mies	25	keskiaste	konteksti jälkeen	3,00	4,00
24	nainen	24	korkea-aste	konteksti jälkeen	4,00	5,00
25	mies	22	korkea-aste	konteksti jälkeen	5,00	6,00
26	mies	22	korkea-aste	konteksti jälkeen	4,00	4,50
27	mies	26	keskiaste	konteksti jälkeen	3,00	2,00
28	mies	20	keskiaste	konteksti jälkeen	3,00	5,00
29	nainen	22	korkea-aste	konteksti jälkeen	4,00	3,00
30	mies	20	korkea-aste	konteksti jälkeen	3,00	4,00

Liite I3: Tunteiden säätelykysely

Tunne mekanismit ovat kehittyneet evoluution myötä muokkaamaan nopeasti ihmisten ja eläinten psykologista ja fysiologista tilaa siten, että se vastaa ympäristössä havaittuja uhkia ja mahdollisuuksia. Vaikka tunteet ovatkin useimmiten hyödyllisiä, ne voivat joskus olla ristiriidassa yksilön tavoitteiden kanssa. Esimerkiksi tenttitilanteessa koettu ahdistuneisuus ei välttämättä paranna tenttisuoritusta, erityisesti jos ahdistuksen kokemus on hyvin voimakas. Ihmiset voivat säädellä ei-toivottuja tunnereaktioitaan hyvin monella tavalla. Viimeaikaisissa tutkimuksissa on kuitenkin havaittu, että kaksi tunteiden säätelystrategiaa ovat hyvin yleisiä: tunteen aiheuttaman tilanteen uudelleenarviointi sekä tunneilmauksen (esimerkiksi kasvonilmeiden) peittämiseen perustuva ekspressiivinen suppressio. On havaittu, että uudelleenarviointiin perustuva strategia on tehokkaampi säätelämään tunnereaktion koettua voimakkuutta, ja lisäksi sen käyttäminen kuormittaa esimerkiksi kardiovaskulaarista järjestelmää vähemmän kuin ekspressiivinen suppressio. James Gross työryhmineen on kehittänyt Emotion Regulation Questionnairen (ERQ), jonka avulla voidaan mitata, kuinka paljon henkilö käyttää näitä kahta säätelystrategiaa arkielämässään. Tutkittava vastaa kuinka hyvin erilaiset uudelleenarviointiin (esimerkiksi ”Säätelen tunteitani siten, että koitan ajatella tilanteita toisella tavalla”) ja ekspressiivisen suppressioon (esimerkiksi ”Säätelen tunteitani siten, että en näytä niitä”) liittyvät väitteet sopivat häneen. Vastaamisessa käytetään asteikkoa 1 (täysin eri mieltä) – 7 (täysin samaa mieltä). Tässä aineistossa on esitetty ERQ-lomakkeen tulokset kahdeltakymmeneltä lomakkeen täyttäneeltä henkilöltä.

Muuttujaluettelo

1. KH = tutkittavan juokseva järjestysnumero
2. U1–U6 = uudelleenarviointiin perustuvaa säätelystrategiaa mittaavat väittämät
3. S1–S4 = ekspressiiviseen suppressioon perustuvaa säätelystrategiaa mittaavat väittämät
4. USUM = uudelleenarviointia mittaavien väittämien summapistemäärä
5. SSUM = ekspressiivistä suppressiota mittaavien väittämien summapistemäärä

KH	U1	U2	U3	U4	U5	U6	S1	S2	S3	S4	USUM	SSUM
1	1	6	6	5	5	5	2	2	3	4	28	11
2	3	3	4	4	4	4	2	1	2	2	22	7
3	5	5	4	5	5	5	5	3	5	3	29	16
4	5	5	5	3	3	6	2	2	2	2	27	8
5	6	6	5	6	5	6	5	2	4	4	34	15
6	5	3	6	5	5	5	2	1	2	1	29	6
7	2	2	7	5	5	4	2	1	4	2	25	9
8	2	1	7	2	2	1	2	2	1	3	15	8
9	6	5	6	4	2	5	5	3	5	4	28	17
10	7	3	6	6	2	5	2	1	1	6	29	10
11	2	2	2	3	2	2	2	1	1	2	13	6
12	4	4	3	2	2	3	1	1	1	3	18	6
13	5	5	6	3	5	3	6	3	5	3	27	17
14	6	6	6	4	4	6	2	1	3	5	32	11
15	3	3	4	3	3	3	6	2	2	2	19	12
16	5	6	6	3	4	6	1	2	1	1	30	5
17	6	2	2	4	2	3	1	1	2	2	19	6
18	2	1	4	2	4	5	6	3	5	5	18	19
19	1	6	6	5	5	5	2	2	3	4	28	11
20	3	3	4	4	4	4	2	1	2	2	22	7