Diskreettien osiovastausten monipiirreanalyysi yleistetyllä Raschin mallilla

Tilastotieteen pro gradu -tutkielma

Vesa Kiviniemi
Jyväskylän yliopisto
Tilastotieteen laitos
Syyskuu 2000

Tiivistelmä

Vesa Kiviniemi: Diskreettien osiovastausten monipiirreanalyysi yleistetyllä Raschin mallilla.
Tilastotieteen pro gradu -tutkielma, Jyväskylän yliopisto, syyskuu 2000.
Sivuja 72.

Tutkielmassa tarkastellaan teoreettisesti diskreettien osiovastausten analyysia yleistetyllä Raschin mallilla, johon sisältyen tutkitaan yleistetyn Raschin osioanalyysimallin rakennetta, erilaisten osioanalyysimallien spesifiointia yleistetystä Raschin mallista sekä mallin parametrien estimointia. Lisäksi keskeisinä tarkastelun kohteena ovat mallin oppilasparametrien ennustamismenetelmä ja yleistetyn Raschin mallin osioparametrien yhteensopivuustestit.

Tutkielman alussa käydään yleistettyyn Raschin malliin johdattaen läpi paljon koulutustutkimuksessa käytettävän osioanalyysin peruskäsitteistö, Raschin malli, mallin parametrien estimointimenetelmät, sekä iteratiiviset algoritmit ja Raschin mallin lisäparametriset vaihtoehdot.

Perinteisten osioanalyysimallien, kuten Raschin mallin, ongelmana on kuitenkin niiden vaatima dikotominen mitta-asteikko ja oletus osiokokeella oppilailta mitattavan latentin piirteen yksiulotteisuudesta. Ratkaisuesitys näihin osioanalyysin ongelmallisuuksiin on yleistetty Raschin malli, jota tutkielmassa keskeisimmin tarkastellaan. Yleistetyn Raschin mallin yhteydessä käsitellään mallin rakenne, sekä havainnollistetaan mallin asetelmamatriisin ja pisteytysmatriisin käyttöä erilaisten osioanalyysimallien spesifioinnissa. Keskeisenä tarkastelun kohteena on myös yleistetyn Raschin mallin parametrien estimointi suurimman uskottavuuden menetelmällä, joka yleistetyn Raschin mallin tilanteessa edellyttää myös EM-algoritmin modifikaation ja numeeristen integrointimenetelmien tarkastelua. Lisäksi estimoinnin yhteydessä käydään läpi odotettujen aposteriori estimaattien laskeminen oppilasparametreille ja edelliselle vaihtoehtoisena tekniikkana rajoitettu suurimman uskottavuuden estimointi mallin oppilasparametreille. Tutkielman lopussa on esitelty yleistetyn Raschin mallin osioparametrien yhteensopivuuden tutkiminen.

Tutkielman keskeisenä tarkoituksena on selvittää osioanalyysin, Raschin mallin ja erityisesti yleistetyn Raschin mallin parissa käytettäviä käsitteitä, sekä niihin liittyvien tilastollisten tekniikkojen ja erilaisten numeeristen menetelmien teoriaa.

Avainsanat: osioanalyysi, Raschin malli, yleistetty Raschin malli, suurimman uskottavuuden menetelmä, marginaaliposteriori, numeerinen integrointi, EM-algoritmi

Sisältö

1 Johdanto **1**

2 Raschin malli osioanalyysiaineistolle **4**

2.1 Latentin piirteen esiinsaaminen Raschin mallilla 5

2.1.1 Oletus latentin piirteen yksiulotteisuudesta 5

2.1.2 Oletus osioiden ominaiskäyristä 6

2.1.3 Oletus osioiden riippumattomuudesta 9

2.2 Raschin ominaisfunktio 10

2.3 Mallin identifioituvuus 14

2.4 Osioparametrien tarkentuva estimointi 15

2.5 Oppilasparametrien tarkentuva estimointi 18

2.6 Iteratiiviset laskenta-algoritmit 20

2.6.1 Newton-Raphson 21

2.6.2 Fisherin scoring-menetelmä 22

2.7 Lisäparametrisia vaihtoehtoja Raschin mallille 23

2.8 Yhteenveto Raschin mallista 25

3 Yleistetty Raschin malli **27**

3.1 Yleistetyn Raschin mallin ominaisfunktio 28

3.2 Mallin asetelmamatriisin ja pisteytysmatriisin käyttö erilaisten mallien spesifioinnissa 31

3.2.1 Raschin mallin spesifioiminen yleistetystä Raschin mallista 32

3.2.2 Polytomisesti pisteytettyjen osiovastausten mallintaminen Yleistetyllä Raschin mallilla 34

3.2.3 Moniulotteisen Raschin mallin konstruointi yleistetystä Raschin mallista 37

3.3 Populaatiomalli 39

3.4 Yleistetyn Raschin mallin identifioituvuusehdot 41

3.5 Yhteenveto yleistetystä Raschin mallista 43

4 Yleistetyn Raschin mallin parametrien estimointi ja kykyparametrien ennustaminen **44**
4.1 Suurimman uskottavuuden menetelmä 45
4.2 Integraalien approksimointi 50
4.3 EM-algoritmi 52
4.4 Parametrien estimaattien keskivirhetarkastelut 56
4.5 Henkilön kykyparametrin estimointi ja ennustaminen 60
4.6 Yhteenveto estimoinnista 62

5 Mallin osioparametrien yhteensopivuus **63**

6 Pohdinta **66**

7 Kiitosmaininnat **69**

Lähteet

Luku 1

Johdanto

Tutkielmassa tarkastellaan yleistettyyn Raschin osioanalyysimalliin sisältyvää teoriaa. Keskeisimpinä tutkimusongelmina ovat yleistetyn Raschin mallin rakenne, erilaisten mallivaihtoehtojen spesifiointi yleistetystä Raschin mallista ja mallin parametrien estimointi, jossa tarvittavat tekniikat ovat suurimman uskottavuuden menetelmä, EM-algoritmi, numeerinen integrointi ja iteratiiviset laskentaalgoritmit. Tutkielman alussa tarkastellaan myös perinteistä Raschin mallia ja sen lisäparametrisia vaihtoehtoja, joiden tarkastelu helpottaa yleistetyn Raschin mallin teoreettista ymmärtämistä.

Erilaisia osioanalyysimalleja käytetään yleisesti koulutustutkimuksissa, joissa tutkimusaineisto kerätään testaamalla oppilaiden kognitiivista suorituskykyä tai osaamista monivalintatehtävin eli osioin. Traditionaalisemmat osioanalyysimallit kuten yksi-, kaksi- ja kolmiparametriset logistiset mallit (Rasch, 1960; Andersen, 1980) ovat soveltuvia dikotomisesti koodattuun tutkimusaineistoon ja niillä voidaan mitata yhtä latenttia eli ei-havaittavissa olevaa piirrettä koehenkilöiltä. Täten ne ovat yksiulotteisia latentin piirteen osioanalyysimalleja. Kyseisillä osioanalyysimalleilla voidaan mallintaa osioiden vaikeustaso ja oppilaiden kykytaso dikotomisen mitta-asteikon tilanteessa, jolloin mitattava latentti piirre oletetaan yksiulotteiseksi.

Osioanalyysimallien ongelmana on, että niiden vaatima dikotominen mitta-asteikko voi joissain tilanteissa olla suhteellisen epäinformatiivinen. Lisäksi tutkimusaineistosta mallinnettava latentti struktuuri ei välttämättä ole yksiulotteinen, vaan todellisuudessa osiot voivat mitata oppilailta useita ominaisuuksia, joita traditionaalisilla osioanalyysin malleilla ei ole mahdollista mallintaa.

Ratkaisuesitys kompleksisemman latentin struktuurin mallintamiseen on yleistetyn Raschin mallin sovittaminen aineistoon, jolloin osiovastaukset voivat olla myös polytomisesti koodattuja moniulotteisen latentin struktuurin lisäksi. Käytännössä yleistetyllä Raschin mallilla oppilaiden osiovastauksista on mahdollista mallintaa useita eri mitattavia ominaisuuksia. Mitta-asteikon ollessa polytominen, ominaisuuksien mittaukseen saadaan lisää tarkkuutta.

Yleistetyn Raschin mallin keskeiset komponentit ovat mallin asetelmamatriisi ja vastausten pisteytysmatriisi, joiden sopivilla valinnoilla on mahdollista spesifioida lukuisia jo latentin piirteen teorian parissa tunnettuja malleja. Yksiulotteisista malleista saadaan esimerkiksi Raschin osioanalyysimalli (Rasch, 1960), moniulotteinen Raschin malli (Kelderman & Rijkes, 1994) ja polytomisesti koodattujen vastausten osioanalyysimalli (Masters, 1982). Mallin lisäominaisuutena on siihen yhdistettävä populaatiomalli eli latentin muuttujan tiheysfunktio. Populaatiomallin avulla malli voidaan johtaa ei-ehdolliseksi marginaaliosioanalyysimalliksi, josta johdetaan marginaaliuskottavuusfunktio. Marginaaliuskottavuusfunktiosta johdetaan suurimman uskottavuuden menetelmin parametrien estimaattorit. Parametrien numeeriset estimaatit lasketaan kaksivaiheisen EM-algoritmin (Dempster, Laird & Rubin, 1977) modifikaatiolla (Bock & Aitkin, 1981), jolloin estimoinnin tuloksina ovat osioita kuvaavien parametrien sekä populaatiomallin parametrien estimaatit. Populaatiomallin eli latentin muuttujan jakauman keskiarvo voidaan myös korvata lineaarisella regressiomallilla osioanalyysiaineistosta, joka mahdollistaa esimerkiksi sukupuolten välisten kognitiivisten tasoerojen suoran mallintamisen osioaineistosta.

Marginaaliosioanalyysimalli ei kuitenkaan sisällä oppilaskohtaisia kykyä kuvaavia parametreja, sillä ne integroidaan marginaaliosioanalyysimallin johtamisen yhteydessä siitä pois. Tällöin oppilaskohtaiset parametrit ennustetaan oppilaan posteriorikeskiarvosta tai ratkaistaan SU-menetelmällä osioparametrien ollessa tunnetut.

Tutkielman luvussa 2 esitellään latentin piirteen teoriassa jo hyvin tunnettu Raschin malli, kuvataan mallin osioparametrien tarkentuva estimointi ehdollisen suurimman uskottavuuden menetelmällä ja oppilasparametrien tarkentuva estimointi marginaaliuskottavuusmenetelmällä. Lisäksi luvussa 2 kuvataan iteratiiviset algoritmit parametrien numeeriseksi ratkaisemiseksi, joita ovat Newton-Raphson laskenta-algoritmi ja Fisherin Scoring menetelmä. Luvun 2 lopussa kuvataan lyhyesti lisäparametrisia vaihtoehtoja Raschin mallille, joita ovat kaksi- ja kolmiparametrinen osioanalyysimalli. Luvussa 3 on esitellään yleistetty Raschin malli. Malli esitellään aluksi yleisessä muodossa, jonka jälkeen mallin asetelmamatriisin ja pisteytysmatriisin käyttöä havainnollistetaan spesifioimalla erilaisia malleja mallin yleisestä muodosta. Näiden yhteydessä on kuvattu tilanteita, missä kunkin spesifioidun mallin käyttö on aiheellista oppilaiden osiovastausten mallintamiseksi. Luvussa 3 esitellään myös populaatiomalli ja sen käyttötarkoitus yleistetyssä Raschin mallissa. Luvun 3 lopussa esitetään yleistetyn Raschin mallin identifioituvuusehdot. Luvussa 4 on kuvattu parametrien estimointi. Luvussa tarkastellaan SU-menetelmää mallin parametrien estimaattorien johtamiseksi, sekä niiden numeeristen estimaattien laskemista EM-algoritmilla. Luvussa 4 on kuvattu myös parametrien keskivirhetarkastelut ja latentti ennustaminen oppilaiden kykyparametrien arvioimiseksi oppilaan posteriorikeskiarvosta. Luvussa 5 on kuvattu mallin osioparametrien yhteensopivuustarkastelut. Tutkielman luvussa 6 on yhteenvedon merkeissä pohdittu tutkielmaa.

Yleisesti tutkielman lähestymistapa on täysin teoreettinen ja tutkielman tavoitteena on havainnollistaa osioanalyysin, Raschin mallin ja erityisesti yleistetyn Raschin mallin teoriaa.

Luku 2

Raschin malli osioanalyysiaineistolle

Georg Raschin (1960) ehdottama yksiparametrinen logistinen osioanalyysimalli on yksinkertaisin kaikista osioanalyysimalleista ja sitä käytetään yhä koulutuksen ja käyttäytymistieteiden tutkimuksessa. Raschin mallilla saadaan estimoiduksi dikotomisesti koodatuista osiovastauksista osiokohtaiset parametrit, jotka kuvaavat osion vaikeustasoa, sekä oppilasparametrit (yleisemmin henkilöparametrit), jotka kuvaavat oppilaiden osaamista tai yleisemmin sitä, kuinka paljon kullakin oppilaalla on mitattavaa latenttia piirrettä. Raschin mallissa suurimman uskottavuuden estimaatit eivät ole tarkentuvia eli aineiston kasvattaminen ei johda estimaattien tarkkuuden paranemiseen. Tällöin mallin osioparametrit ratkaistaan ehdollisen suurimman uskottavuuden menetelmällä ja oppilasparametrit marginaaliuskottavuusmenetelmällä.

Tässä luvussa on havainnollistettu yhden latentin piirteen esiin saamista Raschin mallilla, joka edellyttää kolmen oletuksen tekemistä. Lisäksi luvussa kuvataan ehdollisen suurimman uskottavuuden menetelmää parametrien estimoimiseksi sekä iteratiivisia algoritmeja estimaattien numeerisen ratkaisun tuottamiseksi. Luvun lopussa on esitetty Raschin mallin lisäparametrisia vaihtoehtoja, joita ovat kaksi- ja kolmiparametrinen osioanalyysimalli. Yleisesti luvussa käydään läpi latentin piirteen teorian peruskäsitteistöä ja osioanalyysin perusperiaatteita, jotka on välttämättömiä ymmärtää ennen yleistetyn Raschin mallin täydellistä ymmärtämistä.

2.1 Latentin piirteen esiinsaaminen Raschin mallilla

Osiovastauksien analysointi eli oppilaiden osaamistason ja osion ominaisuuksien esiin saaminen ei ole mahdollista aivan suoraan, vaan osioaineistosta tarvitaan joitakin tilannetta selventäviä ja yksinkertaistavia oletuksia (ks. esim. Gustafsson, 1977; Gustafsson, 1979; Gustafsson, 1981).

Raschin malliin, kuten myös muihin perinteisiin osioanalyysimalleihin liittyy kolme kaikille malleille yhteistä oletusta, jotka ne toteuttavat. Nämä ovat Hambletonin & Swaminathanin (1985) mukaan seuraavat:

1) *Oletus latentin piirteen yksiulotteisuudesta* eli oletetaan, että osiovastauksilla mitataan ainoastaan yhtä latenttia piirrettä, joka on oppilaan kyvykkyys osiokokeella mitattavassa asiassa, esimerkiksi matematiikassa.

2) Malliperusteisesta lähestymistavasta johtuva *oletus osioiden ominaiskäyrien muodosta.*

3) *Oletus osioiden riippumattomuudesta.* Näin ollen tietyn henkilön vastaukset eri osiokokeen eri osioihin oletetaan riippumattomiksi toisistaan.

2.1.1 Oletus latentin piirteen yksiulotteisuudesta

Dikotomisen mitta-asteikon (0, jos vastaus väärin; 1, jos vastaus oikein) ohella logististen osioanalyysimallien ehkä tärkein oletus on, että oppilaiden osiovastausten oletetaan riippuvan systemaattisesti vain ja ainoastaan yhdestä tekijästä, eli heidän kyvystään tuottaa oikeita osiovastauksia osiokokeella mitattavassa asiassa.

Latenttia ominaisuutta tai piirrettä vastaavaa muuttujaa nimitetään piilomuuttujaksi ja sen vastaavia arvoja piirrearvoiksi, jotka kertovat, kuinka paljon oppilaalla on kykyä tuottaa oikeita vastauksia eri osioihin. Merkitään tätä piirrearvoa eli oppilasparametria thetalla (θ). Piirre oletetaan jatkuvaksi ja vähintään välimatka-asteikolliseksi, mutta jakaumaoletusta siitä ei sen sijaan Raschin mallissa tehdä (Gustafsson, 1977).

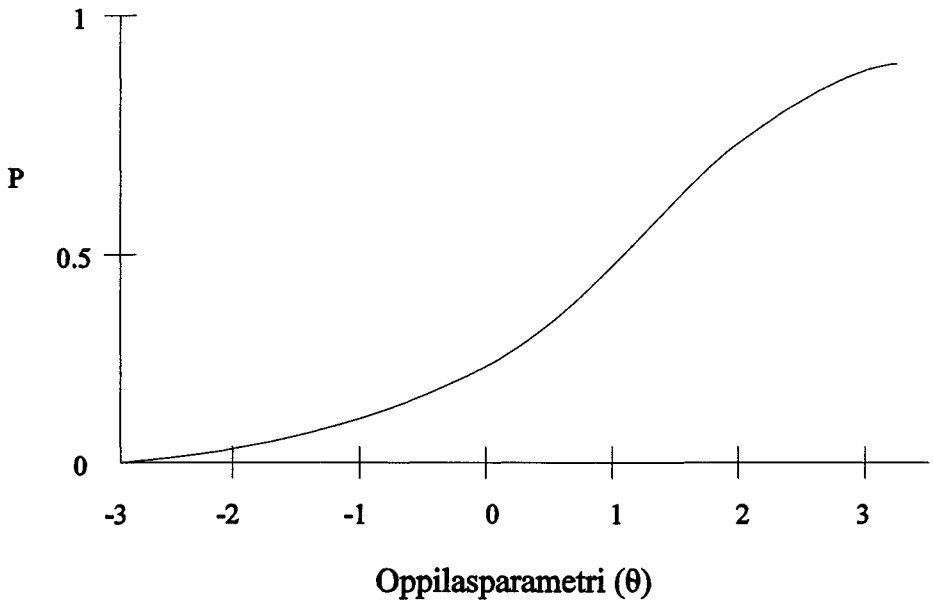
Osioanalyysimallit saavat todennäköisyyslaskennallisen luonteensa siitä, miten osiovastausten oletetaan riippuvan latentista piirteestä. Jonkin piirteen arvon omaavat oppilaat voivat tosin tuottaa erilaisia osiovastauksia, mutta jokaisen erilaisen osiovastauksen todennäköisyyden oletetaan riippuvan kyseisestä piirrearvosta eli oppilasparametrin arvosta. Näin malli määrittelee todennäköisyyden vastata tiettyyn osioon oikein jokaisella eri oppilaan piirrearvolla.

2.1.2 Oletus osioiden ominaiskäyristä

Malliperusteisesta lähestymistavasta johtuen osiovastausten todennäköisyyksien oletetaan riippuvan latentista piirteestä funktionaalisesti osion ominaisfunktion olettamalla tavalla. Raschin mallissa formaalisti

$$P(X = x|\theta, \alpha), \tag{2.1.2.1}$$

missä θ kuvaa henkilön piirrearvoa ja α osiota, osion ja henkilön ollessa kiinteitää. Osion ominaisfunktion kuvaa oikean osiovastauksen todennäköisyyttä oppilasparametrin funktiona. Sen kuvaaja on osion ominaiskäyrä ja sitä merkitään g(θ - α):lla, missä α on tiettyä osiota eli osion vaikeustasoa kuvaava parametri ja θ oppilasparametri. Osion ominaiskäyrän sijaintiin vaikuttaa suoraan osion vaikeustasoparametrin α arvo, joka siirtää käyrää origon suhteen. Seuraavalla sivulla on havainnollistettu osion ominaiskäyrän kuvaajaa yhden osion tilanteessa.



Kuva 2.1. Osion ominaiskäyrä

Yksiparametrinen logistinen osioanalyysimalli on dikotominen malli, joka mallintaa todennäköisyydet vastata osioon oikein (1) tai väärin (0). Näin ollen niiden on oltava suurempia tai yhtäsuuria kuin 0 (ks. esim. Holland, 1990), formaalisti

$$P(X = 1|\theta,\alpha) \geq 0 \text{ ja } P(X = 0|\theta,\alpha) \geq 0. \qquad (2.1.2.2)$$

Tässä kontekstissa ei tarkastella testikysymyksiä, joihin oppilaat eivät mahdollisesti ole vastanneet, joten oletuksena on, että oppilas tuottaa varmasti oikean tai väärän vastauksen, jolloin seuraa, että

$$P(X = 1|\theta,\alpha) + P(X = 0|\theta,\alpha) = 1. \qquad (2.1.2.3)$$

Raschin mallissa kiinnostus kohdistuu ainoastaan oikeisiin vastauksiin ($X = 1 | \theta, \alpha$) eli piirrettä tai ominaisuutta indikoiviin osiovastauksiin, jolloin osion ominaisfunktio voidaan formaalisti kirjoittaa kuten (2.1.2.1). Osion ominaisfunktio kuvaa suoraan, miten oppilaan todennäköisyys ratkaista osio riippuu latentista muuttujasta. Riippuvuus on funktionaalista eli jokaista eri oppilasparametrin θ arvoa vastaa tietty todennäköisyyden arvo. Osion ominaisfunktio oletetaan myös samaksi kaikilla samaa latenttia piirrettä mittaavilla osioilla (Gustafsson, 1977).

Meneteltäessä edellä esitellyn rajauksen mukaisesti voidaan valita mikä tahansa jatkuva funktio, jonka arvot vaihtelevat välillä [0,1]. Oppimistulosten mittaus on eräänlainen "kykytesti", joten tässä kontekstissa on syytä olettaa funktio kasvavaksi ja monotoniseksi. Tällaisen funktion valinnasta riippuen on mahdollista muodostaa erilaisia osioanalyysimalleja, joita voi periaatteessa kehittää rajattomasti (Hambleton & Swaminathan, 1985, 52). Mallien rajattomasta kehittämismahdollisuudesta myös hyvä esimerkki on tässä tutkielmassa havainnollistettava yleistetty Raschin malli. Hyvin usein koulutustutkimuksessa mallin valinnassa on kuitenkin rajoituttu joko standardoidun normaalijakauman kertymäfunktioon

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-u^2}{2}} du \qquad (2.1.2.5)$$

tai logistisen jakauman kertymäfunktioon, joka on muotoa

$$\psi(x) = \frac{e^x}{1+e^x}. \qquad (2.1.2.6)$$

Näissä on yleisimmin on yhdestä kolmeen osiokohtaista parametria (Hambleton & Swaminathan, 1985). Raschin mallihan on johdettu logistisesta kertymäfunktiosta ja siinä on ainoastaan yksi osiota kuvaava parametri (Gustafsson, 1977).

2.1.3 Oletus osioiden riippumattomuudesta

Oletetaan, että θ on kiinteä ja osiokokeen osioita on k kappaletta, $j = 1,...,k$. Tällöin osiovastaukset ovat toisistaan ehdollisesti riippumattomia, ehdolla, että θ ja α tunnetaan, jos

$$\begin{aligned} &P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k | \theta, \alpha) \\ &= P(X_1 = x_1 | \theta, \alpha) P(X_2 = x_2 | \theta, \alpha) \ldots P(X_k = x_k | \theta, \alpha). \end{aligned} \tag{2.1.3.1}$$

Tällöin marginaalitodennäköisyydet määräävät yksikäsitteisesti koko osiokokeen todennäköisyyden. Näin esimerkiksi todennäköisyys, että vastaaja saa kaksi osiota oikein on sama, kuin yksittäisten osioiden ratkaisutodennäköisyyksien tulo.

Edellä esitetystä myös seuraa, että osioilla ei oleteta myöskään olevan mitään järjestystä eli kontekstiefektiä. Näin osiovastausten oletetaan pysyvän samoina riippumatta siitä, missä järjestyksessä tai missä yhteydessä ne esitetään. Osioiden vaihdannaisuus on kuitenkin riippumattomuutta lievempi oletus.

Tämän tutkielman kontekstissa on kuitenkin syytä muistaa, että edellä esitetyt kolme oletusta koskevat nimenomaan Raschin mallia ja sen tässä luvussa esiteltäviä lisäparametrisia vaihtoehtoja. Oletukset ovat täysin siis samat myös Raschin mallin lisäparametrisissa vaihtoehdoissa, ainoana erona on, että näissä lisäehtona on, että myös mallien ominaiskäyrien muut osiota kuvaavat parametrit tunnetaan.

Yleistetyssä Raschin mallissa oletusta latentin piirteen yksiulotteisuudesta ei tarvitse tehdä. Ominaiskäyrän muotokaan ei ole ilmaistavissa x-y-tasona, jos kykyä kuvaava parametri θ oletetaan moniulotteiseksi, jolloin se on vektoriarvoinen. Tällöin osiovastauksista mallinnetaan useata eri piirrettä, esimerkiksi matematiikan ja kemian osaamista, jolloin kykyä kuvaava parametri θ olisi kaksiulotteinen. Tällaisessa tilanteessa osion ominaiskäyrä on pinta (kutsumme jatkossa moniulotteista mallia myös käyräksi, vaikka se olisi piirteiden lukumäärän dimensioita vastaava pinta). Jos latentin piirteen ulottuvuuksia on paljon, niin sen hahmottaminen ja piirtäminen on hankalaa. Matemaattisesti osion ominaiskäyrää moniulotteisessa tilanteessa voidaan kuvata piirreavaruudessa kulkevana vektorina.

2.2 Raschin ominaisfunktio

Indeksoidaan nyt oppilasta i:llä, $i = 1,...N$ ja osiota j:llä, $j = 1,...,k$. Tällöin Raschin mallissa todennäköisyys

$$P(X_{ij} = x_{ij} | \theta_i, \alpha_j) \tag{2.2.1}$$

oletetaan erotuksen (θ_i - α_j) funktioksi, jossa θ_i on oppilasparametri ja α_j on osiokohtainen parametri (ks. Andersen, 1980). Tämä kyseisen erotuksen funktio on osion ominaisfunktio, jonka arvot vaihtelevat nollan ja ykkösen välillä, koska ominaisfunktion arvot ovat todennäköisyyksiä. Ominaisfunktion oletetaan saavan arvon 0 kun oppilasparametrin arvo lähestyy - ∞:tä ja vastaavasti arvon 1 kun oppilasparametrin arvo lähestyy + ∞:tä. Yksinkertaisin muoto tällaisesta funktiosta on yksiparametrinen logistinen osioanalyysimalli, joka on Raschin malli.

Ominaisfunktio on mahdollista johtaa itse tai se voidaan valita erilaisista valmiiksi kehitetyistä ominaisfunktioista. Tässä kontekstissa havainnollistettava malli on Raschin (1960) kehittämä. Perusteet, joilla Rasch johti mallinsa on kuvattu hänen tunnetuksi tulleessa artikkelissaan (1960).

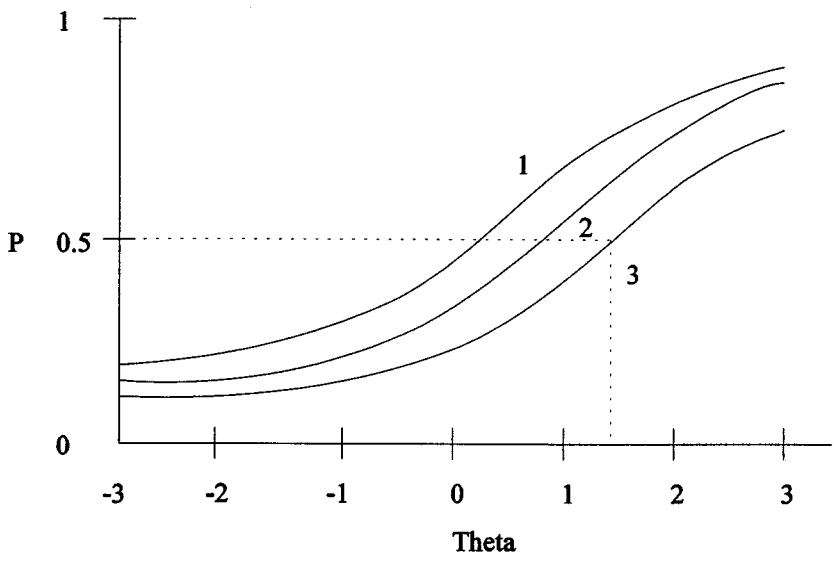
Raschin mallissa kiinnostus suuntautuu yleensä ainoastaan oikeisiin vastauksiin ($X_{ij} = 1$). Tällöin Andersenin (1980) mukaisesti logistisen kertymäfunktion kaavasta (2.1.2.6) voidaan kirjoittaa logistinen jakaumamalli (logistic ogive model), joka osiovastauksen ollessa oikea on formaalisti muotoa

$$P(X_{ij} = 1) = \psi(\theta_i - \alpha_j). \tag{2.2.2}$$

Tällöin Raschin mallin ominaisfunktio eli Raschin malli kirjoitetaan eksplisiittisesti muodossa

$$P(X_{ij} = 1|\theta_i, \alpha_j) = \frac{e^{(\theta_i - \alpha_j)}}{1 + e^{(\theta_i - \alpha_j)}} = \frac{1}{1 + e^{(\alpha_j - \theta_i)}}. \qquad (2.2.3)$$

Tällaisen ominaisfunktioiden kuvaajat eli osion ominaiskäyrät ovat samanmuotoiset ja ainoastaan niiden sijainti vaihtelee. Sijaintiin vaikuttaa osion vaikeustaso eli osiokohtainen parametri α_j. Seuraavassa on esitetty osioiden ominaisfunktioista esimerkkikuva Gustafssonin (1977) mukaisesti.



Kuva 2.2: Esimerkki kolmen eri osion ominaiskäyristä Raschin mallissa

Näin Raschin mallissa, jossa osiota kuvataan yhdellä parametrilla, jokaisen osion ominaiskäyrän jyrkkyys on sama. Ainoastaan niiden sijainti vaihtelee

origon suhteen. Ominaiskäyrän siirtymään origon suhteen vaikuttaa ainoastaan osiokohtainen parametri α_j, joka kuvaa siis osion vaikeustasoa.

Nyt kuvasta 2.2 voidaan havaita, että osion 1 saa oikein todennäköisyydellä 0.5 piirrearvon 0.5 omaava oppilas. Sen sijaan osion 3 ratkaisemiseksi todennäköisyydellä 0.5 vaaditaan jo piirrearvoa 1.5. Näin siis osio 3 on vaikeampi, kuin osio 1 tai osio 2. Tästä syystä α_j - parametria nimitetään osion vaikeustasoparametriksi. Vaikeustasoparametri kertoo sen oppilasparametrin arvon, jota pienemmillä arvoilla oppilas ei todennäköisesti ratkaise osiota, jolloin ratkaisutodennäköisyys $P < 0.5$. Ja vastaavasti osion vaikeustasoparametria suuremmilla θ_i:n arvoilla oppilas ratkaisee monivalintatehtävän oikein, koska ratkaisutodennäköisyys $P > 0.5$. Raschin ominaisfunktiosta eli Raschin mallista voidaan laskea myös eri osioiden vaikeustasot tietyn θ_i:n kohdalla.

Raschin ominaisfunktiossa, joka on siis muotoa

$$P(X_{ij} = 1|\theta_i, \alpha_j) = \frac{e^{(\theta_i - \alpha_j)}}{1 + e^{(\theta_i - \alpha_j)}}, \tag{2.2.4}$$

α_j osoittaa myös funktion käännepisteen θ_i arvon. Käännepisteen ympäristössä θ_i:n muutos vaikuttaa suhteellisesti eniten ratkaisutodennäköisyyden P_{ij} arvoon. Tästä johtuen tietty osio erottelee oppilaita, kun oppilasparametri ja osion vaikeustasoparametri ovat yhtä suuria, $\theta_i = \alpha_j$ (Andersen, 1973a).

Raschin ominaisfunktio eli Raschin malli on käytännöllisesti katsoen kahden tekijän logit-malli, jossa tekijöinä ovat oppilaat ja osiot eli osiovastauksia selitetään oppilailla ja osioilla. Tämä voidaan osoittaa seuraavasti:

Yleisesti tunnetusti logistinen regressiomalli on muotoa

$$\operatorname{logit} \prod (x) = \alpha + \beta x, \tag{2.2.5}$$

missä

$\prod(x) = P(Y = 1|x)$, y on dikotominen havaittu muuttuja ja x voi olla mikä tahansa reaaliluku. Jos nyt tehdään logistinen muunnos Raschin mallille, jolloin

$$\text{logit}\, P(X_{ij} = 1|\theta_i, \alpha_j) = \ln \frac{P(X_{ij} = 1|\theta_i, \alpha_j)}{1 - P(X_{ij} = 1|\theta_i, \alpha_j)} = \ln \frac{P(X_{ij} = 1|\theta_i, \alpha_j)}{P(X_{ij} = 0|\theta_i, \alpha_j)}, \quad (2.2.6)$$

missä todennäköisyyksien suhde kuvaa vedonlyöntisuhdetta (odds ratio) parametrin ollessa θ_i. Kun tähän sijoitetaan (2.2.3), saadaan

$$\text{logit}\, P(X_{ij} = 1|\theta_i, \alpha_j) = \theta_i - \alpha_j, \quad (2.2.7)$$

jossa i:llä on indeksoitu oppilaita ja j:llä osioita. Ilmaisemalla Raschin malli lineaarisena lausekkeena selittävistä muuttujista edellä kuvatulla tavalla, havaitaan, että Raschin malli on kahden tekijän logit-malli. Malli yhtyy logistiseen regressiomalliin tilanteessa, jolloin β_j on parametrisoitu ykköseksi (Andersen, 1980). Näin Raschin mallissa osioiden erottelukyvyt oletetaan vakioksi. Kaksiparametrisessa logistisessa osioanalyysimallissa, josta keskustellaan myöhemmin tässä luvussa, β_j:tä ei kiinnitetä ykköseksi, vaan se vaihtelee osioittain. Kaksiparametrinen malli siis olettaa osioiden erottelukyvyn heterogeeniseksi, kun taas Raschin mallissa erottelukyky oletetaan homogeeniseksi. Tätä homogeenisuusoletusta kutsutaan Raschin homogeenisuudeksi.

Edellä kuvattu Raschin mallin yhteys logit-malliin verrantaa Raschin mallin ristiintaulukkomenetelmien eli loglineaaristen mallien kanssa, joita mm. Holland (1990) havainnollistaa artikkelissaan osioanalyysimalleista.

2.3 Mallin identifioituvuus

Mallin epäidentifioituvuudessa on kyse yliparametrisoinnista, joka tyypillisesti johtaa parametrien estimaattoreiden keskinäiseen korrelaatioon tai siihen, että estimointiyhtälöillä ei ole ratkaisua lainkaan. Tämä tarkoittaa sitä, että ainakin yhden parametrin estimaattori voidaan esittää muiden parametrien estimaattorien funktiona. Heuristisesti ajatellen mallin kompleksisuus on suurempi kuin aineisto sallii.

Raschin mallissa identifioituvuusongelma on seuraavanlainen: Raschin mallissa ollaan kiinnostuneita ainoastaan oppilaiden ja osioiden kokonaispistemäärien muodostamista marginaaleista eli rivi- ja sarakepistemääristä, koska ne ovat tyhjentäviä tunnuslukuja mallin parametreille. Nyt koska kiinnostus suuntautuu havaintomatriisin marginaaleihin, niin mallille jää parametreja $n + k$ kappaletta eli yhtä monta kuin osioita ja oppilaita on yhteensä. Kuitenkin havaintomatriisista yhteenlasketut rivi- ja sarakepistemäärät ovat yhtäsuuret, joten malli ei ole näin identifioituva, sillä siinä on yksi ulottuvuus liikaa ja malli ei ole tällaisenaan identifioituva.

Yleisesti ottaen tällainen identifioituvuusongelma voidaan ratkaista kiinnittämällä yksi parametreista nollaksi. Yleensä Raschin mallissa on ollut tapana kiinnittää yksi osioparametreista nollaksi, esim. $\alpha_1 = 0$, jolloin malli saadaan identifioituvaksi ja estimoitavien parametrien lukumäärä on $n+ k$-1 kappaletta (Andersen, 1980; Gustafsson, 1978).

Kaksiparametrisessa logistisessa osioanalyysimallissa mallista estimoitavien parametrien lukumäärä on $n + 2k$-2 ja kolmiparametrisessa mallissa $n+ 3k$-2, jolloin malleista saadaan identifioituvat.

2.4 Osioparametrien tarkentuva estimointi

Aiemmin 60- ja 70- luvulla Raschin mallia sovellettaessa psykologian ja kasvatustieteen tutkimuksissa on käytetty suurimman uskottavuuden menetelmää, jossa osio- ja oppilasparametrit estimoidaan samanaikaisesti. Menetelmällä on kuitenkin vakava puute, sillä osioparametrien $\alpha_1, \ldots, \alpha_k$ ei-ehdolliset suurimman uskottavuuden estimaatit eivät ole tarkentuvia ja ne ovat asymptoottisesti harhaisia, jolloin harhan approksimatiivinen suuruusluokka on k / (k-1). (Andersen, 1980)

Osioparametrien tarkentumattomuus on vakava ongelma, koska aineiston kasvattaminen johtaa erheellisiin estimaatteihin. Ongelma voidaan kuitenkin ratkaista ehdollistamalla uskottavuusfunktio, jolloin saadaan tarkentuvia kelvollisia osioparametrien estimaatteja. Ehtona on, että θ_i:t tunnetaan. Ehdollisen suurimman uskottavuuden estimointimenetelmän (conditional maximum likelihood estimation) perusajatuksena on estimoida osioparametrit α_j erikseen olettaen, että oppilasparametrit θ_i tunnetaan. Tämä on mahdollista, koska oppilaan saama testin kokonaispistemäärä eli oppilassummapistemäärä, merkitään sitä $T_i = X_{i.} = t_i$, on tyhjentävä tunnusluku θ_i:lle ja näin ollen se on riippumaton θ_i:sta. Seuraavassa on havainnollistettu osioparametrien α_j ehdollisen suurimman uskottavuuden estimointia Andersenin (1973a,1980) mukaisesti.

Johdetaan ehdollinen tiheysfunktio ehdolla $X_{i.} = t_i$, jolloin se saadaan esitetyksi ilman θ:aa. Kirjoitetaan aluksi tiheysfunktio, joka on muotoa

$$f(x_{i1},\ldots,x_{ik}) = \frac{e^{(\theta_i t_i - \sum_{j=1}^{k} \alpha_j x_{ij})}}{\prod_{j=1}^{k}(1+e^{(\theta_i-\alpha_j)})}. \qquad (2.4.1)$$

Huomataan myös, että Raschin mallin otantamalli kuuluu eksponenttiseen jakauma-perheeseen, jolloin tämän perheen teoreettiset ominaisuudet ovat käytettävissä.

Nyt tiheysfunktiosta saadaan oppilassummapistemäärän jakauma, joka on

$$f(t_i) = P\left\{\sum_{j=1}^{k} x_{ij} = t_i\right\} = \frac{e^{\theta_i t_i}}{\prod_{j=1}^{k}(1+e^{(\theta_i-\alpha_j)})\sum_{x_{i.}=t_i}^{n} e^{-\sum_{j=1}^{k}\alpha_j x_{ij}}}. \tag{2.4.2}$$

Jos nyt merkitään

$$\gamma(t_i,\alpha_1,\ldots,\alpha_k) = \sum_{x_{i.}=t_i}^{n} e^{-\sum_{j=1}^{k}\alpha_j x_{ij}} \tag{2.4.3}$$

niin ehdollinen tiheysfunktio voidaan esittää ilman parametria θ ja ehdollinen tiheysfunktio ehdollistettuna oppilassummapistemäärillä voidaan nyt kirjoittaa muotoon

$$f(x_{i1},\ldots,x_{ik}|t_i) = \frac{e^{-\sum_{j=1}^{k}\alpha_j x_{ij}}}{\gamma(t_i,\alpha_1,\ldots,\alpha_k)}. \tag{2.4.4}$$

Tiheysfunktio (2.4.4) on riippumaton θ_i:sta, joten nyt uskottavuusfunktio voidaan kirjoittaa muotoon

$$L_c = \prod_{i=1}^{n} f(x_{i1},\ldots,x_{ik}|t_i) = \frac{e^{-\sum_{j=1}^{k}\alpha_j x_{.j}}}{\prod_{r=0}^{k}\left[\gamma(t_i,\alpha_1,\ldots,\alpha_k)\right]^{n_r}}, \tag{2.4.5}$$

jossa n_r on oppilassummapistemäärien lukumäärä, $t_i = r$.

Nyt voidaan havaita, että uskottavuusfunktio (2.4.5) kuuluu ehdollistamisen jälkeenkin eksponentiaaliseen malliperheeseen. Koska oppilaiden saamat kokonaispistemäärät $t_1, t_2, \ldots$ tunnetaan, $n_0, n_1, \ldots, n_k$ tiedetään ja lisäksi ainoat uskottavuusfunktion 2.4.5 parametrit ovat sen nimittäjässä, niin uskottavuusfunktio voidaan kirjoittaa muotoon

$$L_c = c^{-1}(\alpha_1, \ldots, \alpha_{k-1}) e^{\sum_{j=1}^{k-1} \alpha_j(-x_{.j} + x_{.k})}, \tag{2.4.6}$$

jossa α_j:n minimaalisti tyhjentävinä tunnuslukuina ovat

$$y_j = x_{.k} - x_{.j}, \; j = 1, \ldots, k-1.$$

Näin eksponenttisen perheen ominaisuuksia käyttäen uskottavuusyhtälöryhmäksi saadaan

$$x_{.j} - x_{.k} = E\left[x_{.j} - x_{.k} | n_0, \ldots, n_k\right], \; j = 1, \ldots, k-1. \tag{2.4.7}$$

Ratkaisemalla edellä kuvattu yhtälöryhmä (2.4.7), saadaan osioparametreille $\alpha_1, \ldots, \alpha_k$:lle tarkentuvat estimaatit. Uskottavuusyhtälöryhmä ei kuitenkaan ratkea suljetussa muodossa, vaan se ratkaistaan numeerisesti käyttäen jotain iteratiivista laskenta-algoritmia. Kyseisen yhtälöryhmän ratkaisemiseksi voidaan käyttää Newton-Rapson-laskenta-algoritmia tai Fisherin scoring-algoritmia, joita tarkastellaan kappaleessa 2.6.

Andersenin (1980) mukaisesti ehdollisen suurimman uskottavuuden menetelmäkään ei ole aivan aukoton, sillä siinä teoreettisessa tilanteessa, jossa otoksen otoskoko n kasvaisi äärettömyyteen, myös parametrien lukumäärä kasvaisi äärettömyyteen. Tähänkin ongelmaan on kehitetty ratkaisu, osioanalyysi tehdään latentin luokan malleilla (latent class models). Tällöin oletetaan, että oppilasparametrien θ_i määrä on rajallinen, jolloin niillä oletetaan olevan m eri mahdollista arvoa. Tällaisessa tilanteessa populaatio jaetaan m:ään eri luokkaan ja näitä kyseisiä luokkia käytetään apuna parametrien estimoinnissa. Latentin luokan malleista on enemmän

kerrottu mm. Andersenin (1980) ja Bartholomevin (1987) teoksissa. Toisena vaihtoehtona on ajatella riviefektit satunnaisina, jolloin mallista saadaan muodostettua ehdollinen Poisson-malli. Tällöin henkilöpistemäärille johdetaan jakauma, jolloin riviefektit oletetaan satunnaisiksi ja tätä jakaumaa hyväksikäyttäen esitetään Raschin malli ei-ehdollisena. Ehdollisessa Raschin mallissahan riviefektit ovat kiinteitä. Raschin mallin esittämiseen ehdollisena Poisson-mallina voi tutustua mm. Tjurin (1980) artikkelin avulla.

2.5 Oppilasparametrien tarkentuva estimointi

Andersenin (1973a, 1980) mukaisesti oppilasparametrien tarkentuva estimointi voidaan tehdä kahdella eri tavalla. Mahdollisuutena on estimoida oppilasparametrit oppilaan vastauksista k:hon eri osioon. Tällöin käytettävä menetelmä on marginaaliuskottavuus menetelmä, jota tässä aliluvussa havainnollistetaan.

Vaihtoehtoisena tapana on tavallinen suurimman uskottavuuden estimointi, jolloin tarvitaan oletus θ_i:den jakaumasta, $\psi(\theta)$:sta. Andersen (1980, 264-268) on tarkastellut tätä estimointimenetelmää johtamalla osion ominaisfunktiolle uudenlaisen esityksen kriittisen pisteen käsitteen avulla. Tätä estimointimenetelmää ei tässä kontekstissa kuitenkaan tarkemmin käsitellä, vaan oppilasparametrien estimoimiseksi käytetään marginaaliuskottavuus menetelmää, jota seuraavassa havainnollistetaan.

Estimointi perustuu oppilaan i marginaaliuskottavuusfunktion

$$L(\theta_i) = P\{x_{i.} = t_i\}, \tag{2.5.1}$$

maksimointiin. Näin ollen oppilaan i marginaaliuskottavuus, kun osioparametrit tunnetaan kirjoitetaan muotoon

$$L(\theta_i) = \frac{e^{\theta_i t_i} \gamma(t_i, \alpha_1, \ldots, \alpha_k)}{\prod_{j=1}^{k} (1 + e^{(\theta_i - \alpha_j)})}, \tag{2.5.2}$$

jonka logaritmin derivaatta θ_i:n suhteen asetetaan nollaksi eli

$$\frac{\partial \ln L(\theta_i)}{\partial \theta_i} = \frac{t_i - \sum_{j=1}^{k} e^{(\theta_i - \alpha_j)}}{1 + e^{(\theta_i - \alpha_j)}} = 0, \tag{2.5.3}$$

jonka ratkaisuksi tulee yhtälöryhmä

$$t_i = \frac{\sum_{j=1}^{k} e^{(\theta_i - \alpha_j)}}{1 + e^{(\theta_i - \alpha_j)}}. \tag{2.5.4}$$

Nyt yhtälöryhmästä (2.5.4) oppilasparametrit θ_i voidaan ratkaista suoraan jos osioparametrit $\alpha_1, \ldots, \alpha_k$ tunnetaan.

Jos osioparametrit ovat tuntemattomia, niin tällöin α_j:t korvataan niiden ehdollisen suurimman uskottavuuden estimaateilla, jolloin ratkaistava yhtälöstö on muotoa

$$t_i = \frac{\sum_{j=1}^{k} e^{(\theta_i - \hat{\alpha}_j)}}{1 + e^{(\theta_i - \hat{\alpha}_j)}}. \tag{2.5.5}$$

Yhtälöryhmän (2.5.5) tuntemattomat parametrit voidaan ratkaista Newton-Raphson-algoritmilla (ks. aliluku 2.6).

Nyt jos osioparametrien estimaattien keskivirheet ovat pieniä, niin oppilasparametreille θ_i saadaan ominaisuuksiltaan samanlaisia estimaattoreita yhtälöryhmästä (2.5.5), kuin yhtälöryhmästä (2.5.4), jolloin osioparametrit α_j olivat tunnetut (Andersen, 1980).

Kappaleissa 2.4 ja 2.5 kuvatulla tavalla Raschin mallista voidaan estimoida oikeelliset osio- ja oppilasparametrien estimaatit. Estimaattorien keskivirhetarkasteluja ei tässä asiayhteydessä tarkemmin esitetä. Estimaattorien asymptoottiset varianssit lasketaan informaatiomatriisin käänteismatriisin diagonaalialkioista, joiden neliöjuuret ovat parametrien estimaattorien asymptoottisia keskivirheitä.

Keskivirhetarkasteluihin voi tutustua Wrightin (1977) avulla, joka on kehittänyt "nyrkkisäännön" parametrien estimaattien keskivirheiden arvioimiseksi. Tarkemmista keskivirhetarkasteluista kiinnostuneille Andersen (1980) ja Hambleton & Swaminathan (1985) tarjoavat tyhjentävän esityksen yksiparametrisen logistisen osioanalyysimallin keskivirhetarkasteluista ja mallin informaation mitoista.

2.6 Iteratiiviset laskenta-algoritmit

Yleisestikin suurin osa SU-estimoinnissa tai sen modifikaatioissa, kuten esimerkiksi ehdollisessa SU-estimoinnissa saaduista yhtälöryhmistä ratkeaa jollakin iteratiivisella laskenta-algoritmilla. Iteratiivisella yhtälöryhmien ratkaisutekniikalla tarkoitetaan käytännössä sitä, että laskenta-algoritmilla tuotetaan jono estimaatteja, alkaen ennalta kiinnitetystä alkuestimaatista. Estimaatteja päivitetään, kunnes estimaattijono saavuttaa konvergenssin, eli se ei enää muutu edellisestä estimaatista annettua rajaa enemmän. Seuraavassa on havainnollistettu kahta yleisimmin käytettyä iteratiivista laskenta-algoritmia, jotka ovat Newton-Raphson-algoritmi ja Fisherin scoring-menetelmä. Yleisesti algoritmeilla ratkaistaan yhtälöä

$$V(\theta;\underline{x}) = 0 \tag{2.6.1}$$

θ:n suhteen, missä $\underline{x} = x_1, x_2, \ldots, x_n$ on otos.

2.6.1 Newton-Raphson

Raschin mallissa yleisimmin käytetään Newton-Raphson algoritmia osio- ja oppilasparametrien ratkaisemiseksi (Andersen, 1980). Seuraavassa on Searlen, Casellan ja McCullochin (1992) mukaisesti havainnollistettu lyhyesti algoritmin toimintaa.

Vaihe 1: Asetetaan alkuarvo $\theta^{(0)}$

Vaihe 2: Päivitetään kaavalla, joka on muotoa

$$\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} - \left[\frac{\partial V}{\partial \theta}(\hat{\theta}^{(t)})\right]^{-1} V(\hat{\theta}^{(t)}), \tag{2.6.1.1}$$

jossa

$$V(\hat{\theta};\underline{x}) = \frac{\partial \ln L(\hat{\theta};\underline{x})}{\partial \theta}. \tag{2.6.1.2}$$

Vaihe 3: Lopetetaan iterointi, kunnes $|\hat{\theta}^{(t+1)} - \hat{\theta}^{(t)}| \leq \varepsilon$, missä ε käyttäjän ennalta määrittämä algoritmin pysäytystarkkuus.

Yleisesti ottaen algoritmin konvergointinopeus riippuu alkuarvon valinnasta. Matemaattisesti ajatellen on mahdollista, että algoritmi ei konvergoi ollenkaan, jos se on todella huonosti valittu. Jos alkuarvo valitaan "läheltä" oikeaa ratkaisua, niin algoritmi toimii hyvin ja konvergoi oikeaan kohtaan. Useissa nykyohjelmistoissa on oma tekniikkansa juuri alkuarvojen valintaan (Hambleton & Swaminathan, 1985). Tähän seikkaan ei kuitenkaan tässä tutkimuksessa keskitytä enempää.

2.6.2 Fisherin scoring-menetelmä

Fisherin scoring-menetelmä on Newton-Raphson-algoritmin kaltainen, mutta siinä vaiheessa 2, päivitysaskeleen (2.6.1.1) termi, score-matriisin derivaatta

$$\frac{\partial V(\theta;\underline{x})}{\partial \theta}$$

korvataan sen odotusarvolla. Tällöin vaiheen 2 päivityskaava saa Fisherin scoringissa muodon

$$\hat{\theta}^{(t+1)} = \hat{\theta}^{(t)} + E\left[\frac{\partial V(\hat{\theta}^{(t)})}{\partial \theta}\right]^{-1} V(\hat{\theta}^{(t)}), \tag{2.6.2.1}$$

jossa

$$E\left[\frac{\partial V(\theta;\underline{x})}{\partial \theta}\right] = -E\left\{-\frac{\partial^2 \ln L(\theta;\underline{x})}{\partial \theta^2}\right\}. \tag{2.6.2.2}$$

Käytännössä siis Fisherin scoringin algoritmi hakee ratkaisua uskottavuusfunktion odotetun tangentin suunnassa, jolloin käännettävänä oleva matriisi riippuu edellisen askeleen estimaatin arvosta, mutta ei havainnosta. Newton-Raphson-algoritmihan hakee ratkaisua havaitun tangentin suunnassa.

Yleensä Fisherin scoringing-menetelmässä konvergointi on hitaampaa eli siinä iterointikierroksia tulee enemmän kuin Newton-Raphson-algoritmissa, mutta sen yhden päivitysaskeleen laskeminen on nopeampaa, sillä Fisherin scoring on havainnoista riippumaton, eikä niitä tarvitse huomioida jokaisella päivitysaskeleella. Käänteismatriisi riippuu siis ainoastaan edellisen parametrin arvosta, ei suoraan aineistosta.

Lisäksi on hyvä huomata, että (2.6.2.2):n käänteismatriisi lähestyy informaatiomatriisin käänteismatriisia. Sen diagonaalialkiot antavat parametrien estimaattorien asymptoottiset varianssit. Vastaavasti myös havaitun informaatiomatriisin käänteismatriisi on yhtäpitävä (2.6.2.2):n kanssa suurilla otoksilla. Siten sekä Newton-Raphson, että Fisherin scoring antavat sivutuotteina estimaattorien asymptoottiset varianssit.

2.7 Lisäparametrisia vaihtoehtoja Raschin mallille

Kuten jo kappaleessa 2.1 mainittiin, niin usein koulutusta tutkittaessa on käytetty logistisia jakaumamalleja, joissa on yhdestä kolmeen osiota kuvaavaa parametria latentin avaruuden ollessa yksiulotteinen. Onhan toki olemassa paljon muitakin mallivaihtoehtoja kuten nominaalimalli ja intervalliasteikollisen latentin piirteen mallit, joissa osion oletetaan mittaavan latenttia piirrettä välimatka-asteikon tasolla. Näiden kaltaisiin malleihin palataan yleistetyn Raschin mallin parissa. Seuraavassa on lyhyesti esitetty koulutuksen tutkimuksessa yleisimmin käytetyt kaksi- ja kolmiparametriset osioanalyysimallit.

Hieman Raschin mallia kompleksisemman mallit kuten kaksiparametriset osioanalyysimallit ovat kehittyneet logistisesta mallista. Malliin lisätään tällöin toinen osiokohtainen parametri β, jota kutsutaan osion erottelukykyparametriksi. Tällaista mallia kutsutaan Birnbaumin (1968) malliksi. Malli on muotoa

$$P(X_{ij} = 1|\theta_i, \alpha_j, \beta_j) = \frac{e^{d\beta_j(\theta_i - \alpha_j)}}{1 + e^{d\beta_j(\theta_i - \alpha_j)}}, \tag{2.7.1}$$

jossa α_j on osion vaikeustasoparametri ja β_j osion erottelukykyparametri. Malli on logistinen silloin, kun $d = 1$ ja approksimatiivisesti normaalijakaumamalli, jos $d \approx 1{,}7$. Tällä tarkoitetaan, että malli approksimoi kaavan (2.1.2.5) mukaista mallia parametrein θ_i, α_j, β_j, jos $d \approx 1{,}7$.

Erottelukykyparametri β_j on suoraan verrannollinen P_j:n tangentin kulmakertoimeen eli kuvaajan jyrkkyyteen kun $\theta_i = \alpha_j$. Mitä suurempi erottelukykyparametrin arvo on, sitä paremmin osio erottelee. Tosin on todettava, että vaikka erottelukykyparametrin arvo olisikin suuri, se ei välttämättä erottele hyvin kaikilla θ_i:n arvoilla. Erottelukykyparametri kertoo sen, kuinka hyvin osio erottelee parhaiten erottelevalla alueella, jolloin osio- ja kykyparametrien arvot ovat lähellä toisiaan. Voidaan myös todeta, että mitä suurempi β_j:n arvo on, sitä varmemmin ja johdonmukaisemmin osio hylkää arvon α_j alapuolella olevat ja läpäisee yläpuolella olevat. Yleisesti ottaen vaikeustasoparametrin arvot rajataan välille [0,2], vaikka se voisi saada arvoja äärettömyyteenkin saakka.

Mallin parametrit voidaan estimoida maksimoimalla marginaaliuskottavuusfunktio tai tehdä estimointi Bayesilaisittain. Näitä molempia estimointitapoja on havainnollistettu esim. Hambletonin & Swaminathanin (1985) teoksessa.

Jos kyseessä on monivalintaosio, niin silloin on mahdollista, että oppilas voi vastata osioon oikein myös arvaamalla, vaikka oppilasparametrin arvo olisi pienempi kuin osioparametrin arvo, $\theta < \alpha_j$. Tätä varten on kehitetty kolmiparametrinen osioanalyysimalli, jossa vastaamistodennäköisyydet esitetään arvauskomponentin ja kaksiparametrisen logistisen käyrän sekoituksena. Arvauskomponentti on kolmas osiosta toiseen vaihteleva parametri, joka kuvastaa oppilaan mahdollisuuksia saada osio oikein arvamalla.

Kolmiparametrisessa mallissa oletetaan, että henkilö tietää, milloin hän osaa ratkaista osion oikein. Jos hän arvelee ratkaisevansa osion, hänen todennäköisyytensä ratkaista osio noudattaa kaksiparametrista mallia. Jos hän puolestaan arvaa, hänen mahdollisuutensa saada osio oikein riippuu todennäköisyyttä kuvastavasta parametrista γ_j. Näin ollen kolmiparametrinen osioanalyysimalli on muotoa

$$P(X_{ij} = 1|\theta_i, \alpha_j, \beta_j, \gamma_j) = \gamma_j + (1 - \gamma_j)\frac{e^{d\beta_j(\theta_i - \alpha_j)}}{1 + e^{d\beta_j(\theta_i - \alpha_j)}}, \qquad (2.7.2)$$

missä γ_j on osion nk. pseudoarvoparametri (= sekoitusparametri).

Pseudoarvoparametri on osion ominaiskäyrän ala-asymptootti, jota pienemmäksi ei osion läpäisemistodennäköisyys voi tulla. Huomataan myös, että β_j-parametri riippuu α_j:n kohdalla erotuksesta (1-γ_j) eli osio erottelee sitä paremmin, mitä pienemmät mahdollisuudet henkilöllä on saada osio oikein, silloin kun hän ei sitä osaa. Mallin käyttö edellyttää isoja aineistoja ja kehittyneitä estimointialgoritmeja, esim. Bayesilaista estimointia (ks. Hambleton & Swaminathan, 1985; Mislevy, 1991).

2.8 Yhteenveto Raschin mallista

Raschin mallia ja sen lisäparametrisia vaihtoehtoja, käytetään monivalintakokeiden analysointimenetelmänä testin mitta-asteikon ollessa dikotominen. Raschin mallissa ja myös kaksi- ja kolmiparametrisessa mallissa latentti piirre oletetaan yksidimensionaaliseksi. Mallien puutteina on, että ne olettavat latentin struktuurin

yksidimensionaaliseksi, vaikka todellisuudessa se voi olla kompleksisempi. Toisena ongelmana on dikotomisen mitta-asteikon epäinformatiivisuus. Traditionaalisempien osioanalyysimallien etuna on, että parametrit ovat tulkittavia, joka helpottaa tutkimustulosten raportointia ja johtopäätösten tekemistä.

Raschin mallissa osioparametrien estimointi tapahtuu ehdollisen suurimman uskottavuuden menetelmää käyttäen ja oppilasparametrien estimointi puolestaan marginaaliuskottavuusmenetelmällä.(Andersen, 1973a). Osioparametrien estimointiin on myös ehdotettu marginaaliuskottavuumenetelmää (Thissen, 1982). Mallin yleistä yhteensopivuustestiä eli mallin homogeenisuustestiä ja parametrien keskivirhetarkasteluja ei tarkasteltu tässä asiayhteydessä, sillä yhteensopivuustestausta ja keskivirhetarkasteluja havainnollistetaan yleistetyn Raschin mallin parissa. Kyseinen mallin yhteensopivuustesti soveltuu myös Raschin mallille. Yleisistä yhteensopivuustesteistä kiinnostuneille suosittelen tutustumista näihin Andersenin (1973b tai 1980) teosten tai Gustafssonin (1977) julkaisun avulla, joista löytyy myös parametrien keskivirhetarkastelut.

Raschin malli eli yksiparametrisen logistisen osioanalyysimalliln rinnalle on kehitetty myös lisäparametrisia vaihtoehtoja, joista käytetyimpiä ovat kaksi- ja kolmiparametrinen osioanalyysimalli.

Luku 3

Yleistetty Raschin malli

Yleistetty Raschin malli on kehitetty tilanteisiin, joissa mahdollisesti polytomisesti koodatusta aineistosta voidaan mallintaa yhtäaikaisesti useita eri latentteja piirteitä. Koulutustutkimuksessa tämä tarkoittaa, että oppilaiden osaamista on mahdollista mallintaa usealta eri osa-alueelta, esimerkiksi algebran, analyysin ja geometrian osa-alueelta. Raschin mallilla rajoituttaisiin tällaisessa tilanteessa mallintamaan oppilailta ainoastaan esimerkiksi yleistä matemaatiikan osaamista.

Mallin yleisestä muodosta, jota kutsutaan yleistetyksi Raschin malliksi, on mahdollista spesifioida lukuisia erilaisia osioanalyysimalleja, mukaan lukien Raschin malli. Mallin valinta riippuu siitä, kuinka käyttäjä määrittää mallin komponentit, jotka ovat sen asetelmamatriisi **A** ja pisteytysmatriisi **B**, joista jälkimmäinen supistuu pisteytysvektoriksi **b**, jos latentti piirre oletetaan yksiulotteiseksi. Asetelmamatriisilla **A** määritetään osiokohtaiset parametrit osioparametrivektorista ξ. Pisteytysvektorilla **b** pisteytetään eri osioiden vastauskategoriat siten, kuinka niiden oletetaan heijastavan yhtä spesifiä latenttia piirrettä θ. Pisteytysmatriisilla **B** pisteytetään eri osioiden vastauskategoriat siten, kuinka niiden oletetaan heijastavan kutakin latenttia piirrettä eli osaamistasoa eri osiokokeen osa-alueilla. Tällaisessa moniulotteisessa tilanteessa piilomuuttuja ei ole enää skalaari, vaan se laajenee vektoriksi,

$$\theta^T = (\theta_1, \ldots, \theta_D).$$

Yleistetty Raschin malli yhden osion tilanteessa kuvaa todennäköisyyttä tuottaa tietty osiovastaus ja mallin yleinen muoto kuvaa vastaavasti todennäköisyyttä tuottaa koko testin eli osiokokeen τ kattava vastausvektori ehdolla, että piilomuuttujan tai piilomuuttujavektorin arvot tunnetaan. Näin ollen malli on ehdollinen osioanalyysimalli, kuten luvussa 2 esitelty Raschin malli ja sen lisäparametriset vaihtoehdot. Malli voidaan kuitenkin määritellä ei-ehdolliseksi eli marginaaliosioanalyysiymalliksi, jolloin piilomuuttujat oletetaan olevan jostain jakaumasta. Piilomuuttujan tai piilomuuttujavektorin jakaumaa kutsutaan populaatiomalliksi.

Populaatiomallin mukaan ottamisella osiovastauksista voidaan suoraan osioanalyysin yhteydessä estimoida erilaisten kiinteiden oppilastason muuttujien vaikutuksia osiovastauksiin. Tämä mahdollistaa esimerkiksi sukupuolten välisten erojen mittaamisen suoraan osioaineistosta.

Tässä luvussa esitetään aluksi yleistetty Raschin malli, jonka jälkeen sen asetelmamatriisin **A** ja pisteytysmatriisin **B** formulointia on havainnollistettu esimerkkitilantein spesifioimalla Raschin mallista tavanomaisia osioanalyysimalleja. Luvun lopussa on esitetty populaatiomalli ja yleistetyn Raschin mallin identifioituvuusehdot. Luvun kuvaukset on pääpiirteittäin tehty tulkitsemalla Wu:n, Adamsin ja Wilsonin (1998) kirjoitusta, johon erillistä viittausta ei yleisesti ole tehty.

3.1 Yleistetyn Raschin mallin ominaisfunktio

Käydään seuraavassa lyhyesti läpi yleistetyssä Raschin mallissa käytettävä perusnotaatio ja sen komponentit. Aikaisemmin tutkimuksessa oppilasta indeksoitiin i:llä. Yleistetyssä Raschin mallissa osioita indeksoidaan i:llä, $i = 1,...,I$. Näissä kussakin i:ssä osiossa oletetaan olevan $Ki + 1$ vastausvaihtoehtoa, $k = 0,...,Ki$. Ki:n arvo riippuu siitä, kuinka monen eri vastausvaihtoehdon oletetaan indikoivan jotain latenttia piirrettä. Raschin mallissahan $Ki = 1$ jokaiselle osiolle, sillä siinä kategorioita on vain kaksi eli tällöin on mahdollista antaa joko oikea tai väärä vastaus. Polytomisesti pisteytetyssä osioaineistossa mallissa on vastauskategorioita useampia.

Yleistetyssä Raschin mallissa vektoriarvoisella satunnaismuuttujalla

$$X_i^T = (X_{i1}^T, X_{i2}^T, \ldots, X_{ik_i}^T),$$

missä

$$X_{ij} = \begin{cases} 1 & \text{jos vastaus osioon i on kategoriassa j} \\ 0 & \text{muulloin,} \end{cases} \tag{3.1.1}$$

mitataan kaikkia tietyn osion i vastauksia eri vastauskategorioihin. Osiovastauksia mittaavat satunnaismuuttujan arvot kaikille testin τ osioille, voidaan kirjoittaa koko testin τ kattavaksi vektoriksi

$$X^T = (X_1^T, X_2^T, \ldots, X_I^T). \tag{3.1.2}$$

Vektoria (3.1.2) kutsutaan vastausvektoriksi. Tällöin mahdollisia vastausvektoreita X koko testissä τ on yhteensä

$$\prod_{i=1}^{I} (Ki + 1). \tag{3.1.3}$$

Mahdollisten vastausvektorien lukumäärä riippuu näin ollen aineiston pisteytystavasta ja testin τ osioiden lukumäärästä. Testin τ kaikkia vastausvektoreita ei ole mahdollista empiirisesti havaita, jos osioiden lukumäärä on suuri ja otoskoko n pieni, sillä mahdollisten vastausvektoreiden määrä kasvaa hetkessä suuremmaksi, kuin otoskoko n. Tietyltä yksilöltä havaittua vastausvektoria merkitään tutkielmassa X_n:llä.

Seuraavassa on esitetty yleistetyn Raschin mallin ominaisfunktio aluksi rajaten se yksiulotteiseksi. Siitä tehdään suora laajennus moniulotteiseksi malliksi, jota kutsutaan yleistetyksi Raschin malliksi. Mallilla voidaan analysoida polytomisesti koodatuista osiovastauksista useita eri latentteja piirteitä.

Oletetaan, että oppilaille on tehty osiokoe τ, jossa osiota on indeksoitu i:llä, $i = 1,\ldots,I$. Näissä oletetaan kussakin olevan j eri vastauskategoriaa, $j = 1,\ldots,k$. Tällöin yksiulotteiseksi rajatussa yleistetyssä Raschin mallissa asetelmamatriisilla **A** määritellään jokainen osiokohtainen parametri kullekin testin τ osiolle osioparametrivektorista $\xi^T = (\xi_1,\ldots,\xi_p)$.

Pisteytysvektorilla **b** puolestaan määritellään, kuinka voimakkaasti vastaukset oletetaan heijastavan latenttia piirrettä θ. Nythän θ on skalaari, koska malli oletetaan yksiulotteiseksi (ks. luku 2).
Yleistetyn Raschin mallin ominaisfunktio kuvaa todennäköisyyttä, että tietty osiovastaus yhteen osioon i on vastauskategoriassa j ehdolla, että piilomuuttujan θ arvo tunnetaan. Se kirjoitetaan muodossa

$$P(X_{ij} = 1; A, b, \xi | \theta) = \frac{e^{(b_{ij}\theta + a_{ij}{}^T \xi)}}{\sum_{k=1}^{K_i} e^{(b_{ik}\theta + a_{ik}{}^T \xi)}}, \tag{3.1.4}$$

ja tästä seuraten tietylle vastausvektorille X, ominaisfunktio kirjoitetaan muodossa

$$f(x; \xi | \theta) = \frac{e^{[x^T (b\theta + A\xi)]}}{\sum_{z \in \Omega} e^{[z^T (b\theta + A\xi)]}}, \tag{3.1.5}$$

jossa Ω on kaikkien mahdollisten testin τ vastausvektoreiden arvot. Nyt malli (3.1.5) kuvaa todennäköisyyttä tuottaa tietty vastausvektori testissä τ ehdolla, että piilomuuttujan θ arvo tunnetaan.

Yksiulotteinen yleistetty Raschin malli laajenee suoraviivaisesti moniulotteiseksi, jolloin piilomuuttuja on vektoriarvoinen, $\theta^T = (\theta_1, \ldots, \theta_D)$ .Tällöin oletetaan, että osiovastauksista mallinnetaan D:tä eri latenttia piirrettä. Tällöin mallin asetelmamatriisi **A** on kuten yksiulotteisessa mallissa, sillä vastauksista estimoitavien osioparametrien lukumäärä ei muutu vaikka piilomuuttuja θ määritellään vektoriarvoiseksi. Pisteytysvektori muuntuu pisteytysmatriisiksi **B**, jolla pisteytetään eri osioiden vastauskategoriat siten, kuinka ne heijastavat kutakin latenttia piirrettä piilomuuttujavektorista $\theta^T = (\theta_1, \ldots, \theta_D)$. Tällöin todennäköisyys sille, että osiovastaus osioon i on vastauskategoriassa j ehdolla, että piilomuuttujavektorin $\theta^T = (\theta_1, \ldots, \theta_D)$ arvot tunnetaan, voidaan kirjoittaa muotoon

$$P(X_{ij} = 1; A, B, \xi | \theta) = \frac{e^{(b_{ij}{}^T\theta + a_{ij}{}^T\xi)}}{\sum_{k=1}^{K_i} e^{(b_{ik}{}^T\theta + a_{ik}{}^T\xi)}}. \tag{3.1.6}$$

Vastaavasti saadaan ominaisfunktio, ts. todennäköisyys, että vastaaja tuottaa tietyn vastausvektorin X testissä τ, ehdolla, että piilomuuttujavektorin $\theta^T = (\theta_1, \ldots, \theta_D)$ arvot tunnetaan. Se kirjoitetaan muotoon

$$f(x; \xi | \theta) = \frac{e^{[x^T(B\theta + A\xi)]}}{\sum_{z \in \Omega} e^{[z^T(B\theta + A\xi)]}}. \tag{3.1.7}$$

Mallia (3.1.7) sanotaan yleistetyksi Raschin malliksi ja siihen ei liity mitään rajauksia. Mallissa (3.1.7) Ω on kaikkien mahdollisten testin τ vastausvektoreiden arvot.

Ainoa ero yksiulotteisella ja moniulotteisella esityksellä eli yleistetyllä Raschin mallilla on, että yksiulotteisessa mallissa kykyä kuvaava latentti muuttuja θ on skalaari ja täydellisessä yleistetyssä mallissa se on vektoriarvoinen eli D - ulotteinen vektori, jolloin pisteytysvektori **b** laajenee pisteytysmatriisiksi **B**.

3.2 Mallin asetelmamatriisin ja pisteytysmatriisin käyttö erilaisten mallien spesifioinnissa

Yleistetyn Raschin mallin käyttö osioaineiston mallintamiseksi edellyttää sen keskeisten komponenttien eli asetelmamatriisin **A** ja pisteytysmatriisin **B** määrittä-

mistä. Näiden käytön havainnollistaminen jää kuitenkin hyvin abstraktiksi ilman käytännöllisiä esimerkkejä. Tässä alaluvussa on esitelty tavallisen Raschin mallin, polytomisesti pisteytettyjen vastausten osioanalyysimallin (partial credit model) ja moniulotteisen Raschin mallin spesifioiminen yleistetystä Raschin mallista (3.1.7). Lisäksi mallien spesifioinnin yhteydessä on esitetty esimerkkitilanteita, joissa mallien sovittaminen aineistoon on mahdollisesti aiheellista.

3.2.1 Raschin mallin spesifioiminen yleistetystä Raschin mallista

Ajatellaan esimerkkitilannetta, että on tehty testi, jossa matematiikan osaamista mitataan neljällä osiolla, jolloin $i = 1,...,4$. Jos aineistoon sovitetaan tavallinen Raschin malli (ks. Luku 2), niin koko neljän osion testille τ, asetelmamatriisi **A** ja pisteytysvektori **b** kirjoitetaan seuraavasti:

$$A = \begin{Bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \quad \text{ja} \quad b = \begin{Bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{Bmatrix} \tag{3.2.1.1}$$

Kun testissä on 4 osiota ja vastaukset ovat dikotomisesti koodatut, niin osiota eli osion vaikeustasoa kuvaavia parametreja on neljä, yksi jokaista osiota kohden. Tällöin asetelmamatriisissa **A** on neljä saraketta. Ensimmäisellä sarakkeella määrite-

tään osioparametrivektorista ξ osion 1 vaikeustasoa kuvaava parametri. Vastaavasti asetelmamatriisin **A** toisella sarakkeella määritellään osioparametrivektorista ξ vaikeustasoparametri osiolle 2 ja edelleen osioille 3 ja 4 vastaavalla tavalla.

Vastauskategorioita on osiokohtaisesti kaksi eli henkilö voi vastata joko oikein tai väärin, jonka mukaan vastaukset pisteytetään. Tällöin pisteytysvektorissa **b** on kahdeksan riviä ja luonnollisesti yksi sarake. Nyt esimerkiksi pisteytysvektorin **b** ensimmäisen rivin alkio pisteytetään nollaksi, joka indikoi, että vastaus ensimmäiseen osioon on väärä ja vastauksen ei oleteta heijastavan latenttia piirrettä, jota testillä τ mitataan. Pisteytysvektorin **b** toisen rivin alkio puolestaan pisteytetään ykköseksi, jolloin se Raschin mallissa kuvaa, että tähän kategoriaan sijoittuva vastaus osioon 1 on oikein , jolloin sen oletetaan heijastavan latenttia piirrettä θ. Pisteytysvektorin **b** kolmannen ja neljännen rivin alkiot puolestaan kuvaavat vastaavaa osiolle 2 ja loput alkiot vastaavalla tavalla osioille 3 ja 4.

Ennen kuin malli kirjoitetaan eksplisiittisesti, on syytä palauttaa mieleen vastaustodennäköisyyden kaava (3.1.4), joka on siis muotoa

$$P(X_{ij} = 1; A,b,\xi|\theta) = \frac{e^{(b_{ij}\theta + a_{ij}{}^T\xi)}}{\sum_{k=1}^{K_i} e^{(b_{ik}\theta + a_{ik}{}^T\xi)}}, \qquad (3.2.1.2)$$

josta tavallinen Raschin malli yhdelle osiolle voidaan spesifioida. Nyt esimerkiksi todennäköisyys, että tietty osiovastaus esimerkkitilanteen ensimmäiseen osioon on oikein ehdolla, että piilomuuttujan arvo tunnetaan on

$$P(X_{12} = 1; A,b,\xi|\theta) = \frac{e^{(1\cdot\theta + 1\cdot\xi)}}{e^{(0\cdot\theta + 0\cdot\xi)} + e^{(1\cdot\theta + 1\cdot\xi)}} = \frac{e^{(\theta+\xi)}}{1 + e^{(\theta+\xi)}}, \qquad (3.2.1.3)$$

joka on luvussa 2 kuvattu Raschin malli (vrt. kaava 2.2.3). Ainoa erohan on, että mallista estimoitava osioparametrin arvo kuvaa suoraan osion vaikeustasoa eli se saa suuria negatiivia arvoja, jos osio on vaikea ja vastaavasti positiivisia, jos osio on helppo.

Esimerkkitilanteessa tietylle vastausvektorille X, ominaisfunktion kirjoittaminen eksplisiittisesti ei ole järkevää kaavasta (3.1.4), sillä kaikkia mahdollisia vastausvektoreita on jo neljän osion osiokokeessa τ yhteensä 16, joten kaavaan sijoitettavat matriisit kasvavat jo näinkin pienen osioaineiston tilanteessa suuriksi.

3.2.2 Polytomisesti pisteytettyjen osiovastausten mallintaminen yleistetyllä Raschin mallilla

Kuten jo aikaisemmin mainittiin, niin dikotominen mitta-asteikko on joissain tilanteissa epäinformatiivinen, sillä tällöin ajatellaan, että ainoastaan täysin oikeat vastaukset indikoivat osiokokeella mitattavaa latenttia piirrettä. Seuraavassa on Mastersin (1982) esimerkin avulla havainnollistettu informatiivisempaa osioiden pisteytystapaa. Esimerkissä käsitellään matematiikan osaamista mittaavaa osiota , jonka pisteytystapa on hierarkkinen.

Esimerkki:

Osiokysymys: $\sqrt{7.5 / 0.3} = ?$

Pisteytys: Täysin väärä vastaus 0

$7.5 / 0.3 = 25$ 1

$\sqrt{25} = 9$ 2

Edellisessä esimerkissä oppilaan osiovastaus pisteytetään sen mukaisesti, kuinka paljon hänen vastauksensa oletetaan indikoivan latenttia piirrettä. Pisteytystapa etenee vaiheittain. Tehtävän ratkaiseminen edellyttää ensin jakolaskutoimituksen tekemistä, jonka jälkeen neliöjuuren ottamista jakolaskutoimituksen tuloksesta. Ideana on, että oppilaan vastaus pisteytetään sen mukaisesti, kuinka paljon hän on osannut tehtävää ratkaistessaan. Jos vastaus olisi dikotomisesti koodattu, niin kaikki muut vastaukset olisivat koodattu nollaksi, paitsi viimeinen eli täysin oikea vastaus. Näin ollen dikotominen pisteytystapa olisi yksipuolisen "tiukkalinjainen" ja jättäisi huomiotta vastaukset, jotka on tehty tiettyyn vaiheeseen asti oikein.

Jatketaan edelleen esimerkillä asetelmamatriisin **A** ja pisteytysvektorin **b** konstruoinnista polytomisesti koodattujen osiovastausten tilanteessa, jossa piilomuuttuja oletetaan edelleen yksidimensionaaliseksi. Osioita on nyt kolme kappaletta, joiden vastausvaihtoehdot oletetaan olevan pisteytetyt 0,1,2 sen mukaisesti, kuinka voimakkaasti niiden oletetaan indikoivan latenttia piirrettä, jota testillä τ mitataan. Tällöin asetelmamatriisi **A** ja pisteytysvektori **b** muodostetaan seuraavasti:

$$A = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{Bmatrix} \qquad b = \begin{Bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 1 \\ 2 \\ 0 \\ 1 \\ 2 \end{Bmatrix} \tag{3.2.2.1}$$

Asetelmamatriisissa **A** on 6 saraketta, koska osioita on kolme ja nyt kutakin osiota kuvataan kahdella osioparametrilla, sillä kussakin osiossa on kaksi vastauskategoriaa, joiden oletetaan indikoivan latenttia piirrettä. Matriisin **A** rivit 1- 3 määrittelevät

osion 1 parametrit osioparametrivektorista ξ, rivit 4-6 puolestaan osion 2 parametrit ja rivit 7-9 puolestaan osion 3 parametrit.

Pisteytysvektorin **b** kolmen ensimmäisen rivin alkiot määrittävät voimakkuuden, jolla tietty osiovastaus osioon 1 heijastaa latenttia piirrettä. Rivit 4 - 6 puolestaan osiolle 2 ja rivit 7-9 osiolle 3 vastaavasti.

Kirjoitetaan mallista (3.1.3) todennäköisyys, että vastaus osioon 1 on on täysin väärin eli saa arvon 0. Tällöin todennäköisyys ehdolla, että θ tunnetaan voidaan kirjoittaa muotoon

$$P(X_{11}=1;A,b,\xi|\theta)=\frac{e^{(0\cdot\theta+0\cdot\xi_{1o})}}{e^{(0\cdot\theta+0\cdot\xi_{1o})}+e^{(1\cdot\theta+1\cdot\xi_{11})}+e^{(2\cdot\theta+1\cdot\xi_{12})}}$$

$$=\frac{1}{1+e^{(\theta+\xi_{11})}+e^{(2\theta+\xi_{12})}}. \qquad (3.2.2.2)$$

Vastaavasti ehdollinen todennäköisyys, että vastaus osioon 1 on ratkaistu osittain oikein voidaan kirjoittaa kaavasta (3.1.3) muotoon

$$P(X_{12}=1;A,b,\xi|\theta)=\frac{e^{(\theta+\xi_{11})}}{1+e^{(\theta+\xi_{11})}+e^{(2\theta+\xi_{12})}}. \qquad (3.2.2.3)$$

Edelleen todennäköisyys, että osio 1 on ratkaistu täysin oikein, saadaan muotoon

$$P(X_{13}=1;A,b,\xi|\theta)=\frac{e^{(2\theta+\xi_{12})}}{1+e^{(\theta+\xi_{11})}+e^{(2\theta+\xi_{12})}}. \qquad (3.2.2.4)$$

3.2.3 Moniulotteisen Raschin mallin konstruointi yleistetystä Raschin mallista

Tähän asti tässä aliluvussa käsitellyt mallit ovat olleet yksiulotteisia osioanalyysimalleja. Sekä Raschin malli, että polytomisesti koodattujen osiovastausten malli voidaan laajentaa tilanteeseen, jossa oletetaan, että osiokokeen osiot mittaavat eri latentteja piirteitä. Tässä kontekstissa esitellään ainoastaan moniulotteisen Raschin mallin johtaminen tilanteeseen, jossa vastaukset ovat siis koodattu dikotomisesti.

Jos osiokokeessa on paljon osioita ja osiokokeella halutaan testata erilaisia kognitiivisia kykyjä oppilailta, on melko rohkeaa olettaa, että osiot mittaavat kaikki samaa latenttia piirrettä. Esimerkinomaisesti voidaan ajatella tilannetta, jossa osiolla 1 mitataan oppilailta kykyä ratkaista matemaattisia yhtälöitä, osiolla 2 puolestaan mitataan tilastotieteen osaamista ja osiolla 3 puolestaan jotain, jossa oikean vastauksen tuottaminen edellyttää kummankin osiossa 1 ja 2 tarvittavan taidon hallitsemista. Tällöin asetelmamatriisi **A** ja pisteytysmatriisi **B** konstruoidaan seuraavasti:

$$A = \begin{Bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{Bmatrix} \quad \text{ja} \quad B = \begin{Bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{Bmatrix}. \qquad (3.2.3.1)$$

Nyt asetelmamatriisi **A** on samanlainen kuin tavallisessakin Raschin mallissa eli sillä määritellään osioparametrivektorista ξ parametri kuvaamaan kunkin kolmen osion vaikeustasoa.

Edellisessä esimerkkitilanteessa erona tavalliseen Raschin malliin on, että pisteytysvektori **b** on monidimensionaalisessa Raschin mallissa pisteytysmatriisi **B**. Nyt pisteytysmatriisin kaksi ensimmäistä riviä koskevat osiota 1, rivit 3 ja 4 osiota 2 ja rivit 5 ja 6 osiota 3. Nyt osio 1 pisteytetään heijastamaan latenttia piirrettä θ_1, jos vastaus on oikein. Osio 2 puolestaan pisteytetään siten, että oikean osiovastauksen oletetaan heijastavan latenttia piirrettä θ_2. Oikea vastaus osioon 3 puolestaan pisteytetään heijastamaan molempia latentteja piirteitä, sekä θ_1:tä, että θ_2:ta.

Esimerkkitilanteessa oikean vastauksen tuottamisen todennäköisyys osioon 1 ehdolla, että $\theta^T = (\theta_1, \theta_2)$ tunnetaan kirjoitetaan kaavasta (3.1.5) muotoon

$$P(X_{12} = 1; A, B, \xi | \theta) = \frac{e^{\left([1 \;\; 0]\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} + 1\cdot\xi_1\right)}}{e^{\left([0 \;\; 0]\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} + 0\cdot\xi_1\right)} + e^{\left([1 \;\; 0]\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} + 1\cdot\xi_1\right)}}$$

$$= \frac{e^{(\theta_1 + \xi_1)}}{1 + e^{(\theta_1 + \xi_1)}} \qquad (3.2.3.2)$$

ja vastaavasti osiolle 2 todennäköisyys, että oppilas tuottaa oikean osiovastauksen ehdolla, että $\theta^T = (\theta_1, \theta_2)$ tunnetaan, kirjoitetaan muotoon

$$P(X_{22} = 1; A, B, \xi | \theta) = \frac{e^{\left([0 \;\; 1]\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} + 1\cdot\xi_2\right)}}{e^{\left([0 \;\; 0]\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} + 0\cdot\xi_2\right)} + e^{\left([0 \;\; 1]\begin{bmatrix}\theta_1\\ \theta_2\end{bmatrix} + 1\cdot\xi_2\right)}}$$

$$= \frac{e^{(\theta_2 + \xi_2)}}{1 + e^{(\theta_2 + \xi_2)}}. \qquad (3.2.3.3)$$

Edelleen osiolle 3 todennäköisyys, että oppilas tuottaa oikean osiovastauksen ehdolla, että $\theta^T = (\theta_1, \theta_2)$ tunnetaan, kirjoitetaan muotoon

$$P(X_{32} = 1; A, B, \xi | \theta) = \frac{e^{\left([1 \;\; 1]\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix} + 1 \cdot \xi_3\right)}}{e^{\left([0 \;\; 0]\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix} + 0 \cdot \xi_3\right)} + e^{\left([1 \;\; 1]\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix} + 1 \cdot \xi_3\right)}}$$

$$= \frac{e^{((\theta_1 + \theta_2) + \xi_3)}}{1 + e^{((\theta_1 + \theta_2) + \xi_3)}}. \qquad (3.2.3.4)$$

3.3 Populaatiomalli

Tähän mennessä on tarkasteltu osiovastausten todennäköisyysmallia olettaen latentin piirteen olevan vakio. Kuitenkin populaatiossa se vaihtelee. Tämän vaihtelun mallintaminen voidaan tehdä kahdella eri tavalla, kiinteillä efekteillä (kuten Raschin mallissa) tai satunnaisefekteillä. Jälkimmäisen vaihtoehdon etuna on, että parametrien määrä ei kasva rajattomiin aineiston kasvaessa. Käsiteltäessä piilomuuttujaa θ satunnaisena, siitä täytyy tehdä jakaumaoletus, jota kutsutaan populaatiomalliksi. Se on piilomuuttujan tai piilomuuttujavektorin tiheysfunktio. Populaatiomallin määrittäminen on myös välttämättömyys, jotta malli voidaan kirjoittaa ei-ehdolliseksi marginaalitodennäköisyyttä kuvaavaksi marginaaliosioanalyysimalliksi. Seuraavassa on esitetty erilaisia populaatiomalleja.

Yksidimensionaalisisssa malleissa, kuten Raschin mallissakin, kykyä kuvaavan latentin muuttujan θ tiheysfunktio oletetaan usein olevan normaalijakauma (Masters, 1982; Thissen, 1982; Holland, 1990), jolloin

$$f_\theta(\theta;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\theta-\mu)^2}{2\sigma^2}}, \tag{3.3.1}$$

ja tällöin edellistä (3.3.1) vastaavasti

$$\theta = \mu + E, \tag{3.3.2}$$

jossa $E \sim N(0,\sigma^2)$.

Vastaavasti, jos vastauksista halutaan mallintaa useita eri latentteja piirteitä, jolloin piilomuuttuja on vektoriarvoinen. Tällöin (3.3.1):n sijasta populaatiomallina on moniulotteinen normaalijakauma eli formaalisti

$$f_\theta(\theta;\mu,\Sigma) = (2\pi)^{-\frac{d}{2}} |\Sigma|^{-\frac{1}{2}} e^{-1/2[(\theta-\mu)^T \Sigma^{-1} (\theta-\mu)]}. \tag{3.3.3}$$

Nythän keskiarvoja kuvaava parametri μ on vektori ja $\sum$ on thetojen eli piilomuuttujien kovarianssimatriisi.

Nyt, jos osiovastauksista halutaan tutkia vaikkapa sukupuolien välisiä eroja, sosiaaliluokan tai iän yms. yksilötason kiinteiden muuttujien vaikutusta osiovastauksiin, niin yksiulotteisessa tilanteessa populaatiomallin (3.3.1) keskiarvo μ voidaan korvata lineaarisella regressiomallilla. Tällöin populaatiomalli tietylle oppilaalle n on muotoa

$$f_\theta(\theta_n;Y_n,\beta,\sigma^2) = (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}(\theta_n - Y_n^T\beta)^T(\theta_n - Y_n^T\beta)}, \tag{3.3.4}$$

joka voidaan havainnollistavammin ilmaista muodossa

$$\theta_n = Y_n^T\beta + E_n, \tag{3.3.5}$$

jossa jäännökset oletetaan riippumattomiksi $N(0,\sigma^2)$ - jakautuneiksi satunnaismuuttujiksi kuten regressiomallissa yleensä..

Nyt mallissa (3.3.5) vektori Y_n sisältää tunnettujen, kiinteiden muuttujien u arvot oppilaalle n. Kuten tavallisessa regressiomallissakin, β on näitä vastaavien regressiokertoimien vektori.

Malli (3.3.4) laajenee suoraan tilanteeseen, jolloin osiovastauksien oletetaan mittaavan useita eri latentteja piirteitä, jolloin moniulotteiseen tilanteeseen yleistettynä, malli (3.3.4) voidaan kirjoittaa muotoon

$$f_\theta(\theta_n; W_n, \gamma, \Sigma) = (2\pi)^{-d/2}|\Sigma|^{-1/2} e^{-\frac{1}{2}(\theta_n - \gamma W_n)^T \Sigma^{-1}(\theta_n - \gamma W_n)}, \quad (3.3.6)$$

jossa γ on regressiokerrointen matriisi, kokoa $u \times d$, missä d on mallinnettavien latenttien piirteiden lukumäärä. W_n on kiinteiden muuttujien u vektori, kokoa $u \times 1$ ja $\sum$ on $d \times d$ -kovarianssimatriisi.

3.4 Yleistetyn Raschin mallin identifioituvuusehdot

Jotta mallista (3.2.2.2) eli yleistetystä Raschin mallista saadaan identifioituva, sen asetelmamatriisille **A** ja pisteytysmatriisille **B** on asetettava tiettyjä rajoitteita. Seuraavien kolmen rajoitteen voimassaolo on identifioituvuudelle välttämätöntä.

1. Jos mallissa oletetaan olevan D latenttia piirrettä ja osioita kuvaavan parametrivektorin ξ pituuden ollessa P. Lisäksi osiossa i oletetaan mahdollisten vastauskategorioiden lukumäärän olevan $Ki + 1$ ja koko testin τ vastauskategorioiden lukumäärä on

$$K = \sum_{i \in I} Ki,$$

niin mallinnettaessa osiovastauksia testistä τ, jossa on I osiota, yleistetty Raschin malli on identifioituva vain, jos

$$P + D \leq K.$$

Näin siis osioita kuvaavien parametrien ja mallinnettavien latenttien piirteiden summa ei voi olla isompi, kuin latenttia piirrettä indikoivien vastauskategorioiden lukumäärä.

2. Jos edelleen D on oletettujen latenttien piirteiden lukumäärä, joita osioaineistosta mallinnetaan ja P osioparametrivektorin ξ pituus, niin malli on identifioituva ainoastaan, jos seuraavat matriisien **A** ja **B** asteita koskevat ehdot ovat voimassa:

$rank(\mathbf{A}) = P$, $rank(\mathbf{B}) = D$ ja $rank([\mathbf{B}\ \ \mathbf{A}]) = P + D$.

3. Jos yhä D on latenttien dimensioiden lukumäärä ja P osioparametrivektorin ξ pituus, sekä osiossa i oletetaan mahdollisten vastauskategorioiden lukumäärän olevan $Ki + 1$ ja

$$K = \sum_{i \in I} Ki,$$

niin mallinnettaessa osiovastauksia testistä τ, jossa on I osiota, niin malli on identifioituva vain, jos $rank([\mathbf{B}\ \ \mathbf{A}]) = P + D \leq K$.

3.5 Yhteenveto yleistetystä Raschin mallista

Yleistetyn Raschin mallin keskeiset komponentit ovat mallin asetelmamatriisi **A** ja pisteytysmatriisi **B**, joiden erilaisilla formuloinneilla voidaan kirjoittaa useita erilaisia osioanalyysin malleja, joista tässä luvussa esiteltiin esimerkkien valossa Raschin malli, polytomisesti pisteytettyjen osiovastausten malli ja moniulotteinen Raschin malli. Käytännössä mallin yleisestä muodosta voidaan kirjoittaa erilaisia käyttäjän määrittämiä osioanalyysimalleja lähes rajattomasti.

Yleistetty Raschin malli kuvaa todennäköisyyttä tuottaa tietty osiovastaus ehdolla, että mallista riippuen latentti muuttuja tai muuttuja vektori tunnetaan. Jotta malli voidaan ilmaista ei-ehdollisena mallina, niin kykyä kuvaavasta parametrista/parametrivektorivektorista tarvitaan jakaumaoletus. Tätä jakaumaa kutsutaan populaatiomalliksi, joka yleistetyssä Raschin mallissa mahdollistaa myös yksilötason kiinteiden muuttujien vaikutusten mittaamisen suoraan osioanalyysin yhteydessä. Populaatiomallin määritys on välttämätöntä myös parametrien estimoinnissa, jota käsitellään seuraavassa luvussa (ks. Luku 4).

Luku 4

Yleistetyn Raschin mallin parametrien estimointi ja kykyparametrien ennustaminen

Yleistetyn Raschin mallin parametrien estimaattorit ratkaistaan suurimman uskottavuuden menetelmällä. SU-menetelmässä malli johdetaan ei-ehdolliseksi marginaaliosioanalyysimalliksi, konstruoidaan tästä uskottavuusfunktio ja ratkaistaan parametrien estimaattorit. Estimaattien numeeriset ratkaisut saadaan EM-algoritmin (Dempster, Laird & Rubin, 1977) modifikaatiolla (Bock & Aitken, 1981). Lisäksi mallin parametrien SU-estimaattorien lausekkeihin jäävät integraalit approksimoidaan. Integraalit approksimoidaan jakamalla integrointiväli kiinteisiin osaväleihin ja suorittamalla integrointi approksimatiivisesti tai käyttämällä Monte-Carlo-integrointia.

Parametrien keskivirheet estimoidaan SU-menetelmästä tuttuun tapaan laskemalla Fisherin empiirinen informaatiomatriisi, jonka käänteismatriisin diagonaalialkioiden neliöjuuret ovat parametrien estimaattorien keskivirheitä. Parametrien estimaattorien keskivirheet saadaan laskettua myös approksimatiivisesti jättämällä huomioimatta estimaattorien kovarianssit, jolloin keskivirheiden laskeminen helpottuu.

Ei-ehdollinen marginaaliosioanalyysimalli ei sisällä oppilaskohtaisia kykyparametreja, joten estimointialgoritmi ei tuota estimaatteja latentille muuttujalle θ. Tällöin parametrin arvo/arvot ennustetaan. Optimaaliset ennusteet oppilaiden kykyparametreille saadaan laskemalla odotetut aposteriori estimaatit (ks.esim. Bock & Aitken, 1981; Holland, 1990)

Tässä luvussa on aluksi esitetty mallin johtaminen marginaaliosioanalyysimalliksi, josta seuraten on esitetty estimaattorien lausekkeiden johtaminen SU - menetelmin. Tämän jälkeen kuvataan SU-estimaattorien lausekkeiden integraalien numeerisia approksimointimenetelmiä. Luvussa on myös esitetty parametrien estimaattien numeeristen ratkaisujen tuottaminen EM - algoritmilla, jonka jälkeen on kuvattu parametrien keskivirhetarkastelut. Luvun lopussa on esitetty oppilasparametrien ennustamismenetelmä.

4.1 Suurimman uskottavuuden menetelmä

Palautetaan aluksi mieleen ehdollinen yleistetty Raschin malli, joka ilmaisee ehdollisen todennäköisyyden tuottaa tietty vastausvektori testissä τ. Kiinteälle latentille piirteelle θ se oli muotoa

$$f_x(x;\xi|\theta) = \frac{e^{[x^T(B\theta + A\xi)]}}{\sum_{z\in\Omega} e^{[z^T(B\theta + A\xi)]}}. \tag{4.1.1}$$

Palautetaan mieleen myös populaatiossa olevaa vaihtelua mallintava populaatiomalli (3.3.6) , joka on muotoa

$$f_\theta(\theta_n; W_n, \gamma, \Sigma) = (2\pi)^{-d/2}|\Sigma|^{-1/2} e^{[-1/2(\theta_n - \gamma W_n)^T \Sigma^{-1}(\theta_n - W_n)]}. \tag{4.1.2}$$

Yleistetty Raschin mallihan kuvaa ehdollista todennäköisyyttä saada tietty vastausvektori testissä τ, ehdolla, että θ tunnetaan. Malli voidaan johtaa aluksi muotoon, jossa θ ei ole enää ehtona. Tämä tehdään todennäköisyyslaskennan ketjusääntöä käyttäen kombinoimalla yleistetty Raschin malli ja populaatiomalli, jolloin lauseke saadaan muotoon

$$f_x(x,\theta,\xi,\gamma,\Sigma) = f_x(x;\xi|\theta)f_\theta(\theta;\gamma,\Sigma). \tag{4.1.3}$$

Nyt malli johdetaan ei-ehdolliseksi marginaaliosioanalyysimalliksi edellisestä, integroimalla saatu lauseke (4.1.3) yli koko jatkuvan θ:n jakauman, jolloin se saadaan esitettyä ilman θ:aa. Marginaalijakauma on

$$f_x(x;\xi,\gamma,\Sigma) = \int_\theta f_x(x;\xi|\theta)f_\theta(\theta;\gamma,\Sigma)d\theta. \tag{4.1.4}$$

Marginaaliosioanalyysimalli kuvaa nyt todennäköisyyttä havaita tietty vastausvektori X, jonka argumentteina ovat osioparametrivektori, populaatiomallien regressioparametrivektori ja niiden kovarianssimatriisi. Lausekkeesta (4.1.4) johdetaan uskottavuusfunktio, joka saadaan SU-menetelmästä tutulla tavalla. Olettaen oppilaat riippumattomiksi uskottavuusfunktio on muotoa

$$\Lambda = \prod_{n=1}^{N} f_x(x_n;\xi,\gamma,\Sigma), \tag{4.1.5}$$

missä N on otoskoko.

Määritellään ennen estimaattorien johtamista θ_n:n marginaaliposteriori, joka sisältää kaiken empiirisen informaation θ_n:sta eli oppilaskohtaisesta piilomuuttujasta, joka kuvaa kykyä. Toisin sanoen marginaaliposteriori on θ_n:n empiirinen jakauma ehdolla, että oppilaan vastausvektori X_n tunnetaan. Marginaaliposteriori määritetään Bayesin kaavasta

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}, \tag{4.1.6}$$

soveltaen sitä jatkuvalle vektoriarvoiselle muuttujalle θ_n. Tällöin oppilaan n marginaaliposteriori on muotoa

$$h_\theta(\theta_n; W_n, \xi, \gamma, \Sigma | x_n) = \frac{f_x(x_n; \xi | \theta_n) f_\theta(\theta_n; W_n, \gamma, \Sigma)}{f_x(x_n; W_n, \xi, \gamma, \Sigma)}, \tag{4.1.7}$$

jonka osoittajassa on Bayesin kaavaa seuraten ehdollinen yleistetty Raschin malli ja priorijakaumana populaatiomalli. Lausekkeen (4.1.7) nimittäjässä on uskottavuusfunktio, joka saatiin johtamalla yleistetty Raschin malli ei-ehdolliseksi marginaaliosioanalyysimalliksi yhdistäen siihen siihen populaatiomalli ja integroimalla kombinaatio yli koko populaation.

Nyt mallin estimaattorien lausekkeet saadaan derivoimalla uskottavuusfunktion logaritmia

$$\log \Lambda = \sum_{n=1}^{N} \log f_x(x_n; \xi, \gamma, \Sigma) \tag{4.1.8}$$

kunkin mallin parametrin suhteen eli osioparametrivektorin ξ, regressiokerroinvektorin γ ja regressiokerrointen kovarianssimatriisin Σ suhteen ja asettamalla derivaatta nollaksi, josta kunkin parametrin estimaattori voidaan ratkaista. Seuraavassa on esitetty SU - menetelmällä tuotetut estimaattorit ja lyhyesti käyty läpi SU - menetelmän periaate, jolla ne on johdettu.

Kun lauseketta (4.1.8) derivoidaan osioparametrivektorin ξ suhteen ja asetetaan derivaatta nollaksi, eli formaalisti

$$\frac{\partial \log \Lambda}{\partial \xi} = 0, \tag{4.1.9}$$

niin Wu, Adams ja Wilson (1998) ovat osoittaneet, että edellisen ratkaisu tuottaa osioparametrivektorin estimaattien laskemiseksi seuraavan yhtälöryhmän

$$A^T \sum_{n=1}^{N} \left[x_n - \int_{\theta_n} E_z(z|\theta_n) h_\theta(\theta_n; Y_n, \xi, \gamma, \Sigma | x_n) d\theta_n \right] = 0, \tag{4.1.10}$$

jossa

$$E_z(z|\theta_n) = \frac{\sum_{z \in \Omega} z e^{\left[z^T (b\theta_n + A\xi)\right]}}{\sum_{z \in \Omega} e^{\left[z^T (b\theta_n + A\xi)\right]}}. \tag{4.1.11}$$

Edellisessä yhtälöstössä (4.1.10) asetetaan havaitusta vastausvektorista x_n saatu informaatio vastausvektorin ehdollista odotusarvoa vasten, ehdolla, että θ_n tunnetaan. Tähän asiaan palataan EM-algoritmin yhteydessä.

Wu, Adams ja Wilson (1998) ovat niinikään osoittaneet, että derivoimalla uskottavuusfunktion logaritmi (4.1.8) regressiokerroinvektorin suhteen ja asettamalla tämä nollaksi, formaalisti

$$\frac{\partial \log \Lambda}{\partial \gamma} = 0, \tag{4.1.12}$$

ja ratkaisemalla edellinen, tuottaa estimaattorin regressiokerroinvektorin lausekkeen

$$\hat{\gamma} = \left(\sum_{n=1}^{N} \bar{\theta}_n W_n^T \right) \left(\sum_{n=1}^{N} W_n W_n^T \right)^{-1}, \tag{4.1.13}$$

jossa

$$\overline{\theta}_n = \int_{\theta_n} \theta_n h_\theta(\theta_n; Y_n, \xi, \gamma, \Sigma | x_n) d\theta_n, \tag{4.1.14}$$

joka on θ_n:n posteriorikeskiarvo.

Populaatiomallin regressiokerroinvektorin kovarianssimatriisin estimaattori saadaan edelleen Wu:n, Adamsin ja Wilsonin (1998) mukaisesti derivoimalla uskottavuusfunktion logaritmia (4.1.8) ja asettamalla tämä nollaksi eli formaalisti

$$\frac{\partial \log \Lambda}{\partial \Sigma} = 0, \tag{4.1.15}$$

jonka ratkaisu tuottaa regressiokerrointen kovarianssimatriisille estimaattorin, joka on muotoa

$$\hat{\Sigma} = \frac{1}{N} \sum_{n=1}^{N} \int_{\theta_n} (\theta_n - \gamma W_n)(\theta_n - \gamma W_n)^T h_\theta(\theta_n; Y_n, \xi, \gamma, \Sigma | x_n) d\theta_n. \tag{4.1.16}$$

Lauseke (4.1.16) on toisin sanoen posteriorikovarianssien keskiarvo yli populaation.

Edellä esitetyt lausekkeet (4.1.10), (4.1.13) ja (4.1.16) muodostavat implisiittisen "yhtälöstösysteemin", josta parametrien estimaatit ratkaistaan iteratiivisesti EM-algoritmilla (Bock & Aitken, 1981). Ennen EM-algoritmin kuvausta käydään lyhyesti läpi yhtälöissä olevien integraalien numeeriset approksimointimenetelmät, sillä käytännössä niiden suora ratkaisu on usein mahdotonta (Mislevy, 1984).

4.2 Integraalien approksimointi

Uskottavuusyhtälöiden integraalien numeerinen approksimointi voidaan tehdä esimerkiksi kahdella menetelmällä (Mislevy, 1984), jotka tässä kontekstissa esitetään. Nämä ovat numeerinen integrointi jakamalla integraali kiinteisiin osaväleihin (vrt. Simpsonin integroimiskaava, Gaussin integroimismenetelmät, ks. Haataja, Käpyaho & Rahola, 1993) ja Monte Carlo-approksimointi. Seuraavassa on esitetty lyhyesti molemmat approksimointimenetelmät sovellettuna yleistettyyn Raschin malliin. Tällöin menetelmillä approksimoidaan lausekketta (4.1.4).

Jaettaessa integrointiväli kiinteisiin osaväleihin, integraalit θ:n suhteen yli koko jatkuvan θ:n jakauman korvataan seuraavin vaihein:

1) Määritellään θ:n jatkuvasta jakaumasta äärelliseltä väliltä, äärellinen ja kiinteä määrä pisteitä, $q = 1,...,Q$ jokaista jakauman D:tä ulottuvuutta kohden. Koko θ:n moniulotteisesta jakaumasta pisteet määritellään vektorina, jota merkitään

$$\Theta_r = (\Theta_{1j_1}, \Theta_{2j_2}, \ldots, \Theta_{Dj_D}), \tag{4.2.1}$$

jossa

$$j_1 = 1,\ldots,Q; j_2 = 1,\ldots,Q;\ldots; j_D = 1,\ldots,Q.$$

Yhdessä ulottuvuudessa eli ulottuvuuden pysyessä vakiona näitä pisteitä merkitään neliömuodossa vektorilla

$$(\Theta_{d1}, \Theta_{d2}, \ldots, \Theta_{dQ}). \tag{4.2.2}$$

Ja vastaavasti jos moniulotteisessa jakaumassa piste q vakioidaan, niin tällöin muodostuvaa vektoria merkitään

$$\Theta_p = (\Theta_{q1}, \Theta_{q2}, \ldots, \Theta_{qD}). \tag{4.2.3}$$

2) Jokaista vektorin pistettä kohden määritellään sopiva paino $W_q(\gamma, \Sigma)$. Paino määritetään siten, että painotetut pisteet approksimoivat populaatiomallin jakaumaa, joka on muotoa

$$W_r = K(2\pi)^{-d/2}|\Sigma|^{-1/2} e^{[-1/2(\Theta_r-\gamma W_n)^T \Sigma^{-1}(\Theta_r-\gamma W_n)]}, \tag{4.2.4}$$

jossa K on skaalaustekijä, joka määrittää painojen summan ykköseksi. Näin jakauman hännissä olevat pisteet saavat pienemmän painon, kun taas jakauman tiheässä kohdassa painot ovat suurempia. Esimerkkejä integraalien painotaulukoista on annettu yleisesti Haatajan ym. (1993) teoksessa ja yleisestä painojenlaskuperiaatteesta on kerrottu Mislevyn (1984) ja Bockin& Aitkenin (1981) artikkeleissa.

3) Approksimoidaan ei-ehdolliseksi eli marginaaliosioanalyysimalliksi johdettua mallia (4.1.4) seuraavasti

$$f_x(x;\xi,\gamma,\Sigma) = \sum_{p=1}^{Q} f_x(x;\xi|\Theta_p)W_p(\gamma,\Sigma) \tag{4.2.5}$$

ja vastaavasti marginaaliposterioria

$$h_\theta(\Theta_q;W_n,\xi,\gamma,\Sigma|x_n) = \frac{f_x(x_n;\xi|\Theta_q)W_q(\gamma,\Sigma)}{\sum_{p=1}^{Q} f_x(x_n;\xi|\Theta_p)W_p(\gamma,\Sigma)}. \tag{4.2.6}$$

Monte-Carlo-menetelmässä integraalia (4.1.4) approksimoidaan seuraavasti:

1) Simuloidaan M vektoriarvoista funktionarvoa, $\{\varphi_{mn}\}_{m=1}^{M}$, standardoidusta moniulotteisesta normaalijakaumasta jokaiselle oppilaalle n.

2) Approksimoidaan integraalia (4.1.4), joka esiintyy myös marginaaliposteriorin nimittäjässä seuraavasta kaavasta:

$$\int_{\theta} f_x(x;\xi|\theta) f_\theta(\theta,\gamma,\Sigma) d\theta \approx \frac{1}{M}\sum_{m=1}^{M} f_x(x;\xi|\varphi_{mn}) \equiv \Im. \qquad (4.2.7)$$

Ero Monte Carlo-menetelmässä ja kiinteiden osavälien avulla approksimoinnissa on, että jälkimmäisessä pisteet ja vektorit/painot, joita edellisessä kuvattiin ovat kiinteästi määritettyjä, mutta Monte Carlo approksimoinnissa vastaavat vektorit ovat aluksi satunnaisotos standardoidusta moniulotteisesta normaalijakaumasta ja paino on kaikilla vakio, 1/M. Monte Carlo approksimointi on iteratiivinen menetelmä, joka tuottaa lopuksi setin vektoreita, jotka ovat satunnaisotos monimuuttujaisesta normaalijakaumasta keskiarvolla γWn ja varianssilla Σ. Yleisesti ottaen korkeaulotteisempiin integraaleihin käytetään Monte Carlo-menetelmää (ks. haataja, ym., 1993).

Asiasta enemmän kiinnostuneille, numeerisista menetelmistä on saatavilla tilastotiteellistä kirjallisuutta (ks. esim. Haataja ym., 1993) ja niiden sovelluksia on nähtävillä useihin eri konteksteihin. Tämän tutkimuksen yhteydessä asioita ei kuitenkaan enempää käsitellä.

4.3 EM-algoritmi

EM-algoritmin (E = expectation, M = maximization) ehdottivat eksponenttisen jakaumaperheen tapauksessa Dempster, Laird ja Rubin (1977). EM-algoritmia käytetään SU-estimaattien laskemiseksi, jos aineistossa on puuttuvia, aggregoituja, katkaistuja (sensuroituja) havaintoja tai yleisemmin havaitsemattomia muuttujia.

EM-algoritmin keskeisenä ideana on palauttaa SU-yhtälöiden ratkaisu iteratiiviseksi jonoksi "yksinkertaisia" optimointitehtäviä. EM-algoritmi on kaksivaiheinen. Algoritmi sisältää E-askeleen (Expectation step) ja M-askeleen (Maximization step). Seuraavassa on lyhyesti kuvattu EM-algortimin toimintaa yleisellä tasolla.

EM-algoritmi on iteratiivinen proseduuri ja toimii yleisesti ottaen seuraavasti:

1) Asetetaan/lasketaan alkuarvo parametrille, jota estimoidaan. Merkitään tätä $\hat{\lambda}_k$.

2) E-askel: lasketaan parametrin tyhjentävän tunnusluvun t arvo jakaumasta, jossa λ:n arvo on $\hat{\lambda}_k$, jolloin

$$t_k = E_{\hat{\lambda}_k} T. \tag{4.3.1}$$

3) M-askel: Asetetaan havaittu ja odotettu informaatio vastakkain ja maksimoidaan tämä eli ratkaistaan SU-estimaatit $\hat{\lambda}_{k+1}$ yhtälöstä

$$E_\lambda T = \lambda_k \tag{4.3.2}$$

4) palataan askeleeseen 1, kunnes saavutetaan konvergenssi eli

$\left|\hat{\lambda}_k - \hat{\lambda}_{k+1}\right| < \varepsilon$, jossa ε on käyttäjän ennalta määrittelemä pysäytystarkkuus.

EM-algoritmin etuina on, että se on helppo rakentaa. E-askel ohjelmoidaan ja M-askel tehdään rutiininomaisesti optimoimalla pyrkien käyttämään standardiohjelmaa. EM-algoritmin ongelmina on, että se yleisesti ottaen konvergoi hitaasti vaatien paljon iteraatiokierroksia.

Bock ja Aitken (1981) kehittivät EM-algoritmin modifikaation osioparametrien ja populaatiomallin parametrien numeeristen estimaattien tuottamiseksi. Menetelmä ei ole aivan sama kuin EM-algoritmi, mutta siinä on hyvin saman kaltaisia piirteitä. Yleistetyssä Raschin mallissa se toimii seuraavasti:

Askel 1)

Approksimoidaan θn jakaumaa vektoreilla ja niitä vastaavilla painoilla, jotka määritellään riippuen $\gamma^{(t)}$:stä ja $\Sigma^{(t)}$: stä, jotka ovat γ:n ja Σ:n estimaatit iteraatiokierroksella t. (vrt.alkuarvot EM-algoritmissa)

Askel 2)

Lasketaan diskreetti, numeerinen approksimaatio θn:n marginaaliposteriorista iteraatiokierroksella t käyttäen marginaaliposteriorin approksimaatiota

$$h_\Theta(\Theta_q; W_n, \xi^{(t)}, \gamma^{(t)}, \Sigma^{(t)} | x_n) = \frac{f_x(x_n; \xi^{(t)} | \Theta_q) W_q(\gamma^{(t)}, \Sigma^{(t)})}{\sum_{p=1}^{Q} f_x(x_n; \xi^{(t)} | \Theta_p) W_p(\gamma^{(t)}, \Sigma^{(t)})}, \tag{4.3.3}$$

jossa $\xi^{(t)}, \gamma^{(t)}$ ja $\Sigma^{(t)}$ ovat estimaatteja iteraatiokierroksella t (vrt. EM-algortimin E-askel).

Askel 3)

Lasketaan Newton-Raphson-menetelmällä (ks. Luku 2) lausekkeesta

$$A^T \sum_{n=1}^{N} \left[x_n - \sum_{r=1}^{Q} E_z(z|\Theta_r) h_\Theta(\Theta_r; W_n, \xi^{(t)}, \gamma^{(t)}, \Sigma^{(t)} | x_n) \right] = 0 \tag{4.3.4}$$

uudet estimaatit osioparametrivektorille eli $\hat{\xi}^{(t+1)}$ (vrt. EM-algoritmin M-askel).

Askel 4) Estimoidaan populaatiomallin regressioparametrivektorille estimaatit $\hat{\gamma}^{(t+1)}$ kaavasta (EM-algoritmin M-askel jatkuu...)

$$\hat{\gamma}^{(t+1)} = \left(\sum_{n=1}^{N} \overline{\Theta}_n W_n^T \right) \left(\sum_{n=1}^{N} W_n W_n^T \right)^{-1} \tag{4.3.5}$$

ja edelleen regressiokerrointen kovarianssimatriisille $\hat{\Sigma}^{(t+1)}$ kavaasta

$$\hat{\Sigma}^{(t+1)} = \frac{1}{N} \sum_{n=1}^{N} \sum_{r=1}^{Q} (\Theta_r - \gamma^{(t+1)} W_n)(\Theta_r - \gamma^{(t+1)} W_n)^T h_\Theta(\Theta_r; Y_n, \xi^{(t)}, \gamma^{(t)}, \Sigma^{(t)} | x_n) . \tag{4.3.6}$$

Huomataan myös, että kaavassa (4.3.5)

$$\overline{\Theta}_n = \sum_{r=1}^{Q} \Theta_r h_\Theta(\Theta_r; W_n, \xi^{(t)}, \gamma^{(t)}, \Sigma^{(t)} | x_n). \tag{4.3.7}$$

Askel 5) Palataan askeleeseen 1, kunnes

$$\left| \gamma^{(t+1)} - \gamma^{(t)} \right| < \varepsilon \quad \text{ja} \quad \left| \Sigma^{(t+1)} - \Sigma^{(t)} \right| < \varepsilon \tag{4.3.8}$$

eli kunnes algoritmi konvergoi, jolloin saadut estimaatit ovat lopullisia. Kaavassa (4.3.8) ε on käyttäjän ennalta määrittelemä pysäytystarkkuus.

4.4 Parametrien estimaattien keskivirhetarkastelut

Parametrien estimaattien asymptoottiset keskivirheet lasketaan havaitusta Fisherin informaatiomatriisista, jonka käänteismatriisin diagonaalialkioiden neliöjuuret ovat parametrien estimaattorien asymptoottiset keskivirheet. Seuraavassa on esitetty Wu:n, Adamsin ja Wilsonin (1998) esittämät Fisherin havaitun informaatiomatriisin komponentit, jotka ovat uskottavuusfunktion toisia derivaattoja kunkin parametrin suhteen.

Yksiulotteisessa tilanteessa, jossa populaatiomalli on kuten kaava (3.3.4), havaittu informaatiomatriisi on muotoa

$$\mathrm{I} = \begin{bmatrix} \mathrm{I}_{\xi\xi}{}^{T} & \mathrm{I}_{\beta\xi}{}^{\mathrm{T}} & \mathrm{I}_{\sigma^2\xi}{}^{T} \\ \mathrm{I}_{\xi\beta}{}^{\mathrm{T}} & \mathrm{I}_{\beta\beta}{}^{\mathrm{T}} & \mathrm{I}_{\sigma^2\beta}{}^{T} \\ \mathrm{I}_{\xi\sigma}{}^{2} & \mathrm{I}_{\beta\sigma}{}^{2} & \mathrm{I}_{(\sigma^2)}{}^{2} \end{bmatrix}. \tag{4.4.1}$$

Matriisin (4.4.1) komponentit ovat muotoa (ks. Wu, Adams & Wilson 1998):

$$\begin{aligned} \mathrm{I}_{\xi\xi}{}^{T} = & -A^{T}\sum_{n=1}^{N}[\int_{\theta_n} E_z(zz^{T}|\theta_n)h_\theta(\theta_n;Y_n,\hat{\xi},\hat{\beta},\hat{\sigma}^2|x_n)d\theta_n \\ & - 2\int_{\theta_n} E_z(z|\theta_n)E_z(z^{T}|\theta_n)h_\theta(\theta_n;Y_n,\hat{\xi},\hat{\beta},\hat{\sigma}^2|x_n)d\theta_n \\ & + \int_{\theta_n} E_z(z|\theta_n)h_\theta(\theta_n;Y_n,\hat{\xi},\hat{\beta},\hat{\sigma}^2|x_n)d\theta_n \cdot \\ & \int_{\theta_n} E_z(z^{T}|\theta_n)h_\theta(\theta_n;Y_n,\hat{\xi},\hat{\beta},\hat{\sigma}^2|x_n)d\theta_n]A, \end{aligned} \tag{4.4.2}$$

$$\mathrm{I}_{\beta\beta}{}^{T} = \sum_{n=1}^{N}\frac{Y_nY_n{}^{T}}{\hat{\sigma}^2}\left[\frac{\hat{E}_\theta(\theta_n{}^2) - \hat{E}_\theta(\theta_n)^2}{\hat{\sigma}^2} - 1\right], \tag{4.4.3}$$

$$\mathbf{I}_{(\sigma^2)^2} = -\frac{N}{2\hat{\sigma}^4} + \frac{1}{4\hat{\sigma}^8}\sum_{n=1}^{N}[\int_{\theta_n}(\theta_n - Y_n^T\hat{\beta})^4 h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n) d\theta_n$$

$$-\left(\int_{\theta_n}(\theta_n - Y_n^T\hat{\beta})^2 h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n) d\theta_n\right)^2],$$

(4.4.4)

$$\mathbf{I}_{\beta\xi^T} =$$

$$-A^T\sum_{n=1}^{N}\left(\int_{\theta_n}\theta_n\hat{E}_z(z|\theta_n)h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n - \hat{E}_\theta[\hat{E}_z(z|\theta_n)\hat{E}_\theta(\theta_n)]\right)\frac{Y_n^T}{\hat{\sigma}^2},$$

(4.4.5)

$$\mathbf{I}_{\sigma^2\beta^T} =$$

$$-\frac{A^T}{2\sigma^4}\sum_{n=1}^{N}[\int_{\theta_n}\hat{E}_z(z|\theta_n)(\theta_n - Y_n^T\hat{\beta})^3 h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n$$

$$-\int_{\theta_n}\hat{E}_z(z|\theta_n)h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n\int_{\theta_n}(\theta_n - Y_n^T\hat{\beta})^2 h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n]$$

(4.4.6)

ja

$$\mathbf{I}_{\sigma^2\xi^2} =$$

$$-\frac{1}{2\hat{\sigma}^6}\sum_{n=1}^{N}[\int_{\theta_n}Y_n(\theta_n - Y_n^T\hat{\beta})^3 h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n$$

$$-\int_{\theta_n}Y_n(\theta_n - Y_n^T\hat{\beta})h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n \cdot \tag{4.4.7}$$

$$\int_{\theta_n}(\theta_n - Y_n^T\hat{\beta})^2 h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n)d\theta_n].$$

Parametrien estimaattoreiden asymptoottiset varianssit saadaan matriisin (4.4.1) käänteismatriisista, jonka diagonaalialkiot ovat parametrien estimaattorien varianssit ja varianssien neliöjuuret ovat puolestaan estimaattien keskivirheet. Kuitenkin käytännössä asymptoottisten keskivirheiden laskeminen havaitusta informaatiosta on hyvin aikaa vievää ja hankalaa, sillä matriisi (4.4.1) on $(p+r+2)\times(p+r+2)$ - neliömatriisi, jossa p on osioparametrien lukumäärä ja r on regressioparametrien lukumäärä. Lisäksi jokaisen elementin (4.4.2) - (4.4.7) laskeminen edellyttää integrointia yli posteriorijakauman.

Helpompi ja käytännöllisempi ratkaisu on approksimoida estimaattorien asymptoottisia variansseja, joista keskivirheet lasketaan. Approksimaatiot ovat johdettu siten, että ne eivät ota huomioon estimaattorien asymptoottisia kovariansseja. Seuraavassa on esitetty Wu:n, Adamsin ja Wilsonin (1998) johtamat parametrien estimaattorien varianssien approksimaatiot:

$$\begin{aligned}\mathrm{var}(\hat{\xi}_i) = \\ \sum_{n=1}^{N} \{\mathrm{diag}[A^T(\int_{\theta_n} E_z(zz^T|\theta_n)h_\theta(\theta_n;Y_n,\hat{\xi},\hat{\beta},\hat{\sigma}^2|x_n)d\theta_n \\ - \int_{\theta_n} E_z(z|\theta_n)E_z(z^T|\theta_n)h_\theta(\theta_n;Y_n,\hat{\xi},\hat{\beta},\hat{\sigma}^2|x_n)d\theta_n)A]\}^{-1}\end{aligned} \tag{4.4.8}$$

$$\mathrm{var}(\hat{\beta}_i) = \hat{\sigma}^2(\sum_{n=1}^{N} Y_n Y_n^T)^{-1}, \tag{4.4.9}$$

ja

$$\mathrm{var}(\hat{\sigma}^2) = \frac{2\hat{\sigma}^4}{N}. \tag{4.4.10}$$

Osioparametrien asymptoottiset, tässä kontekstissa approksimatiiviset keskivirheet saadaan seuraavasti kaavasta (4.4.8):

$$s.e.(\hat{\xi}_i) = \sqrt{\mathrm{var}(\hat{\xi}_i)}\,. \tag{4.4.11}$$

Regressioparametrien estimaattoreille vastaavasti kaavasta (4.4.9), jolloin

$$s.e.(\hat{\beta}_i) = \sqrt{\mathrm{var}(\hat{\beta}_i)} \tag{4.4.12}$$

ja niin ikään niiden variansseille kaavan (4.4.10) neliöjuuresta eli

$$s.e.(\hat{\sigma}^2) = \sqrt{\mathrm{var}(\hat{\sigma}^2)}\,. \tag{4.4.13}$$

Kuten aikaisemmin mainittiin, niin edellä esitetyt approksimaatiot eivät ota huomioon ollenkaan estimaattien kovariansseja. Niitä tulkittaessa ja raportoitaessa on pidettävä mielessä seuraavat asiat. Kaavasta (4.4.8) lasketut osioparametrien varianssit yleisesti ottaen aliestimoivat otantavirhettä etenkin intervalliasteikollisissa malleissa, joissa pisteytettäviä vastauskategorioita on useita. Kaavojen (4.4.9) ja (4.4.10) tarkkuus riippuu mittausvirheen suuruudesta, koska se vaikuttaa yksilöiden posteriorijakaumien variansseihin.

Moniulotteisessa tilanteessa, jolloin piilomuuttuja θ on vektoriarvoinen, varianssien ja keskivirheiden laskeminen on todella hankalaa ja työlästä. Näiden käsittely ja johtaminen sivuutetaan tässä tutkimuksessa.

4.5 Henkilön kykyparametrin estimointi ja ennustaminen

Muistutetaan mieliin, että marginaaliosioanalyysimalli (4.1.4) ei sisällä parametreja latentille muuttujalle θn, koska se integroidaan siitä pois, jotta mallin parametrit saadaan toisistaan riippumattomiksi. Näin ollen arvot latentille muuttujalle eli oppilasparametrille täytyy ennustaa tai estimoida rajoitetulla SU-menelmällä (REML) , kun osioparametrit ovat tunnetut.

Seuraavassa on esitetty oppilasparametrien ennustamismenetelmä. Menetelmässä lasketaan jokaiselle oppilaalle odotettu aposteriori (EAP, Expected Aposterior) estimaatti posteriorikeskiarvon approksimaatiosta

$$\theta_n^{EAP} = \sum_{r=1}^{Q} \Theta_r h_\Theta(\Theta_r; W_n, \hat{\xi}, \hat{\gamma}, \hat{\Sigma} | x_n), \tag{4.5.1}$$

joka tuottaa yksilölliset kykyä kuvaavat arvot. Ennusteiden varianssi saadaan kaavasta

$$\mathrm{var}(\theta_n^{EAP}) = \sum_{r=1}^{Q} (\Theta_r - \theta_n^{EAP})(\Theta_r - \theta_n^{EAP})^T h_\Theta(\Theta_r; W_n, \hat{\xi}, \hat{\gamma}, \hat{\Sigma} | x_n) \tag{4.5.2}$$

josta ennusteen keskivirhe lasketaan kaavasta

$$s.e.(\theta_n^{EAP}) = \sqrt{\mathrm{var}(\theta_n^{EAP})}\,. \tag{4.5.3}$$

Oppilasparametrien rajoitetut SU-estimaatit (REML-estimates) ratkaistaan maksimoimalla yleistetty Raschin malli

$$f(x;\xi|\theta) = \frac{e^{[x^T(B\theta+A\xi)]}}{\sum_{z\in\Omega} e^{[z^T(B\theta+A\xi)]}} \tag{4.5.4}$$

θn:n suhteen olettaen osioparametrit tunnetuiksi. Tämä tuottaa seuraavan uskottavuusyhtälöryhmän (Wu, Adams, Wilson, 1998):

$$\sum_{i\in\Omega}\left(b_{ixni} - \sum_{j=1}^{K_i} \frac{b_{ij}e^{(b_{ij}\theta n + a_{ij}{}^T\hat{\xi})}}{\sum_{k=1}^{K_i} e^{(b_{ik}\theta n + a_{ik}{}^T\hat{\xi})}} \right) = 0, \tag{4.5.5}$$

jossa $\hat{\xi}$ on osioparametrien estimaattien vektori. Yhtälöryhmä ratkeaa jo tutkielmassa kuvatulla iteratiivisella Newton-Raphson menetelmällä. Yhtälöstä (4.5.5) ei kuitenkaan saada suoraan äärellisiä estimaatteja tapauksissa, joissa oppilas on vastannut kaikkiin osioihin täysin oikein tai täysin väärin. Tässä tilanteessa vastaava parametrin arvo on parametriavaruuden reunapiste ja tällöin SU-estimaattori ei ole olemassa tai se ei ole yksikäsitteinen. Tällaisissa tapauksissa käytetään seuraavaa menettelyä: Oppilaiden, jotka ovat vastanneet testissä kaikkiin osioihin väärin, pisteisiin lisätään pieni vakio k ja vastaavasti kaikki täysin oikein saaneilta oppilailta, heidän pisteistään vähennetään pieni vakio k, jolloin estimaateille saadaan äärelliset arvot.

4.6 Yhteenveto estimoinnista

Yleistetyn Raschin mallin parametrit estimoidaan SU-menetelmällä, jonka tuottamat ratkaisut palautetaan iteratiiviseksi jonoksi yksinkertaisempia optimointitehtäviä, joista numeeriset estimaatit parametreille on helpompi tuottaa (EM-algoritmi). Uskottavuusyhtälöiden muodostaminen ja johtaminen edellyttää kuitenkin marginaaliposteriorin määrittämistä Bayesin teoriasta, sekä yleistetyn Raschin mallin johtamista marginaaliosioanalyysimalliksi.

Parametrien keskivirheet lasketaan Fisherin havaitusta informaatiosta, mutta käytännön tilanteissa ne usein joudutaan approksimoimaan.

Marginaaliosioanalyysimalli ei sisällä parametreja piilomuuttujalle, joka kuvaa oppilaan kykyä, sillä ne integroidaan siitä pois, jotta muut parametrit saadaan riippumattomaksi toisistaan. Näin ollen ne ennustetaan posteriorikeskiarvosta (EAP-menetelmä) tai ratkaistaan SU-menetelmällä, jos osioparametrien estimaatit ovat tunnetut.

Yleisesti ottaen yleistetyn Raschin mallin parametrien estimointi on suhteellisen kompleksinen, sillä se edellyttää tuntemusta osioanalyysimalleista, todennäköisyyslaskennasta, approksimatiivisista integrointimenetelmistä ja EM-algoritmista, mutta näistä on saatavilla paljon tilastotieteellistä kirjallisuutta asiasta kiinnostuneille.

Luku 5

Mallin osioparametrien yhteensopivuus

Yleistetyssä Raschin mallissa tutkitaan yhteensopivuus jokaiselle osioparametrille, joka estimoidaan. Näin ollen yleistetyssä Raschin mallissa tutkitaan erikseen jokaisen osioparametrin yhteensopivuus.
Seuraavassa on esitetty yleistetyn Raschin mallin osioparametrien yhteensopivuuden tutkiminen yksiulotteisessa tilanteessa Wu:n, Adamsin ja Wilsonin (1998) mukaisesti. Moniulotteisessa tilanteessa yhteensopivuustarkasteluja ei keskivirhetarkastelujen tapaan esitetä niiden laskennallisen hankaluuden ja kompleksoituneen johtamisen vuoksi.

Ensinnäkin, mallin yhteensopivuus perustuu standardoituun residuaaliin

$$z_{np}(\theta_n) = (A_p{}^T x_n - E_{np}) \,/\, \sqrt{V_{np}}, \tag{5.1}$$

jossa Ap on asetelmamatriisin p:s sarake. $A_p{}^T x_n$ on p:nnen osioparametrin tyhjentävä tunnusluku oppilaalle n ja E_{np} on sen ehdollinen odotusarvo ja V_{np} vastaavasti sen varianssi.

Painottamaton mallin yhteensopivuustesti (*FIT*) konstruoidaan edellisten residuaalien neliöiden keskiarvosta yli *N*:n ja integroimalla yli posteriorijakaumien, formaalisti

$$FIT_p =$$

$$\int_{\theta_1}\int_{\theta_2}\cdots\int_{\theta_n}\left[\frac{1}{N}\sum_{n=1}^{N}\hat{z}_{np}^{\,2}(\theta_n)\right]\cdot \tag{5.2}$$

$$\prod_{n=1}^{N} h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n) d\theta_N d\theta_{N-1} \ldots d\theta_1.$$

Vastaavasti painotettu yhteensopivuus (*FITw*) lasketaan painottamalla residuaalia (5.1) latentin muuttujan varianssilla, jolloin se lasketaan kaavasta

$$FIT_{w,p} =$$

$$\int_{\theta_1}\int_{\theta_2}\cdots\int_{\theta_N}\left[\frac{\sum_{n=1}^{N}\hat{z}_{np}^{\,2}(\theta_n)V_{np}(\theta_n)}{\sum_{n=1}^{N}V_{np}(\theta_n)}\right]\cdot \tag{5.3}$$

$$\prod_{n=1}^{N} h_\theta(\theta_n; Y_n, \hat{\xi}, \hat{\beta}, \hat{\sigma}^2 | x_n) d\theta_N d\theta_{N-1} \ldots d\theta_1.$$

Testisuureet yhtälöistä (5.2) ja (5.3) ovat osoitettu olevan approksimatiivisesti χ^2-jakautuneita vapausastein 1.

Yhtälöiden (5.2) ja (5.3) integraalit approksimoidaan Monte Carlo menetelmällä (ks. aliluku 4.2). Tällöin testisuure (5.2) muunnetaan muotoon

$$t_p = \left(FIT^{1/3}{}_p - 1 + \frac{2}{(9rN)} \Big/ \left(\frac{2}{(9rN)}\right)^{1/2}\right), \tag{5.4}$$

jossa r on Monte Carlo integroinnin yhteydessä jakaumasta tehtyjen satunnaisten nostojen lukumäärä. Vastaavasti testisuure (5.3) muotoon

$$t_{w,p} = \left[FIT^{1/3}{}_{w,p} - 1\right] \times \frac{3}{Var(FIT_{w,p})} + \frac{\sqrt{Var(FIT_{w,p})}}{3}, \tag{5.5}$$

jossa

$$Var(FIT_{w,p}) = \left(\frac{1}{\sum_n V_{np}}\right)^2 \left(\sum_n E\left((A_p{}^T x_n - E_{np}{}^4) - V_{np}{}^2\right)\right). \tag{5.6}$$

Edellä esitetyt testisuureet ovat approksimatiivisesti t-jakautuneita vapausastein 1.

Laskemalla edellä esitetyt testisuureet voidaan arvioida osiokokeen osioiden osiokohtaista sopivuutta, jolloin huonosti malliin sopivat osiot voidaan jättää analyysista pois, koska ne eivät todennäköisesti mittaa hyvin oppilailta mitattavaa latenttia piirrettä. Tällaisessa tilanteessa osiot ovat usein joko liian vaikeita tai liian helppoja.

Populaatiomallin regressioparametrien yhteensopivuutta ei yleistetyssä Raschin mallissa yleensä erikseen lasketa (ks. Wu, Adams & Wilson, 1998). Regressioparametrien suhteen ollaan usein kiinnostuneita siitä, eroavatko ne toisistaan tilastollisesti merkitsevästi. Esimerkiksi halutaan tietää ovatko tytöt menestyneet osiokokeessa paremmin kuin pojat. Tämän tutkimiseksi käytetään Wu:n, Adamsin ja Wilsonin (1998) johtamaa χ^2 – jakautunutta testiä, jolloin saadaan tietää eroavatko populaatiomallin regressioparametrien arvot toisistaan tilastollisesti merkitsevästi. Testin teoriasta kiinnostuneille, sen johtaminen on esitetty Wu:n, Adamsin ja Wilsonin (1998) kirjoituksessa. Tässä tutkielmassa sen käsittely sivuutetaan.

Luku 6

Pohdinta

Tutkielmassa tarkasteltiin Raschin osioanalyysimallia, sen parametrien estimointimenetelmiä, iteratiivisia algoritmeja ja Raschin mallin lisäparametrisia vaihtoehtoja, jotka johdattelivat tutkimuksen pääasialliseen aiheeseen eli yleistettyyn Raschin malliin.

Yleistetystä Raschin mallista on mahdollista spesifioida lukuisia erilaisia osioanalyysimalleja mallin asetelmamatriisin ja pisteytysmatriisin sopivilla valinnoilla, joita ovat esimerkiksi tässä tutkielmassa havainnollistetut Raschin malli, polytomisesti pisteytettyjen osiovastausten osioanalyysimalli ja moniulotteinen Raschin malli. Yleistetyn Raschin mallin osioparametrien ja populaatiomallin parametrien estimaattorit johdetaan SU-menetelmin ja niihin jäävät integraalit approksimoidaan numeerisin integrointimenetelmin. Lopulliset estimaattien numeeriset ratkaisut saadaan EM-algoritmin modifikaatiolla (Bock & Aitkin, 1981). Oppilaskohtaiset parametrit ennustetaan laskemalla niille odotetut aposterioriestimaatit tai ne voidaan estimoida rajoitetulla SU-menetelmällä. Tutkielmassa kiinnitettiin lisähuomio myös mallin osioparametrien yhteensopivuuden tutkimiseen.

Tavallista Raschin mallia on käytetty koulutustutkimuksessa, kuten myös muiden ihmistieteiden tutkimuksessa jo 60-luvulta lähtien. Tavallisella Raschin mallilla voidaan mitata/mallintaa oppilailta heidän kognitiivinen osaamistaso esti-

moimalla heille oppilaskohtaiset parametrit. Mallilla voidaan myös tutkia osioiden vaikeustaso, jolloin mallin sovittamisen yhteydessä voidaan testata itse osiokoetta, jonka oppilaat ovat suorittaneet. Raschin mallissa yleensä liian hankalat tai liian helpot osiot sopivat huonosti mittaamaan osiokokeella mitattavaa latenttia ominaisuutta oppilailta, eikä ne sovi hyvin malliin. Tällöin ne voidaan jättää osioanalyysista pois ja tutkia oppilaiden osaamistasoa ilman kyseisiä osioita. Raschin malli, sen parametrien estimointi ja yhteensopivuuden tutkiminen voidaan suorittaa empiirisesti usealla eri valmisohjelmistolla, joita ovat esimerkiksi SAS, PLM, GLIM ja S+. Aineisto tulee olla koodattu dikotomisesti siten, että oppilaat ovat havaintomatriisin vaakariveillä ja eri osiot havaintomatriisin sarakkeissa. Ohjelmistoista ja mallien parametrien tulkinnoista on suhteellisen tyhjentävästi kerrottu esimerkiksi Gustafssonin (1977) julkaisussa ja Hambletonin & Swaminathanin (1985) teoksessa.

Yleistetyn Raschin mallin sovittaminen aineistoon, sekä sen parametrien estimointi ja oppilasparametrien ennustaminen voidaan tehdä toistaiseksi ainoastaan ConQuest-ohjelmistolla, joka on vielä "prototyyppiasteella". Ohjelmiston käytännön toimivuus on toistaiseksi huono ja ohjelmisto on hidas vaatien lukuisia EM-algoritmin iterointikierroksia. Näin ollen tämän tutkielman pääasialliseksi tarkoitukseksi jää osioanalyysiin, Raschin malliin ja yleistettyyn Raschin malliin liittyvän teorian suhteellisen tyhjentävä tarkastelu, valitettavasti ilman empiirisiä tarkasteluja. ConQuest-ohjelmistolla empiirisesti mallien sovittamisesta ja parametrein tulkinnoista kiinnostuneille kehotan tutustumaan Wu:n, Adamsin ja Wilsonin (1998) kirjoitukseen, vaikkakin ohjelmisto vaatii vielä kehittämistä ja käyttäjien palautetta sen kehittämiseksi. Polytomisesti koodattujen vastausten analyysiin toistaiseksi toimiva toimiva ohjelma on BIGSTEP ja moniulotteisen Raschin mallin sovittamiseen pienillä aineistoilla soveltuu LOGIMO (ks. Kelderman & Rijkes, 1994). Kyseiset ohjelmistot eivät käytä kuitenkaan tässä tutkielmassa havainnollistettua yleistettyä Raschin mallia, vaan valmiiksi kyseisiä spesifejä malleja ja erilaisia estimointitekniikoita ja mallin parametrien yhteensopivuusindeksejä.

Yleistetty Raschin malli on kokonaisuudessaan kuitenkin hyvin monipuolinen, idealtaan ja teorialtaan toimiva kokonaisuus. Sen asetelmamatriisi ja pisteytysmatriisi, jotka käyttäjä voi itse määrittää, avaavat tien lukuisten erilaisten latentin piirteen teoriassa tunnettuihin osioanalyysimalleihin. Diskreettien osiovastausten monipiirreanalyysi tarjoaa myös jatkotutkimusongelmia ja kehitysaiheita.

Ensimmäisenä kehitys- ja jatkotutkimusaiheena olisi parametrien estimointi Bayesilaisittain. Toisena varsin relevanttina kehitysaiheena on esimerkiksi tässä

tutkielmassa esitetyn populaatiomallin (3.3.4) korvaaminen lineaarisella sekamallilla (koulutustutkimuksessa monitasomalli, ks. esim. Goldstein, 1987, Goldstein, 1995), jolloin osiovastauksista voitaisiin koulutustutkimuksessa mallintaa oppilastason lisäksi luokkatason tai koulutason muuttujien vaikutuksia osiovastauksiin.

Luku 7

Kiitosmaininnat

Erityiset kiitokset tätä pro gradu-tutkielmaa ohjanneelle FT Antti Penttiselle arvokkaista vinkeistä ja kommenteista, jotka mahdollistivat tutkielman valmistumisen hyvässä aikataulussa. Lisäksi kiitokset Jyväskylän yliopiston koulutuksen tutkimus laitoksen henkilökunnasta Pekka Kuparille, Antero Malinille ja Kari Törmäkankaalle kiinnostavan tutkimusaiheen "virittämisestä", tutkielmassa käytetyn materiaalin käyttöön antamisesta ja yhteistyöstä.

Lähteet

Andersen, E.B. (1973a). Conditional Inference in Multiple Choice Questionaries. *Brittish Journal of Mathematical and Statistical Psychology*, **26**, 31-44.

Andersen, E.B. (1973b). *Conditional Inference and Models for Measuring.* Mentalhygienisk Forlag, Köpenhagen.

Andersen, E.B. (1980). *Discrete Statistical Models with Social Science Applications.* North-Holland, Amsterdam.

Bartholomew, D.J.(1987). *Latent Variable Models and Factor Analysis.* Oxford University Press, New York.

Bock, R.D. & Aitkin, M.(1981). Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm. *Psychometrica*, **46**, 443-459.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society*, series B, **39**, 1-38.

Goldstein, H. (1987). *Multilevel Models in Educational and Social Research.* Oxford University Press, New York.

Goldstein, H. (1995). *Multilevel Statistical Models.* 2nd edition. Oxford University Press, New York.

Gustafsson, J.E. (1977). The Rasch Model for Dichotomous Items: Theory, Applications and a Computer Program. Reports from the Institute of Education, University of Göteborg.

Gustafsson, J.E. (1978). The Rasch Model for Dichotomous Items: A Solution of the Conditional Estimation Problem for Long Tests and some Thoughts about Item Screening Procedures. Reports from the Institute of Education, University of Göteborg.

Gustafsson, J.E. (1979). Testing and Obtaining Fit of Data to the Rasch Model. Reports from the Institute of Education, University of Göteborg.

Gustafsson, J.E. (1981). An Introduction to Rasch Measurement Model. Department of Education, University of Göteborg.

Haataja, J., Käpyaho, J. & Rahola, J. (1993). *Numeeriset menetelmät*. CSC-Tieteellinen laskenta Oy.

Hambleton, R.k. & Swaminathan, H. (1985). *Item Response Theory. Principles and Applications.* Kluwer-Nijhoff Publishing, Boston.

Holland, P.W. (1990). On the Sampling Theory Foundations of Item Response Theory Models. *Psychometrica*, **55**, 577-601.

Kelderman, H. & Rijkes, C.P. (1994). Loglinear Multidimensional IRT Models for Polytomously Scored Items. *Psychometrica*, **59**, 149-176.

Masters, G.N. (1982). A Rasch Model for Partial Credit Scoring. *Psychometrica*, **47**, 149-174.

Mislevy, R.J. (1984). Estimating Latent Distributions. *Psychometrica*, **49**, 359-381.

Mislevy, R.J. (1991). Randomization Based Inference about Latent Variables from Complex Samples. *Psychometrica*, **56**, 177-196.

Rasch, G. (1960). *Probabilistic Models for some Intelligence and Attainment Tests,* Denmarks Pedagogiska Institut, Copenhagen.

Searle, S.R., Casella, G. & McCulloch, C.E. (1992). *Variance Components.* John Wiley & Sons, New York.

Thissen, D. (1982). Marginal Maximum Likelihood Estimation for the One Parameter Logistic Model. *Psychometrica*, **47**, 175-186.

Tjur, T. (1980). A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model. *Scandinavian Journal of Statistics*, **9**, 23-30.

Wright, B.D. (1977). Solving Measurement Problems with the Rasch Model. *Journal of Educational Measurement*, **14**, 97-116.

Wu, M.L., Adams, R.J., Wilson, M.R. (1998). ConQuest. Generalized Item Response Modelling Software. ACER - Australian Council for Educational Research.