

# Järjestysregressiomalli huolto- ja asumisriidoille hovioikeuksissa

Aku Leivonen

17. huhtikuuta 2020



HELSINGIN YLIOPISTO  
HELSINGFORS UNIVERSITET  
UNIVERSITY OF HELSINKI

MATEMAATTIS-LUONNONTIETEELLINEN TIEDEKUNTA  
MATEMATISK-NATURVETENSKAPLIGA FAKULTETEN  
FACULTY OF SCIENCE

Tiedekunta – Fakultet – Faculty		Koulutusohjelma – Utbildningsprogram – Degree programme	
Matemaattis-luonnontieteellinen		Tilastotieteen koulutusohjelma	
Tekijä – Författare – Author			
Aku Leivonen			
Työn nimi – Arbetets titel – Title			
Järjestysregressiomalli huolto- ja asumisriidoille hovioikeuksissa			
Työn laji – Arbetets art – Level		Aika – Datum – Month and year	
Pro gradu -tutkielma		Huhtikuu 2020	
		Sivumäärä – Sidoantal – Number of pages	
		71 sivua + 27 sivua liitteitä	
Tiivistelmä – Referat – Abstract			
<p>Oikeudellisten ilmiöiden tilastollinen mallintaminen on vielä harvinaista Suomessa. Tutkielmassa mallinnetaan lasten huolto- ja asumisriitoja hovioikeuksissa järjestysregressiomallilla. Oikeuden päätökset huolto- ja asumisriidoissa voidaan luokitella järjestysasteikolla, missä asteikon toinen pää käsittää äidin voittoluokat ja toinen isän voittoluokat. Mallin tarkoitus on löytää keskeiset oikeuden päätöstä sekä äitien ja isien voittomahdollisuuksia selittävät tekijät. Mallia voidaan käyttää myös uuden riidan lopputuloksen ennustamiseen. Koska osapuolten vaatimukset rajoittavat sen, millaisen päätöksen oikeus voi antaa, niin mallia laajennetaan skaala- ja nominaalivaikutuksilla. Vastaavaa laajennettua järjestysregressiomallia ei ole sovellettu huolto- ja asumisriitoihin aiemmin.</p> <p>Työssä käytetty oikeustapausaineisto käsittää 500 huolto- ja asumisriitaa Suomen jokaisesta viidestä hovioikeudesta vuosien 2000 ja 2016 väliseltä ajalta. Aineisto on rajattu riitoihin, joissa äiti ja isä ovat eri mieltä lasten asumisesta ja mahdollisesti myös huollosta. Aineiston laajuus ja ilmiön vähäinen mallinnushistoria tekevät aineistosta kansainvälisesti ainutlaatuisen. Tiedot riidoista on poimittu hovioikeuksien ratkaisuista sekä niiden tausta-asiakirjoista. Kaikista riidoista on poimittu tieto myös vastaavasta käräjäoikeuden ratkaisusta, josta äiti tai isä on valittanut hovioikeuteen. Äidit ja isät ovat valittajina yhtä usein, mutta äidit saavat isiä hieman useammin vaatimuksiansa vastaavia päätöksiä hovioikeudessa. Äidit vaativat lasten yksinhuoltoa useammin kuin isät.</p> <p>Aineistossa on paljon muuttujia, joiden yhdistelmistä pyritään muodostamaan mahdollisimman hyviä malleja erilaisiin lähtökohtiin. Selittäjiä tarkastellaan aluksi yhden selittäjän malleilla, mutta varsinaiset analyysit perustuvat usean selittäjän malleihin, jotka muodostetaan yhden selittäjän mallien pohjalta. Usean selittäjän mallien valitsemisessa käytetään tavanomaisia tilastollisten mallien mallinvalintamenetelmiä. Lopputuloksena saadaan kolme mallia, joista ensimmäisen on tarkoitus löytää keskeiset oikeuden päätöstä selittävät tekijät. Toinen malli pyrkii ennustamaan uuden riidan lopputuloksen ja kolmas ennustamaan lopputuloksen riidassa, josta on käräjäoikeuden päätös. Kolmas malli pyrkii myös löytämään tekijät, jotka parhaiten selittävät käräjäoikeuden päätöksen muuttumista.</p> <p>Lasten vakiintunut asuinpaikka, sosiaaliviranomaisten esittämä suositus ja oikeuden todeksi katsoma väkivalta-, päihteidenkäyttö- tai mielenterveyspsykiatrian toisesta osapuolesta ovat merkityksellisimmät oikeuden päätöstä selittävät tekijät. Näiden tekijöiden merkityksen suuruus ei näytä riippuvan siitä, onko kyseessä äiti vai isä. Lasten vakiintunut asuinpaikka on useammin äidin kuin isän luona, mikä selittää sitä, miksi äidit voittavat riitoja hieman useammin kuin isät. Mitä suurempi vanhempien välinen ikäero on, sitä paremmat ovat nuoremman osapuolen voittomahdollisuudet. Mikäli vanhin lapsi on alle kouluikäinen, niin äidin voittomahdollisuudet paranevat. Mikäli äidin avustaja oikeudessa on mies tai mikäli äiti on työtön, niin isän voittomahdollisuudet paranevat.</p> <p>Hovioikeus muuttaa käräjäoikeuden päätöstä asumisriidoissa noin 14 prosentissa valituksista. Koska käräjä- ja hovioikeuden päätökset ovat hyvin vahvasti kytkeytyneet toisiinsa, niin päätöksen muuttumiselle jää vain vähän selittäviä tekijöitä. Tällaisia ovat edeltävä huolto- ja asumistilanne ja vanhempien välinen ikäero.</p> <p>Järjestysregressiomalli toimii ilmiön mallintamisessa melko hyvin ja erityisesti skaala- ja nominaaliselittäjien käyttäminen osoittautuu toimivaksi tavaksi huomioida osapuolten vaatimukset. Mallien ennusteet ovat lupaavia, vaikka ennustekykä on arvioitu vain aineistolla, jolla malli on sovitettu. Aineiston suuren muuttujamäärän vuoksi kaikkia yhteyksiä on vaikea havaita, minkä vuoksi aineistosta olisi mielenkiintoista tehdä vielä useita lisätarkasteluja.</p>			
Avainsanat – Nyckelord – Keywords			
tilastotiede, oikeustilastotiede, huoltoriita, asumisriita, järjestysregressio, järjestysasteikko, sijaintiskaalamalli			
Säilytyspaikka – Förvaringställe – Where deposited			
Muita tietoja – Övriga uppgifter – Additional information			

# Sisältö

<b>1</b>	<b>Johdanto</b>	<b>1</b>
<b>2</b>	<b>Kertymä-logit-regressiomalli</b>	<b>4</b>
2.1	Multinomijakauma . . . . .	4
2.2	Kertymätodennäköisyydet ja kertymävastasuhde . . . . .	5
2.3	Mallin määrittely . . . . .	5
2.4	Piilomuuttuja . . . . .	7
2.5	Mallin laajentaminen skaalavaikutuksilla . . . . .	8
2.6	Mallin laajentaminen nominaalivaikutuksilla . . . . .	8
2.7	Parametrien estimointi . . . . .	9
2.7.1	Tavallinen malli . . . . .	9
2.7.2	Laajennettu malli . . . . .	11
2.8	Mallin parametrien tulkinta . . . . .	12
<b>3</b>	<b>Diagnostiikka ja mallin valinta</b>	<b>14</b>
3.1	Parametrien luottamusvälit ja testit . . . . .	14
3.2	Yhteensopivuusindeksi . . . . .	16
3.3	Yhteiskorrelaatiokerroin . . . . .	17
3.4	Selittävien muuttujien valinta . . . . .	17
3.4.1	Akaiken informaatiokriteeri . . . . .	18
3.4.2	Mallinvalinta-algoritmit . . . . .	18
<b>4</b>	<b>Aineisto ja muuttujat</b>	<b>20</b>
4.1	Päivämäärät ja ikämuuttujat . . . . .	22
4.2	Ulkomaalaistaustaisuus . . . . .	23
4.3	Edeltävä huolto- ja asumistilanne sekä tilanteen vakiintuminen . . . . .	24
4.4	Osapuolten vaatimukset . . . . .	25
4.5	Oikeuden päätös ja riidan voittaja . . . . .	26
4.6	Olosuhdeselvityksen suositus . . . . .	27
4.7	Ratkaisijoiden ja avustajien tiedot . . . . .	28
4.8	Lasten tiedot . . . . .	29

4.9	Vanhempien taustatiedot . . . . .	29
4.10	Osapuolten esittämät syytökset . . . . .	29
4.11	Esimerkki muuttujien kirjaamisesta . . . . .	30
<b>5</b>	<b>Ilmiön mallintaminen</b>	<b>32</b>
5.1	Mallin käyttöperuste ja oletukset . . . . .	32
5.2	Vastemuuttujan määrittely . . . . .	33
5.3	Osapuolten vaatimusten huomioiminen . . . . .	35
5.4	Mallien tulkitsemisesta . . . . .	39
5.5	Selittäjien tarkastelu yhden selittäjän malleilla . . . . .	39
<b>6</b>	<b>Usean selittäjän mallit</b>	<b>45</b>
6.1	Juurimalli . . . . .	45
6.2	Eri lähtökohtiin sovitettut mallit . . . . .	50
6.2.1	Ensimmäinen malli . . . . .	50
6.2.2	Toinen malli . . . . .	52
6.2.3	Kolmas malli . . . . .	55
6.3	Malleilla ennustaminen . . . . .	58
6.3.1	Uuden riidan lopputuloksen ennustaminen . . . . .	59
6.3.2	Lopputuloksen ennustaminen käräjäoikeuden ratkaisemassa riidassa, josta ollaan valittamassa hovioikeuteen . . . . .	60
6.3.3	Ennustekyvyn arviointi . . . . .	61
<b>7</b>	<b>Tulosten pohdintaa</b>	<b>64</b>
7.1	Oikeuden päätöstä selittävät tekijät . . . . .	65
7.1.1	Keskeisimmät taustatekijät . . . . .	65
7.1.2	Syytökset . . . . .	65
7.1.3	Ikätekijät . . . . .	66
7.1.4	Muut tekijät . . . . .	67
7.2	Käräjäoikeuden päätöksen muuttumista selittävät tekijät . . . . .	68
7.3	Lopuksi . . . . .	68
	<b>Lähteet</b>	<b>70</b>
	<b>Liitteet</b>	<b>72</b>
<b>A</b>	<b>Oikeustapausaineiston muuttujat</b>	<b>72</b>
<b>B</b>	<b>Analyyseissa käytetyt muuttujat</b>	<b>82</b>
<b>C</b>	<b>R-koodit ja tulosteet</b>	<b>89</b>
C.1	Juurimalli . . . . .	89
C.2	Ensimmäinen malli . . . . .	92

C.3	Toinen malli . . . . .	95
C.4	Kolmas malli . . . . .	97

# Luku 1

## Johdanto

Lasten huolto- ja asumisriitoja ei ole Suomessa juurikaan tutkittu kvantitatiivisin menetelmin. Huolto- ja asumisriitoja itsessään ovat tutkineet muun muassa Valkama ja Laasola (2008), joilla oli käytössään oikeustapausaineisto huolto- ja asumisriidoista Suomen kaikista hovioikeuksista vuodelta 2006. Palo-Repo (2015) käytti pro gradu -tutkielmassaan aineistoa Helsingin hovioikeuden huolto- ja asumisriitapäätöksistä vuosilta 2003–2006, joissa hän mallinsi äidin ja isän voittomahdollisuuksia huolto- ja asumisriidoissa logistisella regressiomallilla.

Oikeustilastotiede, eli tilastotieteen soveltaminen oikeudellisiin ilmiöihin on yleisesti vielä varsin vähäisesti hyödynnettyä Suomessa. Tilastollisesti mallinnettavia kiinnostavia oikeudellisia ilmiöitä voivat olla esimerkiksi rikosoikeudenkäynneissä tuomion suuruus, tai riitaoikeudenkäynneissä riidan voittaja esimerkiksi dikotomisena muuttujana tai järjestysasteikollisena muuttujana (voitto/tasapeli/tappio).

Tässä tutkielmassa sovelletaan järjestysregressiomallia oikeuden päätösten mallintamiseen huolto- ja asumisriidoissa hovioikeuksissa. Mallilla pyritään selittämään paitsi kummankin osapuolen voittotodennäköisyyttä, myös todennäköisyyttä erilaisille oikeuden päätöksille, joita voidaan huolto- ja asumisriidassa mallintaa järjestysasteikollisella muuttujalla, missä järjestysasteikon toinen pää käsittää äidin voittoluokat ja toinen isän voittoluokat. Järjestysregressiomalleja on erilaisia, ja tässä työssä käytetään kertymälogit-regressiomallia, jota voi pitää logistisen regressiomallin yleistykseenä järjestysasteikolliselle vastemuuttujalle. Koska osapuolten vaatimukset rajoittavat sen, millaisen päätöksen oikeus voi antaa, niin mallista käytetään laajennettua versiota, johon on otettu vaikutteita erityisesti Christensenin (2018a) artikkelista. Itse analyysit on toteutettu R-ohjelmiston ordinal-paketilla (Christensen 2018b), jonka edellä mainittu artikkeli myös esittelee.

Palo-Repo (2015) sovelsi pro-gradu -tutkielmassaan tavallista logistista regressiomallia huolto- ja asumisriitoihin Helsingin hovioikeudessa luokitellen oikeuden päätöksen äidin voitoksi tai tappioksi. Hän muun muassa havaitsi, että Helsingin hovioikeudes-

sa oli vuosina 2003–2006 poikkeuksellisen paljon äidin yksinhuoltopäätöksiä suhteessa kaikkiin hovioikeuksiin vuonna 2006. Äidin voittoa huolto- ja asumisriidoissa selittivät parhaiten lasten edellinen asuinpaikka, sosiaaliviranomaisten suositus sekä lasten ikä. Tutkielman ongelmana oli Palo-Revon mukaan havaintomäärän vähyys, joka saattoi hänen mukaansa aiheuttaa sen, että mahdollisesti tärkeät selittäjät eivät nousseet mallissa esiin tilastollisesti merkitsevinä. Tässä tutkielmassa pyritään osaltaan myös jatkamaan Palo-Revon tutkielmaa, jonka loppusanoissa todettiin, että oikeuden päätösten aikajakson laajentamisella sekä muiden hovioikeuksien päätösten mukaan ottamisella voitaisiin jatkaa tutkimusta. Voi olla kiinnostavaa tutkia myös sitä, eroavatko päätökset toisistaan hovioikeuksien välillä.

Pere, Lahti ja Sutela (2017) mallinsivat käräjäoikeuksien tuomioita törkeissä rattijuopumusrikoksissa kertymä-logit-regressiomallilla tutkien eri selittäjien kuten esimerkiksi tuomarin ja syyttäjän sukupuolten vaikutuksia tuomion suuruuteen suhteessa syyttäjän vaatimukseen. He havaitsivat, että niin tuomarin, syytetyn kuin syyttäjän sukupuoli on yhteydessä annettuun tuomioon. Tässä tutkielmassa pyritään osaltaan myös tutkimaan tuomarien, riidan osapuolten ja osapuolten avustajien mahdollisia sukupuolivaikutuksia annettuihin oikeuden päätöksiin.

Tämän tutkielman aineisto on 500 huolto- ja asumisriitaa, jotka on kerätty Suomen jokaisesta viidestä hovioikeudesta. Päätösten aikajaksona on käytetty vuosia 2000–2016. Vastaavan kaltaista koko Suomen kattavaa aineistoa ei tiettävästi ole koskaan kerätty näin laajalta ajanjaksolta. Aineisto on myös kansainvälisesti ainutlaatuinen, koska huolto- ja asumisriitoja ei juurikaan ole mallinnettu tilastollisesti edes kansainvälisesti. Aineiston keruussa rajoituttiin tutkimaan vain asumisriitoja, jotka ovat mielenkiintoisempia kuin puhtaat huoltoriidat, eli sellaiset riidat, joissa osapuolet ovat yksimielisiä lasten asumisesta ja erimielisyyden on vain lasten huollosta (yhteishuolto/yksinhuolto). Asumisriidoissa osapuolet riitelevät lasten asumisesta ja mahdollisesti myös huollosta. Sen lisäksi, että asumisriidat ovat huoltoriitoja kiinnostavampia, aineiston rajaamisella vältetään se, että malleja ei tarvitse sovittaa erikseen huolto- ja asumisriidoille. Tällöin mallien havaintomäärät saadaan suuremmiksi.

Palo-Repo (2015) kuvailee tutkielmassaan lapsen huoltoon ja asumiseen liittyvät oikeuskäytännöt, joten niihin ei tässä tutkielmassa paneuduta. Laki lapsen huollosta ja tapaamisoikeudesta (361/1983) ei ole merkittävästi muuttunut tässä tutkielmassa käytettyjen päätösten aikajaksolla. Sen sijaan vuodesta 2003 vuoteen 2011 hovioikeuksissa käytössä ollut seulontamenettely on 1.1.2011 voimaan tulleella lailla korvattu jatkokäsittelyluvalla (Laki oikeudenkäymiskaaren muuttamisesta 650/2010). Lisäksi jatkokäsittelylupajärjestelmä laajeni entisestään lokakuussa 2015, jolloin jatkokäsittelyluvan piiriin tulivat kaikki riita- ja hakemusasiat (Laki oikeudenkäymiskaaren muuttamisesta 386/2015). Ennen 1.10.2003 käytössä ei ollut seulontamenettelyä (Laki oikeudenkäymiskaaren muuttamisesta 381/2003). Tässä tutkielmassa käytetyn aineiston päätösten aikajakso kattaa kaikki edellä mainitut ajankohdat. Luvussa 4 pohditaan lyhyesti oi-

keuskäytäntöjen muutosten mahdollisia vaikutuksia analyyseihin.

Tutkielma keskittyy työssä käytetyn tilastollisen mallin määrittelyyn, sovittamiseen, diagnostiikkaan ja arviointiin sekä aineiston kuvailuun. Malli määritellään laajennuksineen luvussa 2. Luvussa 3 käydään läpi mallin diagnostiikkaan ja valintaan liittyvää teoriaa. Luvussa 4 käydään tarkemmin läpi tutkielmassa käytetty aineisto ja sen kerääminen. Luvussa 5 käsitellään ilmiön mallintamista, määritellään mallin vastemuuttuja sekä tutkitaan osapuolten vaatimusten vaikutusta mallin toimivuuteen. Lisäksi tehdään alustavia yhden selittäjän tarkasteluja. Luvussa 6 sovitetaan usean selittäjän malleja erilaisiin lähtökohtiin. Analyysien tulokset ja johtopäätökset kootaan yhteen luvussa 7.

Liitteisiin on koottu alkuperäisen aineiston muuttujalista sekä analyyseissa käytetyt muuttujamuunnokset. Lisäksi mukaan on liitetty keskeisimmät R-koodit sekä ohjelman tuottamat analyyseihin liittyvät tulosteet.



## Luku 2

# Kertymä-logit-regressiomalli

Kun mallinnettava satunnaismuuttuja eli vastemuuttuja on dikotominen, eli mikäli sillä on vain kaksi mahdollista tulosluokkaa, niin ilmiötä mallinnetaan usein logistisella regressiomallilla. Logistinen regressiomalli voidaan kuitenkin yleistää myös useampiluokkaiselle vastemuuttujalle. Tässä työssä kiinnostuksen kohteena on järjestysasteikollinen satunnaismuuttuja, jolla on  $c$  luokkaa, missä  $c > 2$ . Järjestysasteikollisuus tarkoittaa sitä, että satunnaismuuttujan arvot voidaan asettaa suuruusjärjestykseen.

### 2.1 Multinomijakauma

Tarkastellaan  $n$ -kertaista toistokoetta, jossa yksittäisellä toistolla on enemmän kuin kaksi mahdollista tulosluokkaa. Kun toistot ovat toisistaan riippumattomia ja tulosluokkien todennäköisyydet eivät muutu toistojen välillä, niin onnistumisten lukumääriä eri luokissa kuvaava satunnaisvektori  $\mathbf{Y}$  on multinomijakautunut.

Olkoon mahdollisten tulosvaihtoehtojen määrä  $c$ , jolloin satunnaisvektorissa  $\mathbf{Y}$  on  $c$  komponenttia. Merkitään kunkin luokan todennäköisyyksiä  $\pi_1, \dots, \pi_c$ , joille pätee  $\sum_{i=1}^c \pi_i = 1$ . Kun toistoja tehdään  $n$  kappaletta, niin todennäköisyys, että näistä  $n_1$  kappaletta kuuluu luokkaan 1,  $n_2$  kappaletta kuuluu luokkaan 2, ja niin edelleen, ja  $n_c$  kappaletta kuuluu luokkaan  $c$ , on

$$\begin{aligned} P(\mathbf{Y} = \mathbf{y}) &= P(Y_1 = n_1, Y_2 = n_2, \dots, Y_c = n_c) \\ &= \left( \frac{n!}{n_1! n_2! \dots n_c!} \right) \pi_1^{n_1} \pi_2^{n_2} \dots \pi_c^{n_c}, \end{aligned}$$

kun  $\sum_{i=1}^c n_i = n$ . Binomijakauma on multinomijakauman erikoistapaus, kun  $c = 2$ . (Esim. Agresti 2007, 5.)

## 2.2 Kertymätodennäköisyydet ja kertymävastasuhde

Oletetaan, että satunnaismuuttuja  $Y$  on järjestysasteikollinen ja sillä on luokat  $1, \dots, J$  tässä järjestyksessä. Tällöin  $n$ :ää satunnaismuuttujan  $Y$  havaittua arvoa vastaavat lukumäärät luokissa  $1, \dots, J$  on multinomijakautunut satunnaisvektori  $\mathbf{Y} = (Y_1, \dots, Y_J)$ , kun havainnot ovat toisistaan riippumattomia. Merkitään luokkien tuntemattomia todennäköisyyksiä eli multinomijakauman parametreja kirjaimin  $\pi_1, \dots, \pi_J$ . Kiinnostuksen kohteena on näiden parametrien, eli kunkin luokan todennäköisyyden estimointi.

Satunnaismuuttujan  $Y$  järjestysasteikollisuuden vuoksi sille voidaan määritellä kertymätodennäköisyys luokassa  $j$

$$\gamma_j = P(Y \leq j) = \pi_1 + \dots + \pi_j, \text{ missä } j = 1, \dots, J,$$

eli todennäköisyys, että  $Y$  kuuluu luokkaan  $j$  tai kyseisellä järjestysasteikolla tätä alempaan luokkaan. Kun oletetaan, että yksikään luokka ei ole turha, niin todennäköisyyksille pätee

$$0 < \gamma_1 < \gamma_2 < \dots < \gamma_J = 1.$$

Mallintamisen kannalta viimeistä todennäköisyyttä  $\gamma_J$  ei tarvita, sillä järjestysasteikon suurimman luokan kohdalla kertymätodennäköisyys on aina 1.

Määritellään kertymävastasuhde (*cumulative odds*) luokassa  $j$

$$\frac{P(Y \leq j)}{1 - P(Y \leq j)} = \frac{\pi_1 + \dots + \pi_j}{\pi_{j+1} + \dots + \pi_J}, \quad j = 1, \dots, J - 1.$$

Jokaista luokkaa  $j$  vastaavassa kertymävastasuhteessa ilmenevät siten kaikkien vastemuuttujan luokkien todennäköisyydet. (Esim. Agresti 2010, 44–45.)

## 2.3 Mallin määrittely

Oletetaan, että aineisto koostuu järjestysasteikollisen vastemuuttujan satunnaisista havaituista arvoista  $Y_1, \dots, Y_n$  ja tunnetuista selittäjistä  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_p)$ , missä  $x_k = (x_{1k}, \dots, x_{nk})^\top$ , kun  $k = 1, \dots, p$ . Kertymä-logit-regressiomalli kuuluu yleistettyjen lineaaristen mallien joukkoon. Yleistetyssä lineaarisessa mallissa on satunnainen osa ja systemaattinen osa, joista ensimmäinen määrittelee vastemuuttujan ja sen jakauman ja jälkimmäinen lineaarisen ennusteen

$$\eta_i = \mathbf{x}_i^\top \boldsymbol{\beta},$$

missä  $i = 1, \dots, n$ . Mallin kolmas osa, eli linkkifunktio  $g$  muodostaa yhteyden vastemuuttujan odotusarvon ja lineaarisen ennusteen välille siten, että

$$g(\mu) = \eta_i,$$

missä  $\mu$  on vastemuuttujan odotusarvo. (Esim. Agresti 2018, 66–67.)

Olkoon linkkifunktio  $g$  logistinen funktio, eli

$$g : (0, 1) \rightarrow \mathbb{R}, \quad g(x) = \text{logit}(x) = \log\left(\frac{x}{1-x}\right).$$

Kertymä-logit-regressiomalli voidaan määrittellä järjestysasteikolliselle vastemuuttujalle  $Y$  yleistettynä lineaarisena mallina logistisen linkkifunktion käänteisfunktion avulla lausuttuna yhtälöllä

$$\gamma_{ij} = g^{-1}(\eta_{ij}), \quad i = 1, \dots, n, \quad j = 1, \dots, J-1,$$

missä  $\eta_{ij} = \theta_j - \mathbf{x}_i^\top \boldsymbol{\beta} = \theta_j - (x_{i1}\beta_1 + x_{i2}\beta_2 + \dots + x_{ip}\beta_p)$  on mallin lineaarinen ennuste ja  $\gamma_{ij}$  on kertymätodennäköisyys, eli

$$\gamma_{ij} = P(Y_i \leq j) = \pi_{i1} + \dots + \pi_{ij}, \quad \text{jolle} \quad \sum_{j=1}^J \pi_{ij} = 1.$$

Funktio  $g^{-1}$  on logistisen funktion käänteisfunktio, eli

$$g^{-1} : \mathbb{R} \rightarrow (0, 1), \quad g^{-1}(y) = \frac{\exp(y)}{1 + \exp(y)}.$$

Mallin parametreja ovat regressioparametrit  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$  sekä kynnsparametrit, eli vakiot  $\theta_j$ , joille pätee suuruusjärjestys

$$\theta_1 \leq \theta_2 \leq \dots \leq \theta_{J-1}.$$

Logistisen linkkifunktion vuoksi varsinaisena mallinnuksen kohteena on kertymävastausuhde, eli

$$\text{logit}(\gamma_{ij}) = \log\left[\frac{P(Y_i \leq j)}{1 - P(Y_i \leq j)}\right].$$

Kertymätodennäköisyys luokassa  $j$  voidaan nyt kirjoittaa mallin avulla muodossa

$$\gamma_{ij} = P(Y_i \leq j) = g^{-1}(\eta_{ij}) = \frac{\exp(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta})}{1 + \exp(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta})},$$

jolloin kertymätodennäköisyydelle saadaan tulkinta lineaarisesta ennusteesta. Merkitään  $\gamma_{ij}(\mathbf{x}_i) = P(Y_i \leq j \mid \mathbf{x}_i)$ . Koska vastasuhteiden suhde tapahtumalle  $\{Y_i \leq j\}$  selittäjävektorin arvoilla  $\mathbf{x}_1$  ja  $\mathbf{x}_2$ , eli kertymäristisuhde (*cumulative odds ratio*), on

$$\frac{\gamma_j(\mathbf{x}_1)/[1 - \gamma_j(\mathbf{x}_1)]}{\gamma_j(\mathbf{x}_2)/[1 - \gamma_j(\mathbf{x}_2)]} = \frac{\exp(\theta_j - \mathbf{x}_1^\top \boldsymbol{\beta})}{\exp(\theta_j - \mathbf{x}_2^\top \boldsymbol{\beta})} = \exp[(\mathbf{x}_2^\top - \mathbf{x}_1^\top) \boldsymbol{\beta}],$$

joka ei riipu indeksistä  $j$ , niin kertymäristisuhde on verrannollinen vektoreiden  $\mathbf{x}_1$  ja  $\mathbf{x}_2$  väliseen etäisyyteen riippumatta siitä, millä indeksin  $j$  arvolla kertymätodennäköisyyttä  $P(Y_i \leq j)$  tarkastellaan. Näin ollen malli olettaa selittäjävektorin  $\mathbf{x}_i$  vaikutuksen kertymävastasuhteeseen samaksi kaikissa luokissa  $j = 1, \dots, J-1$ . Mallissa on estimoitavana siten vain kynnsparametrit  $\theta_1, \dots, \theta_{J-1}$  sekä kaikille kertymävastasuhteille yhteinen selittäjien vaikutus  $\boldsymbol{\beta}$ . (Esim. Christensen 2018a, 3–5; Agresti 2010, 44–47.)

Yksittäisen luokan todennäköisyys voidaan mallin avulla lausua peräkkäisten luokien kertymätodennäköisyyksien erotuksena

$$P(Y_i = j) = \frac{\exp(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta})}{1 + \exp(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta})} - \frac{\exp(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})}{1 + \exp(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})},$$

missä  $\theta_0 = -\infty$  ja  $\theta_J = \infty$ .

## 2.4 Piilomuuttuja

Regressioparametrin  $\boldsymbol{\beta}$  yhteistä vaikutusta kaikkiin vastasuhteisiin voi perustella myös niin sanotun piilomuuttujan avulla. Järjestysasteikollisen vastemuuttujan  $Y$  taustalla ajatellaan olevan jatkuva satunnaismuuttuja  $Y^*$ , jota ei havaita. Oletetaan, että kiinnitetyllä selittäjävektorin arvolla  $\mathbf{x}_i$  jatkuva piilomuuttuja  $Y_i^*$  saa arvoja jonkin sijaintiparametrin  $\mu$  (esimerkiksi odotusarvon) ympäriltä. Oletetaan lisäksi, että tämä sijaintiparametri riippuu selittäjävektorista lineaarisesti siten, että  $\mu = \mathbf{x}_i^\top \boldsymbol{\beta}$ . Jos sijaintiparametri  $\mu$  on piilomuuttujan odotusarvo, niin selittäjävektorin arvolla  $\mathbf{x}_i$  pätee

$$Y_i^* = \mu + \epsilon_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \epsilon_i,$$

missä  $\epsilon_i$  on satunnaismuuttuja, jolle  $E(\epsilon_i) = 0$ . Erityisesti satunnaismuuttujan  $\epsilon_i$  jakauma ei riipu selittäjistä  $\mathbf{x}_i$ . Olkoon  $G$  satunnaismuuttujan  $\epsilon$  kertymäfunktio. Tällöin satunnaismuuttujan  $Y_i^*$  ehdollinen kertymäfunktio on

$$P(Y_i^* \leq y^* \mid \mathbf{x}_i) = G(y^* - \mu) = G(y^* - \mathbf{x}_i^\top \boldsymbol{\beta}).$$

Tarkastellaan tapahtumaa  $\{Y_i = j\}$ , missä  $j = 1, \dots, J$ . Piilomuuttujatarkastelussa oletetaan, että tätä vastaa tapahtuma  $\{\theta_{j-1} < Y_i^* \leq \theta_j\}$ , missä  $j = 1, 2, \dots, J$  ja

$$\theta_1 < \dots < \theta_{J-1}.$$

Järjestysasteikollinen satunnaismuuttuja  $Y_i$  saa siis arvon  $j$  silloin, kun jatkuva piilomuuttuja  $Y_i^*$  kuuluu  $j$ :nnteen kynnyksparametrien määräämään väliin. Tällöin

$$P(Y_i \leq j \mid \mathbf{x}_i) = P(Y_i^* \leq \theta_j \mid \mathbf{x}_i) = G(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}).$$

Toisaalta tämä tarkoittaa, että

$$G^{-1}[P(Y_i \leq j \mid \mathbf{x}_i)] = \theta_j - \mathbf{x}_i^\top \boldsymbol{\beta},$$

eli piilomuuttujan  $Y_i^*$  kertymäfunktion käänteisfunktiole  $G^{-1}$  saadaan sama tulkin-ta kuin mallin linkkifunktion käänteisfunktiole  $g^{-1}$ . Mikäli piilomuuttujan jakaumaksi oletetaan logistinen jakauma odotusarvolla  $\mu$ , niin  $G$  on logistisen jakauman kertymä-funktio

$$G(\epsilon) = \frac{e^\epsilon}{1 + e^\epsilon}$$

ja tämän käänteisfunktio on logistinen funktio.

Piilomuuttujan kautta tarkastelemalla nähdään, että eri selittäjän arvoilla on vaikutus vastemuuttujan sijaintiin, mutta ei kuitenkaan satunnaismuuttujan  $\epsilon$  skaalaan eli hajontaan. Tämä tarkoittaa, että kertymä-logit-regressiomalli tässä muodossa sopii huonosti tilanteisiin, joissa eri selittäjän arvoilla piilomuuttujan arvojen vaihtelu on voimakkaampaa. (Agresti 2010, 53–55.)

## 2.5 Mallin laajentaminen skaalavaikutuksilla

Joissakin tilanteissa vastemuuttujan äärimmäiset arvot voivat korostua merkittäväällä tavalla eri selittäjän arvoilla ja vastaavasti vastemuuttujan keskimmäiset arvot toisilla selittäjän arvoilla. Tällaisissa tilanteissa voi olla mielekästä mallintaa piilomuuttujan ja sitä kautta vastemuuttujan skaalaa erikseen näillä selittäjillä. McCullagh (1980) ja Cox (1995) mainitsevat malleista, joissa vastemuuttujan skaala huomioidaan eri selittäjän arvoilla.

Myös tällainen malli voidaan perustella piilomuuttujan avulla. Selittäjistä  $\mathbf{x}$  riippuva piilomuuttuja on nyt

$$Y^* = \mathbf{x}^\top \boldsymbol{\beta} + \tau_{\mathbf{x}} \epsilon,$$

missä  $\tau_{\mathbf{x}}$  on piilomuuttujan varianssi, joka saa myös riippua selittäjistä  $\mathbf{x}$ . Vastaavaa mallia kutsutaan sijaintiskaalamalliksi. Nimi tulee siitä, että siinä mallinnetaan taustalla olevan piilomuuttujan sijaintia ja skaalaa. Vastemuuttujalle  $Y_i$  pätee nyt

$$P(Y_i \leq j \mid \mathbf{x}_i) = G\left(\frac{\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}}{\tau_{\mathbf{x}}}\right),$$

missä  $G$  on satunnaismuuttujan  $\epsilon$  kertymäfunktio ja toisaalta mallin linkkifunktio. Hajontaparametrilta  $\tau_{\mathbf{x}}$  vaaditaan tietenkin positiivisuus, minkä vuoksi se määritellään usein olemaan  $\exp(\mathbf{z}^\top \boldsymbol{\zeta})$ , missä parametrivektori  $\boldsymbol{\zeta}$  kuvaa varsinaisia selittäjien skaalavaikutuksia ja  $\mathbf{z}$  vastaavia selittäjiä. Malli voidaan tällöin määritellä yhtälöllä

$$\text{logit}(\gamma_{ij}) = \frac{\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}}{\exp(\mathbf{z}_i^\top \boldsymbol{\zeta})}.$$

Jaksossa 2.3 määritelty malli on tämän mallin erikoistapaus siinä tapauksessa, että  $\boldsymbol{\zeta} = \mathbf{0}$ , mikä tarkoittaa, että skaalavaikutuksia ei oleteta. Sijaintiskaalamallista kuitenkin nähdään, että ennuste ei ole lineaarinen eikä se siten kuulu enää yleistettyjen lineaaristen mallien joukkoon. Tämän vuoksi mallin soveltamiseen tarvitaan myös jokin tähän erityisesti laadittu ohjelma. (Esim. Tutz ja Berger 2016, 3–4; Christensen 2018a, 7.)

## 2.6 Mallin laajentaminen nominaalivaikutuksilla

Tarkastellaan skaalavaikutuksilla laajennettua mallia

$$\text{logit}(\gamma_{ij}) = \frac{\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}}{\exp(\mathbf{z}_i^\top \boldsymbol{\zeta})}.$$

Mallissa oletettiin, että regressioparametrit  $\boldsymbol{\beta}$  eivät riipu luokasta  $j$ , eli selittävien muuttujien  $\boldsymbol{x}$  oletetaan vaikuttavan samalla tavalla jokaisen luokan välillä. Tätä oletusta voidaan kuitenkin lieventää ottamalla mukaan nominaalivaikutukset, jolloin mallissa saadaan myös sellaisia selittäjiä, jotka vaikuttavat eri tavalla eri luokkien välillä. Nominaalivaikutuksilla laajennettu sijaintiskaalamalli määritellään yhtälöllä

$$\text{logit}(\gamma_{ij}) = \frac{\theta_j - \boldsymbol{x}_i^\top \boldsymbol{\beta} - \boldsymbol{w}_i^\top \tilde{\boldsymbol{\beta}}_j}{\exp(\boldsymbol{z}_i^\top \boldsymbol{\zeta})}, \quad j = 1, \dots, J - 1,$$

missä  $\tilde{\boldsymbol{\beta}}_j$  on luokkaa  $j$  vastaava ylimääräinen regressioparametrivektori ja  $\boldsymbol{w}_i$  sisältää näitä vastaavat selittäjät havaintoyksikössä  $i$ . Nominaalivaikutukset voidaan ajatella myös kynnyksparametrien  $\theta_j$  laajennukseksi, jolloin

$$\tilde{\theta}_{ij} = \theta_j - \boldsymbol{w}_i^\top \tilde{\boldsymbol{\beta}}_j,$$

missä  $\tilde{\theta}_{ij}$  on  $j$ :s laajennettu kynnyksparametri havaintoyksikössä  $i$ .

Olkoon  $\boldsymbol{W}$  nominaaliselittäjistä muodostettu matriisi. Tällöin parametrivektorissa  $\tilde{\boldsymbol{\beta}}_j$  on yhtä monta komponenttia kuin matriisissa  $\boldsymbol{W}$  on sarakkeita ja parametrivektorissa  $\tilde{\boldsymbol{\beta}}$  on yhtä monta komponenttia kuin matriisissa  $\boldsymbol{W}$  on sarakkeita kerrottuna  $(J - 1)$ :llä.

Mikäli matriisissa  $\boldsymbol{W}$  on vain yksi sarake, joka on binäärinen, eli indikaattorimuuttuja, niin tällöin  $\tilde{\boldsymbol{\beta}}_j$  on skalaari. Mallin kynnyksparametreja voidaan tällöin tulkita siten, että laajennettu kynnyksparametri  $\tilde{\theta}_{ij} = \theta_j$  silloin, kun indikaattori saa arvon 0 ja  $\tilde{\theta}_{ij} = \theta_j - \tilde{\beta}_j$  silloin, kun indikaattori saa arvon 1. (Christensen 2018a, 6 ja 8–9.)

## 2.7 Parametrien estimointi

### 2.7.1 Tavallinen malli

Tarkastellaan tavallista kertymä-logit-mallia ilman skaala- tai nominaalivaikutuksia. Mallin parametrit estimoidaan aineistosta tyypillisesti suurimman uskottavuuden menetelmällä. Uskottavuusfunktion muodostamista varten oletetaan, että aineistona on  $n$  riippumattomaa järjestysasteikollisesti jakautunutta satunnaismuuttujaa  $Y_1, \dots, Y_n$ , joissa on luokat  $1, 2, \dots, J$ . Niiden yhteisjakauma on multinomijakauma parametreilla  $\pi_1, \dots, \pi_J$ .

Uskottavuusfunktion muodostamista varten määritetään ensin  $y_{i1}, \dots, y_{iJ}$  siten, että  $y_{ij} = 1$ , jos  $Y_i = j$ , eli jos  $i$ :nnessä havainnossa havaitaan luokka  $j$ , ja  $y_{ir} = 0$  kaikilla  $r \neq j$ . Merkitään  $\pi_j(\boldsymbol{x}_i) = P(Y_i = j \mid \boldsymbol{x}_i)$  korostamaan estimoitavien todennäköisyyksien riippuvuutta selittäjistä. Koska havainnot oletetaan riippumattomiksi ja koska

$$P(Y_i = j \mid \boldsymbol{x}_i) = P(Y \leq j \mid \boldsymbol{x}_i) - P(Y_i \leq j - 1 \mid \boldsymbol{x}_i),$$

niin kertymä-logit-mallin uskottavuusfunktio saadaan multinomitodennäköisyyksien tu-

lona

$$\begin{aligned} \prod_{i=1}^n \left[ \prod_{j=1}^J \pi_j(\mathbf{x}_i)^{y_{ij}} \right] &= \prod_{i=1}^n \left\{ \prod_{j=1}^J [P(Y \leq j \mid \mathbf{x}_i) - P(Y \leq j-1 \mid \mathbf{x}_i)]^{y_{ij}} \right\} \\ &= \prod_{i=1}^n \left\{ \prod_{j=1}^J \left[ \frac{\exp(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta})}{1 + \exp(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta})} - \frac{\exp(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})}{1 + \exp(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})} \right]^{y_{ij}} \right\}. \end{aligned}$$

Kun on aineiston kautta havaitaan  $y_{ij}$  ja koska selittäjät oletetaan tunnetuiksi, niin uskottavuusfunktio on kynnysparametrien ja regressioparametrien  $\theta_1, \dots, \theta_{J-1}, \beta_1, \dots, \beta_p$  funktio. Käytetään logaritmiselle uskottavuusfunktiolle merkintää  $l(\boldsymbol{\theta}, \boldsymbol{\beta})$ . Uskottavuusyhtälöt voidaan johtaa derivoimalla logaritmita uskottavuusfunktiota jokaisen parametrin suhteen ja asettamalla jokainen osittaisderivaatta nolaksi. Merkintöjen lyhentämiseksi käytetään merkintöjä

$$G(t) = \frac{\exp(t)}{1 + \exp(t)}, \quad g(t) = \frac{\exp(t)}{[1 + \exp(t)]^2}.$$

Nyt uskottavuusyhtälö regressioparametrille  $\beta_k$  kaikilla  $k = 1, \dots, p$  on

$$\begin{aligned} \frac{\partial l}{\partial \beta_k} &= 0 \\ \Leftrightarrow \sum_{i=1}^n \sum_{j=1}^J y_{ij} x_{ik} \frac{g(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}) - g(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})}{G(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}) - G(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})} &= 0. \end{aligned}$$

Olkoon  $\delta_{jr}$  Kroneckerin deltafunktio, eli  $\delta_{jr} = 1$ , jos  $j = r$  ja 0 muuten. Uskottavuusyhtälö parametrille  $\theta_r$  kaikilla  $r = 1, \dots, J-1$  on

$$\begin{aligned} \frac{\partial l}{\partial \theta_r} &= 0 \\ \Leftrightarrow \sum_{i=1}^n \sum_{j=1}^J y_{ij} \frac{\delta_{jr} g(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}) - \delta_{j,r} g(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})}{G(\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}) - G(\theta_{j-1} - \mathbf{x}_i^\top \boldsymbol{\beta})} &= 0 \end{aligned}$$

(Agresti 2010, 119–121.)

Uskottavuusyhtälöiden ratkaisua ei voi esittää suljetussa muodossa, joten ne on ratkaistava numeerisesti. McCullagh (1980) käytti tähän Fisherin estimointialgoritmia (*Fisher's scoring*). Tässä työssä käytettävissä malleissa hyödynnetään kuitenkin R-ohjelmiston kirjastoa, jossa estimointi perustuu muokattuun Newtonin-Raphsonin algoritmiin. Newtonin-Raphsonin algoritmin perusidea kuvataan esimerkiksi Agrestin kirjassa (Agresti 2015, 138–139). Keskeinen ero Fisherin estimointialgoritmin ja Newtonin-Raphsonin algoritmin välillä on Hessen matriisissa, eli toisen kertaluvun osittaisderivaattojen matriisissa. Newtonin-Raphsonin algoritmissa Hessen matriisi muodostetaan havaitusta informaatiosta, kun taas Fisherin estimointialgoritmissa käytetään Fisherin informaatiota, eli tämän odotusarvoa.

Tässä työssä hyödynnetty muokattu Newtonin-Raphsonin algoritmi, jota voidaan käyttää myös nominaalivaikutuksilla laajennetulle sijaintiskaalamallille, kuvaillaan tarkemmin jaksossa 2.7.2.

## 2.7.2 Laajennettu malli

Tarkastellaan nyt jaksossa 2.6 esitettyä nominaalivaikutuksilla laajennettua sijaintiskaalamallia

$$\gamma_{ij} = \text{logit}(\eta_{ij}), \quad \eta_{ij} = \frac{\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta} - \mathbf{w}_i^\top \tilde{\boldsymbol{\beta}}_j}{\exp(\mathbf{z}_i^\top \boldsymbol{\zeta})}.$$

R-ohjelmiston ordinal-kirjastossa sijaintiskaalamallin parametrit estimoidaan oletusarvoisesti säännöllistetyllä Newton-Raphsonin (NR) algoritmilla askelpuolituksella. Algoritmia voidaan käyttää myös silloin, kun mallissa ei ole skaala- tai nominaalivaikutuksia. R-ohjelmiston ordinal-kirjaston käyttämä algoritmi kuvataan yksityiskohtaisesti Christensenin artikkelissa (Christensen 2018a, 10–12).

Olkoon  $\boldsymbol{\psi} = (\boldsymbol{\theta}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\zeta})$  estimoitava parametrivektori. Olkoon lisäksi

$$\mathbf{g}(\boldsymbol{\psi}; \mathbf{y}) = \left( -\frac{\partial l(\boldsymbol{\psi}; \mathbf{y})}{\partial \psi_1}, \dots, -\frac{\partial l(\boldsymbol{\psi}; \mathbf{y})}{\partial \psi_q} \right)^\top$$

mallin negatiivisen log-uskottavuusfunktion gradientti ja

$$\mathbf{H}(\boldsymbol{\psi}; \mathbf{y}) = \begin{bmatrix} \frac{\partial^2 l(\boldsymbol{\psi}; \mathbf{y})}{\partial \psi_1 \partial \psi_1} & \cdots & \frac{\partial^2 l(\boldsymbol{\psi}; \mathbf{y})}{\partial \psi_1 \partial \psi_q} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 l(\boldsymbol{\psi}; \mathbf{y})}{\partial \psi_q \partial \psi_1} & \cdots & \frac{\partial^2 l(\boldsymbol{\psi}; \mathbf{y})}{\partial \psi_q \partial \psi_q} \end{bmatrix}$$

mallin log-uskottavuusfunktion Hessen matriisi. Skaalavaikutuksilla laajennetun kertymälogit-mallin Hessen matriisin ja gradientin analyttiset lausekkeet on johdettu lähdekirjallisuudessa (Christensen 2012, 60–61).

Säännöllistetty NR-algoritmi tuottaa jonon estimaatteja  $\boldsymbol{\psi}^{(0)}, \dots, \boldsymbol{\psi}^{(i)}, \dots$ , missä suluissa oleva yläindeksi kuvaa iteraation järjestyslukua. Iteraatiossa seuraava estimaatti saadaan edellisestä kaavalla

$$\boldsymbol{\psi}^{(i+1)} = \boldsymbol{\psi}^{(i)} - c_1 \mathbf{h}^{(i)}, \quad \mathbf{h}^{(i)} = \tilde{\mathbf{H}}(\boldsymbol{\psi}^{(i)}; \mathbf{y})^{-1} \mathbf{g}(\boldsymbol{\psi}^{(i)}; \mathbf{y}),$$

missä

$$\tilde{\mathbf{H}}(\boldsymbol{\psi}^{(i)}; \mathbf{y}) = \mathbf{H}(\boldsymbol{\psi}^{(i)}; \mathbf{y}) + c_2(c_3 + \min(\mathbf{e}^{(i)}))\mathbf{I},$$

missä  $\mathbf{e}^{(i)}$  on Hessen matriisin  $\mathbf{H}(\boldsymbol{\psi}^{(i)}; \mathbf{y})$  ominaisarvovektori,  $\mathbf{h}^{(i)}$  on  $i$ :s askel,  $c_1$  on skalaari, joka kontrolloi askelpuolitusta, ja  $c_2$  ja  $c_3$  ovat skalaareja, jotka kontrolloivat Hessen matriisin säännöllisyyttä. Mikäli pienin ominaisarvo  $\min(\mathbf{e}^{(i)})$  on liian lähellä nollaa, niin  $c_2 = 1$  ja muuten  $c_2 = 0$ . Positiivisen skalaarin  $c_3$  valinta on jokseenkin mielivaltainen ja esimerkiksi ordinal-kirjastossa käytetään arvoa  $c_3 = 1$ .

Askelpuolitusta tarvitaan, mikäli askel  $\mathbf{h}^{(i)}$  johtaa uskottavuusfunktiossa nousun sijasta laskuun. Tällöin lukua  $c_1$  puolitetaan, kunnes askel johtaa nousuun uskottavuusfunktiossa tai kunnes puolitusten enimmäismäärä on saavutettu.

Iteraatioiden tuottamien estimaattien jono suppenee kohti optimiarvoja, kunhan negatiivisen log-uskottavuusfunktion gradientti parametrien suhteen on riittävän pieni ja



Hessen matriisi positiivisesti definiitti. Ongelmia optimoinnissa syntyy, jos uskottavuusfunktio on ainakin jonkin parametrin suhteen tasainen. Tavalliselle kertymä-logit-mallille voidaan näyttää, että sen log-uskottavuusfunktio on konkaavi, mistä seuraa, että sillä on yksikäsitteinen globaali ääriarvo. Konkaavius on näytetty esimerkiksi Prattin (1981) artikkelissa. Tavallisen mallin numeerisessa optimoinnissa ei siten ole riskiä sille, että estimaatit suppenisivat kohti lokaalia ääriarvoa. Tämä tulos ei kuitenkaan päde mallille, jossa on skaalavaikutus mukana.

## 2.8 Mallin parametrien tulkinta

Tarkastellaan mallin lineaarisen ennusteen osaa

$$\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}.$$

Kynnysparametrit  $\theta_j$  määräävät todennäköisyydet kullekin vastemuuttujan luokalle, kun parametrivektori  $\boldsymbol{\beta}$  on kiinnitetty. Mitä kauempana kaksi kynnysparametria toisistaan ovat, sitä suurempi on vastaavan väliin jäävän luokan todennäköisyys.

Parametrivektorin  $\boldsymbol{\beta}$  komponentit  $(\beta_1, \dots, \beta_p)$  voidaan tulkita kunkin vastaavan selittäjän  $\mathbf{x}_1, \dots, \mathbf{x}_p$  vaikutukseksi todennäköisyyden  $P(Y_i \leq j)$  logistiseen muunnokseen ja siten luokkaa  $j$  vastaavaan kertymävastasuhteeseen. Sana vaikutus viittaa tässä tilastolliseen yhteyteen, ei välttämättä syy-seuraussuhteeseen. Koska  $\boldsymbol{\beta}$  ei riipu indeksistä  $j$ , niin vastaavan selittäjän vaikutus on yhtä suuri jokaisella indeksillä  $j$ . Koska lineaarinen ennuste on  $\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}$ , niin miinusmerkin vuoksi kunkin selittäjän vaikutus on käänteinen sen etumerkkiin nähden. Esimerkiksi jos parametri  $\beta_k > 0$ , niin se tarkoittaa sitä, että jatkuvan tai järjestysasteikollisen selittäjän  $x_{ik}$  arvojen kasvaessa kertymätodennäköisyyden logistinen muunnos pienenee ja siten myös vastaava kertymätodennäköisyys pienenee kaikilla indekseillä  $j$ . Toisaalta tämä tarkoittaa, että kertymätodennäköisyyttä  $P(Y \geq j)$  vastaava vastausuhde kasvaa kaikilla indekseillä  $j$ . Tämä tarkoittaa, että vastemuuttujan suurten arvojen todennäköisyydet kasvavat suhteessa vastemuuttujan pienten arvojen todennäköisyyksiin. Jos  $\beta_k < 0$ , niin tulkinta on päinvastainen, eli tällöin vastemuuttujan pienten arvojen todennäköisyydet kasvavat suhteessa vastemuuttujan suurten arvojen todennäköisyyksiin. Jaksossa 2.3 todetun kertymäristisuhteen verrannollisuusominaisuuden perusteella yhden yksikön muutos selittäjässä  $x_{ik}$  merkitsee  $\exp(\beta_k)$ -kertaista muutosta jokaisessa kertymävastasuhteessa. (Esim. Agresti 2010, 48–49.)

Jos selittäjä on luokitteluasteikollinen, siten, että sillä on  $r$  luokkaa, niin malli on muotoa

$$\text{logit}[P(Y_i \leq j)] = \theta_j - \tau_{i1}z_1 - \tau_{i2}z_2 - \dots - \tau_{ir-1}z_{r-1},$$

missä  $\tau_1, \dots, \tau_{r-1}$  on jokaista luokkaa, paitsi viimeistä, vastaava parametri ja  $z_l$  on indikaattorimuuttuja, joka saa arvon 1, mikäli vastaava havainto kuuluu luokkaan  $l$ , ja muuten arvon 0 ( $l = 1, \dots, r-1$ ). Viimeinen luokka on vertailuluokka, jolle ei ole omaa

indikaattorimuuttujaa. Tällöin esimerkiksi kahdella kiinnitettyllä parametrin arvolla  $\tau_a$  ja  $\tau_b$ , joille  $\tau_b > \tau_a$ , kertymätodennäköisyys on pienempi luokassa  $b$  kuin luokassa  $a$  kaikilla indekseillä  $j$ . (Esim. Agresti 2010, 50–51.)

Tarkastellaan skaalavaikutuksilla laajennettua mallia

$$\text{logit}(\gamma_{ij}) = \frac{\theta_j - \mathbf{x}_i^\top \boldsymbol{\beta}}{\exp(\mathbf{z}_i^\top \boldsymbol{\zeta})}.$$

Skaalaparametrivektoria  $\boldsymbol{\zeta}$  voidaan tulkita siten, että mikäli sen jokin komponentti on positiivinen, niin vastaavan skaalaselittäjän arvojen kasvaessa vastemuuttujan hajonta kasvaa. Negatiivisella skaalaparametrilla vastaavan skaalaselittäjän arvojen kasvaessa vastemuuttujan hajonta pienenee. Mikäli skaalaselittäjä on indikaattorimuuttuja ja parametri on positiivinen, niin se tarkoittaa, että indikaattorin saadessa arvon 1 vastemuuttujan hajonta on suurempi kuin silloin, jos indikaattori saa arvon 0. (Esim. Christensen 2018a, 22–23; Agresti 2010, 131.)

## Luku 3

# Diagnostiikka ja mallin valinta

Kun aineistossa on paljon muuttujia, niin näistä muodostettavilla selittäjillä saadaan aikaan useita erilaisia malleja, joiden sopivuutta ja ennustekykyä on yleensä vertailtava keskenään. Monesti on myös valittava jollakin kriteerillä paras malli.

### 3.1 Parametrien luottamusvälit ja testit

Tilastollisen mallin parametrien estimaattien hyvyttä voidaan arvioida esimerkiksi tutkimalla niiden keskivirheitä ja niiden avulla muodostettavia luottamusvälejä, sekä testaamalla parametreihin liittyviä hypoteeseja.

Tarkastellaan skaala- ja nominaalivaikutuksilla laajennettua mallia, jonka parametri on  $\boldsymbol{\psi} = (\boldsymbol{\theta}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \boldsymbol{\zeta})$ . Kiinnostuksen kohteena on yksittäinen regressioparametri  $\beta_k$ , missä  $k = 1, \dots, p$ . Yhtä hyvin tämä voisi olla myös kynnysparametri, nominaaliparametri tai skaalaparametri. Merkitään parametrin suurimman uskottavuuden estimaattoria  $\hat{\beta}_k$ . Suurimman uskottavuuden estimaattorin asymptoottisten tulosten nojalla estimaattorin keskivirhe voidaan estimoida kaavalla

$$\text{SE}(\hat{\beta}_k) = \frac{1}{\sqrt{i_{\beta_k}(\hat{\beta}_k)}},$$

missä  $i_{\beta_k}(\hat{\beta}_k)$  on mallin Fisherin informaatiomatriisin

$$i_{ab} = E \left[ \frac{-\partial^2 l(\boldsymbol{\psi})}{\partial \beta_a \partial \beta_b} \right]$$

parametria  $\beta_k$  vastaava diagonaalialkio suurimman uskottavuuden pisteessä  $\hat{\beta}_k$ . Vaihtoehtoisesti informaatiomatriisina voidaan käyttää myös havaittua informaatiomatriisia, jossa log-uskottavuusfunktion toisen kertaluvun osittaisderivaattojen odotusarvot korvataan aineistosta lasketuilla toisen kertaluvun osittaisderivaatoilla.

Olkoon  $\alpha \in (0, 1)$ . Parametrin  $\beta_k$   $100 \cdot (1 - \alpha)$  prosentin Waldin luottamusväli saadaan kaavalla

$$\hat{\beta}_k \pm z_{\alpha/2} \text{SE}(\hat{\beta}_k),$$

missä  $z_{\alpha/2}$  on se piste, jonka oikealle puolelle jää  $\alpha/2$  verran todennäköisyyttä standardinormaalijakaumassa. (Esim. Agresti 2010, 59–60.)

Parempia, mahdollisesti ei-symmetrisiä luottamusvälejä saadaan Christensenin (2018a, 17) mukaan uskottavuusosamäärän neliöjuureen perustuvalla suureella

$$r(\beta_k) = \text{sign}(\hat{\beta}_k - \beta_k) \sqrt{-2[l(\beta^*) - l_P(\beta_k)]},$$

missä  $l_P(\beta_k)$  on logaritminen profiiluskottavuusfunktio. Tämä saadaan tavallisesta loguskottavuusfunktioista suorittamalla jokaisella kiinnostavan parametrin  $\beta_k$  arvolla maksimointi mallin kaikkien muiden parametrien suhteen. Tästä johdettava parametrin  $\beta_k$  luottamusväli on muotoa

$$\{\beta_k : |r(\beta_k)| < z_{\alpha/2}\}.$$

Regressioparametrien  $\beta_k$  ja skaalaparametrien  $\zeta_l$  osalta on mielekästä testata nollahypoteesia, jonka mukaan parametri on nolla, eli onko kyseessä turha selittäjä. Kynnysparametrien  $\theta_j$  osalta tällainen testi puolestaan ei ole kovin mielekäs niiden tulkinnan vuoksi. Testi voidaan perustaa Waldin testisuureeseen

$$w = \frac{\hat{\beta}_k}{\text{SE}(\hat{\beta}_k)},$$

joka noudattaa asympotoottisesti standardinormaalijakaumaa nollahypoteesin pätiessä. (Esim. Agresti 2010, 60.)

Vaihtoehtona Waldin testille on uskottavuusosamäärätesti, joka perustuu testisuureeseen

$$r = 2 [l(\hat{\psi}) - l(\hat{\psi}_0)],$$

missä  $\hat{\psi}$  on vapaa suurimman uskottavuuden estimaatti ja  $\hat{\psi}_0$  rajoitettu, eli nollahypoteesin mukainen suurimman uskottavuuden estimaatti. Myös kahta sisäkkäistä mallia, eli malleja, joista toisen parametrien joukko sisältyy toisen mallin parametrien joukkoon, voidaan vertailla uskottavuusosamäärätestillä. Uskottavuusosamäärän testisuure noudattaa nollahypoteesin pätiessä asympotoottisesti  $\chi_d^2$ -jakaumaa, missä  $d$  on vapaan ja rajoitetun mallin parametrien lukumäärien erotus. Tässä tutkielmassa on tietyissä tilanteissa erityisen kiinnostavaa testata nollahypoteesia

$$H_0 : \beta_k = -\beta_l \quad \text{jollakin } k \neq l.$$

Käytännössä siis testataan sitä, onko kahden eri selittäjän tilastollinen yhteys samansuuruinen, mutta eri suuntainen. Mallille, jolle edellä mainittu nollahypoteesi pätee, lineaarinen ennuste on muotoa

$$\beta_k x_k + \beta_l x_l = \beta_k x_k - \beta_k x_l = \beta_k (x_k - x_l),$$

jota voidaan käyttää uskottavuusosamäärätestissä lineaarisena rajoitteena.

Agrestin (2010, 60) mukaan uskottavuusosamäärätesti on Waldin testiä suositeltavampi silloin, kun otoskoko on pieni tai kun suuri osuus vastemuuttujan havaituista arvoista kuuluu järjestysasteikon pienimpään tai suurimpaan luokkaan.

Koska ordinal-kirjasto käyttää yksittäisten parametrien testauksessa Waldin testiä ja koska uskottavuusosamäärätestin käyttäminen on tietyissä tilanteissa teknisesti monimutkaista ordinal-kirjastossa, niin tässä työssä yksittäisten parametrien testit, joissa testataan parametrin poikkeavuutta nolasta, perustuvat Waldin testiin. Uskottavuusosamäärätestiä käytetään tässä työssä kuitenkin tilanteissa, joissa se on teknisesti yksinkertaisempaa.

## 3.2 Yhteensopivuusindeksi

Mallin ennustekykyä voidaan mitata yhteensopivuusindeksillä (*concordance index*). Tarkastellaan havaittua aineistoa ja kaikkia sellaisia vastemuuttujan arvojen pareja, joiden arvot ovat erisuuret. Yhteensopivuusindeksissä lasketaan osuus niistä mallin ennustamista vastemuuttujan arvoista, jotka ovat yhteensopivia todellisten vastemuuttujan arvojen kanssa. Yhteensopivuus tarkoittaa, että malli osaa erotella kahden havainnon välillä oikein sen, kumpi näistä saa järjestysasteikolla suuremman arvon. Siis havainnolla, jolla vastemuuttujan arvo on korkeampi kuin toisella, myös tätä havaintoa vastaava ennuste toisen havainnon ennusteeseen nähden tulee olla sellainen, että malli osaa erotella tämän suuremmaksi. Käytännössä yhteensopivuusindeksi siis estimoit todennäköisyyttä sille, että malli osaa laittaa kaksi vastemuuttujan arvoa oikeaan suuruusjärjestykseen.

Jos yhteensopivuusindeksin arvo on  $\frac{1}{2}$ , niin se viittaa siihen, että ennustekyky on yhtä hyvä kuin satunnainen arvaus. Mitä lähempänä lukua 1 yhteensopivuusindeksi on, sitä parempi on mallin ennustekyky tällä kriteerillä.

Yhteensopivuusindeksi saadaan lineaarimuunnoksena Somersin  $d$ -suureesta, joka mittaa kahden järjestysasteikollisen satunnaismuuttujan välistä yhteyttä. Somersin  $d$ -suure lasketaan kaavalla

$$d = \frac{C - D}{\frac{n(n-1)}{2} - T_y},$$

missä  $C$  on yhteensopivien parien lukumäärä,  $D$  on ei-yhteensopivien parien lukumäärä ja  $T_y$  on niiden parien lukumäärä, joilla vastemuuttujan arvot ovat yhtä suuret. Somersin  $d$ -suure kuuluu välille  $[-1, 1]$  ja yhteensopivuusindeksi CI saadaan tästä suoraan lineaarimuunnoksella

$$\text{CI} = \frac{d + 1}{2},$$

joka kuuluu välille  $[0, 1]$ .

Yhteensopivuusindeksin etuja ovat sen laskennallinen ja tulkinnallinen yksinkertaisuus. Toisaalta se ei osaa erotella malleja, joissa on käytetty eri linkkifunktioita, sillä mallien ennusteiden suuruusjärjestys riippuu vain lineaarisesta ennusteesta (tai sijaintiskalamallilla epälinearisesta ennusteesta), eikä ollenkaan linkkifunktiosta. (Esim. Agresti ja Tarantola 2018, 218–219.)

Frank Harrellin mukaan mallia, jonka yhteensopivuusindeksi ylittää arvon 0,8, voi pitää ennustekyvyltään hyvänä, kun mallilla ennustetaan yksittäisen tutkimuskohteen

vastemuuttujan arvoja (Harrell 2015, 257).

### 3.3 Yhteiskorrelaatiokerroin

Toinen keino mitata mallin ennustekykä on estimoida jollakin tavalla mallin antamien sovitearvojen ja havaittujen vastemuuttujan arvojen välistä lineaarista riippuvuutta. Kertymä-logit-malli voidaan ajatella vastemuuttujan  $Y$  taustalla olevan piilomuuttujan  $Y^*$  avulla tavallisena lineaarisena mallina, jonka selitysaste on

$$R^2 = \frac{\sum_{i=1}^n (y_i^* - \bar{y}^*)^2 - \sum_{i=1}^n (y_i^* - \hat{y}_i^*)^2}{\sum_{i=1}^n (y_i^* - \bar{y}^*)^2} = \frac{\sum_{i=1}^n (\hat{y}_i^* - \bar{y}^*)^2}{\sum_{i=1}^n (y_i^* - \bar{y}^*)^2},$$

missä  $\hat{y}_i^*$  on mallin sovittama piilomuuttujan arvo havaintoyksikössä  $i$ . Piilomuuttujamallin yhteiskorrelaatiokerroin voidaan estimoida laskemalla neliöjuuri selitysasteesta  $R^2$ . Tämä voidaan kuitenkin estimoida myös vaihtoehtoisella tavalla hyödyntämättä lainkaan piilomuuttujaa. Olkoon  $p_j$  otoksesta laskettu havaintojen osuus luokassa  $j$ , missä  $j = 1, \dots, J$ . Määritellään

$$v_j = \sum_{k=1}^{j-1} p_k + \left(\frac{1}{2}\right) p_j, \quad j = 1, 2, \dots, J$$

kuvaamaan keskimääräistä kumulatiivista havaintojen osuutta luokassa  $j$ . Näille kumulatiivisille osuuksille käytetään myös englanniksi nimitystä *ridits*. Siihen lasketaan siis niiden havaintojen osuus, jotka kuuluvat  $j$ :tä pienempään luokkaan sekä puolet niistä havainnoista, jotka kuuluvat luokkaan  $j$ .

Määritellään keskisijaluku (*midrank*) luokalle  $j$

$$r_j = \frac{\sum_{i=1}^{j-1} n_i + 1 + \sum_{i=1}^j n_i}{2}, \quad j = 1, 2, \dots, J,$$

missä  $n_i$  on havaintojen lukumäärä luokassa  $i$ . Tällöin pätee yhteys

$$r_j = nv_j + 0.5 \quad \text{ja} \quad v_j = \frac{r_j - 0.5}{n}.$$

Laskemalla kumulatiiviset osuudet tai keskisijaluvut sekä havaituista vastemuuttujan arvoista että keskimääräisistä mallin sovittamista arvoista ja laskemalla näiden välinen korrelaatiokerroin saadaan Spearmanin korrelaatiokerroin. (Agresti ja Tarantola 2018, 219–220.)

### 3.4 Selittävien muuttujien valinta

Kun aineistossa on paljon potentiaalisia selittäjiä, niin mallinvalinta on syytä tehdä systemaattisesti. Agresti (2010, 74) toteaa, että järjestysregressiomallin valitsemisessa voidaan käyttää tavanomaisia mallinvalintamenetelmiä. Agresti (2015, 143) esittää kaksi keskeistä tavoitetta selittäjien valinnassa yleistetyssä lineaarisessa mallissa, jotka ovat

samat kuin tavallisessa lineaarisessa mallissa. Mallin tulee toisaalta olla riittävän monipuolinen, jotta se sopii käytettyyn aineistoon. Toisaalta se ei saa olla ylisopiva vain tiettyyn aineistoon, ja sen on oltava myös tulkinnallisesti suhteellisen yksinkertainen.

### 3.4.1 Akaiken informaatiokriteeri

Eräs kriteeri sopivan mallin valinnassa on Akaiken informaatiokriteeri (AIC). Se ottaa huomioon mallin maksimoidun log-uskottavuusfunktion, mutta myös mallin parametrien määrän. Jos mallin parametrivektori  $\psi$  koostuu  $d$ :stä skalaariparametrusta, niin AIC:n perusteella optimaalinen malli minimoi suureen

$$\text{AIC} = -2[l(\hat{\psi}) - d],$$

missä  $l(\hat{\psi})$  on mallin maksimoitu log-uskottavuusfunktio. Selittäjien lisääminen malliin kasvattaa logaritmista uskottavuusfunktiota, joten AIC pyrkii samalla rankaisemaan mallia mitä enemmän siinä on selittäjiä. Tämän tavoitteena on jättää malliin vain aidosti tärkeät selittäjät ja sulkea pois turhat. Kriteerin perusteella valitaan malli, jonka AIC on pienin. (Esim. Held ja Sabanés Bové 2014, 224–227.)

### 3.4.2 Mallinvalinta-algoritmit

Selittäjien valitsemiseen on olemassa erilaisia algoritmeja. Parhaan osajoukon valinnassa (*best subset selection*) valitaan malliin sellaiset selittäjät, joilla jokin diagnostiikkakriteeri täyttyy parhaiten. Jos selittäjäehdokkaiden määrä on suuri, menetelmä voi kuitenkin olla jopa mahdoton. (Agregi 2015, 143–144.)

Eteenpäin askeltavassa valinnassa (*forward selection*) lähdetään liikkeelle yksinkertaisesta mallista ja lisätään selittäjiä askeleittain. Jokaisella askeleella malliin lisätään sellainen selittäjä, joka parantaa parhaiten mallin sopivuutta. Tätä toistetaan, kunnes selittäjien lisääminen ei enää paranna mallin sopivuutta. Sopivuutta voidaan mitata esimerkiksi AIC:llä tai parametrien testeillä. Taaksepäin askeltavassa valinnassa (*backward elimination*) lähdetään liikkeelle mallista, jossa on kaikki selittäjät mukana. Jokaisella askeleella poistetaan huonoin selittäjä. Kriteerinä voidaan käyttää esimerkiksi yksittäisen parametrin testin suurinta  $p$ -arvoa tai pienintä heikennystä mallin sopivuuteen jonkin muun kriteerin perusteella. (Esim. Agresti 2015, 144.)

Hosmer, Lemeshow ja Sturdivant (2013, 90) toteavat, että logistisessa regressiomallissa selittävien muuttujien määrä on syytä pitää suhteellisen pienenä, sillä tällöin malli on usein numeerisesti vakaampi. Lisäksi mitä enemmän selittäjiä mallissa on, sitä enemmän se riippuu käytetystä aineistosta, jolloin mallin soveltaminen yleisemmin voi olla kyseenalaista. He esittävät määrätietoisien mallinvalintamenetelmän (*purposeful selection*), jossa ensiksi valitaan joukko sellaisia selittäjiä, joiden ennestään tiedetään olevan tärkeitä tai jotka yhden selittäjän mallissa osoittavat pientäkään merkkiä siitä, että kyseessä voisi olla sopiva selittäjä. Tämän jälkeen sovelletaan taaksepäin askeltavaa valintaa lähtien liikkeelle kaikilla edellä mainituilla selittäjillä. Selittäjä pidetään

mallissa, mikäli se on merkitsevä jollakin melko tiukalla merkitsevyytasolla tai mikäli sen poistaminen muuttaa jonkin toisen muuttujan estimoitua vaikutusta vähintään 20 prosenttia. Tämän jälkeen katsotaan, onko mikään sellainen selittäjä, jota ei otettu täyteen malliin, merkitsevä, kun se lisätään malliin. Lopuksi tutkitaan vielä mahdollisia yhteisvaikutuksia jäljellä olevien selittäjien osalta ja sovelletaan malliin tavanomaisia diagnostiikkatarkasteluja.



## Luku 4

# Aineisto ja muuttajat

Tutkielmassa käytetty oikeustapausaineisto kerättiin jokaisesta Suomen viidestä hovioikeudesta, jotka ovat Helsingin hovioikeus, Turun hovioikeus, Itä-Suomen hovioikeus, Vaasan hovioikeus ja Rovaniemen hovioikeus. Itä-Suomen hovioikeuden asiakirjat sisälsivät myös vuonna 2014 lakkautetun Kouvolan hovioikeuden asiakirjoja. Aineisto kerättiin vuosien 2017 ja 2018 aikana.

Asiakirjat, joista aineisto kerättiin, sisälsivät huolto- ja asumisriitoja vastaavat hovioikeuden ratkaisut sekä niiden tausta-asiakirjat. Osa tausta-asiakirjoista oli salassa pidettäviä ja niistä saataviin tietoihin oli pyydetty lupa jokaiselta hovioikeudelta erikseen. Salassa pidettävät asiakirjat saattoivat sisältää esimerkiksi sosiaalilautakunnan olosuhdeselvityksiä tai kokonaan salaisiksi julistettuja päätöksiä taustamateriaaleineen. Lupa saatiin jokaiselta hovioikeudelta, vaikkakin näiden yksityiskohdissa saattoi olla pieniä eroja. Lupaa ei pyydetty esittelijöiden muistioihin tai oikeuden omiin muistiinpanoihin, jotka olivat kaikissa tapauksissa salassa pidettäviä.

Aineiston keruu tapahtui pääasiassa paikan päällä kunkin hovioikeuden kirjaamon tiloissa. Aineistoon valikoituneet asiakirjat määräytyivät hovioikeuden arkistonhoitajan tai kirjaajan muodostamien listojen avulla. Listoihin oli koostettu hovioikeuden tietojärjestelmästä haetut diaarinumerot ja nimikkeet vuosien 2000–2016 hovioikeuden ratkaisuista hakukriteereinä lasten huolto ja asuminen. Tutkimuksessa päädyttiin jo aineistonkeruuvaiheessa tutkimaan vain asumisriitoja, eikä asumisriitoja pystynyt karsimaan suoraan listauksesta pelkän nimikkeen perusteella. Listauksessa esimerkiksi nimike ”Lapsen huolto ym.” saattoi koskea huoltoriitaa, asumisriitaa, molempia tai ei välttämättä edes kumpaakaan, vaan esimerkiksi lapsen elatusta. Tämä omalta osaltaan hidasti aineiston keruuta, koska lähes kaikista, myös turhista asiakirjoista, oli luettava ainakin osa hovioikeuden päätöksestä, jotta selvisi, oliko kyseessä tutkimuksen kannalta oleellinen tapaus, eli aito asumisriita. Listauksesta karsittiin myös tapaukset, jotka olivat päättyneet sovintoon tai sellaiset tapaukset, joissa toinen hakija oli peruuttanut valituksensa hovioikeuteen. Myös esimerkiksi sellaiset tapaukset, joiden ratkaisu perustui vain siihen,

että hovioikeus ei ollut toimikelpoinen ratkaisemaan asiaa, sivuutettiin. Näin aineistoon päätyivät ainoastaan riitatapaukset, joissa ainakin toinen osapuolista oli valituksessaan hovioikeuteen vaatinut yhden tai useamman lapsen asumisen muuttamista itselleen.

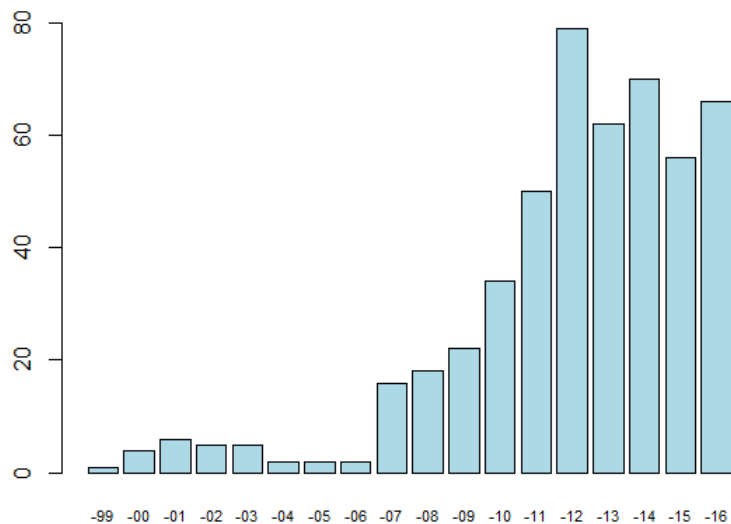
Jonkinlaista aineiston valikoituvuutta on voinut aiheuttaa se, että havaintoja ei ole tasaisesti jokaiselta vuodelta tarkasteluvälillä 2000–2016. Tämä johtuu siitä, että hovioikeuksien oli kaikista helpoin toimittaa uusimpia päätöksiä, jotka olivat sähköisessä muodossa. Selvästi suurin osa havainnoista on vuosilta 2010–2016 ja eniten havaintoja on vuodelta 2012. Myöskään hovioikeuksittain aineiston havainnot eivät jakaudu tasaisesti. Helsingin ja Turun hovioikeudet toimittivat muita hovioikeuksia enemmän sähköisiä asiakirjoja, mikä osittain selittää näiden suurempaa havaintomäärää.

Eräs mahdollinen ongelma aineiston satunnaisuudessa on valituslupaan liittyvässä seurantamenettelyssä ja jatkokäsittelyluvassa tapahtuneet lainsäädännölliset muutokset, jotka kuvailtiin johdantoluvussa. Aineistossa on erityisesti 1.10.2015 jälkeiseltä ajalta mukana sellaisia riitoja, joissa valittaja ei ole saanut jatkokäsittelylupaa hovioikeudelta. Tämä seikka on ilmennyt suoraan asiakirjoista. Tällaiset havainnot eivät siis ole karsiutuneet aineistosta pois. Sen sijaan epäselväksi on jäänyt, ovatko kaikki aineiston havainnot aikaväliltä 1.10.2003–1.11.2011 läpäisseet tuolloin käytössä olleen seurantamenettelyn. On luultavaa, että näin on, sillä toimitetuissa asiakirjoissa ei ollut mainintaa seurantamenettelystä. Seulonnan myötä karsitut havainnot ovat siis luultavasti jääneet kokonaan aineiston ulkopuolelle. Tämä voi tuoda oman harhansa analyysiin, mutta tämä seikka jätetään tässä työssä vain pohdinnan tasolle.

Muilta osin aineistoon ei ole valikoitunut havaintoja, vaan ne ovat päätyneet aineistoon satunnaisesti.

Osa hovioikeuksista toimitti aineistoa sähköisesti. Sähköisinä asiakirjoina pystyttiin toimittamaan ainoastaan uusimpia hovioikeuden ratkaisuja ja näistä tapauksista tausta-asiakirjat jäivät saamatta. Hovioikeuden ratkaisuihin oli kuitenkin aina liitettynä kärjäoikeuden ratkaisu, josta valitus oli tehty. Suurin osa keskeisimmistä muuttujista löytyi suoraan hovi- ja kärjäoikeuden ratkaisuista, mutta esimerkiksi lasten tai vanhempien ikätietoja tai olosuhdeselvityksen tietoja saattoi jäädä puuttumaan niistä havainnoista, joista tausta-asiakirjat puuttuivat.

Aineiston keruu oli kohtalaisen hidasta riippuen asiakirjojen sivumääristä. Erityisesti tapauksissa, joiden käsittely oikeudessa oli kestänyt pitkään, oikeudenkäyntiasiakirjojen sivumäärät tausta-asiakirjoineen saattoivat olla yhteensä useita satoja sivuja ja näistä keskeisten tietojen poimiminen saattoi kestää kauan. Tyypillisesti yhden havainnon kirjaamiseen meni noin puoli tuntia. Osa asiakirjoista oli ruotsinkielisiä ja niiden tulkitsemiseen saattoi kulua normaalia enemmän aikaa. Suuresta osasta tapauksia tehtiin varmuuden vuoksi myös tarkempia muistiinpanoja sen varalta, että niihin olisi tarvinnut palata jälkepäin. Useimmiten hovioikeuden kirjaamossa oli mahdollista tehdä töitä koko kirjaamon aukioloajan ja yhtä päivää kohti havaintoja saatiin paikan päällä kerättyä tyypillisesti noin 15 kappaletta.



Kuva 4.1: Havaintomäärät eri vuosilta

Ongelmiksi aineiston keruussa saattoi joidenkin muuttujien kohdalla muodostua se, että tietoa ei välttämättä löytynyt suoraan asiakirjoista, vaan niistä oli tehtävä tulkintoja. Jyrkkiä tulkintoja ei kuitenkaan tehty ja joskus tieto oli joidenkin muuttujien kohdalla jätettävä puuttuvaksi. Eniten puuttuvuutta oli vanhempien tulo- ja ikätiedoissa.

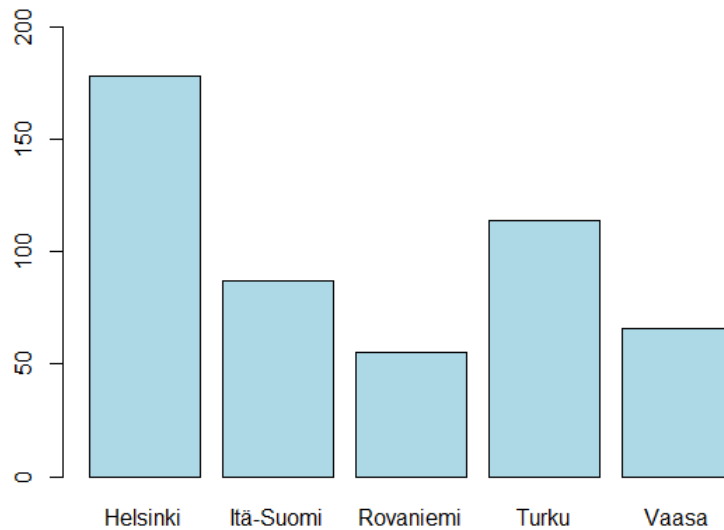
Asumisriita-aineisto koostuu 500 havainnosta ja jokaisessa havainnossa osapuolina ovat äiti ja isä. Näistä ainakin toinen on tehnyt valituksen lapsen huoltoon tai asumiseen liittyvästä käräjäoikeuden päätöksestä hovioikeuteen. On myös mahdollista, että molemmat ovat valittaneet käräjäoikeuden päätöksestä, mutta aineistossa valittajaksi on kirjattu aina se osapuoli, jonka vaatimuksena valituksessa on ollut lapsen asumisen muuttaminen hänen luokseen. Valittaja on siten aina käräjäoikeudessa asumisriidan hävinnyt osapuoli. Asumisriidoissa riidellään monesti myös huollosta ja myös huoltovaatimus sekä päätös huollosta on kirjattu ylös jokaisen havainnon kohdalla.

Aineiston viidestäsadasta asumisriidasta äiti on valittajana 250 riidassa ja isä 250 riidassa. Se, että äiti ja isä ovat aineistossa valittajina yhtä usein, on sattuma.

Kaikki aineiston muuttajat asteikkoineen ovat liitteessä A.

## 4.1 Päivämäärät ja ikämuuttajat

Havainnoista on kirjattu ylös sekä valituksen saapumispäivä hovioikeuteen että päätöksen antamispäivämäärä. Vanhin päätös on vuodelta 1999 ja uusin vuodelta 2016.



Kuva 4.2: Havaintomäärät eri hovioikeuksista

Vanhempien ja lasten syntymäajat on etsitty asiakirjoista ja hovioikeuksien tuomarien syntymäajat on poimittu Lakimiesmatrikelista (Laurila 2011). Kaikki ikätiedot on laskettu vähentämällä syntymäaika hovioikeuden päätöksen antamispäivämäärästä.

Ikätiedoissa on jonkin verran puuttuvuutta erityisesti havainnoissa, jotka on poimittu pelkästään sähköisistä asiakirjoista. Lasten ikätiedot olivat saatavilla noin 99 prosentissa havainnoista. Vanhempien ikätiedot olivat saatavilla noin 90 prosentissa havainnoista ja tuomarien ikätiedot noin 97 prosentissa havainnoista.

## 4.2 Ulkomaalaistaustaisuus

Tieto osapuolten ulkomaalaistaustaisuudesta on pyritty selvittämään mahdollisimman tarkasti suoraan asiakirjoista. Mikäli tieto on jäänyt epäselväksi, on tehty tulkinta osapuolen nimen perusteella. Liian rohkeita tulkintoja ei kuitenkaan ole tehty, eli vierasperäinen nimi ei aina ole johtanut ulkomaalaiseksi luokitteluun.

Noin joka viidennessä aineiston asumisriidassa ainakin toinen osapuoli on ulkomaalaistaustainen. Tällaisissa riidoissa yleisin asetelma on suomalainen isä vastaan ulkomaalaistaustainen äiti. Noin viidessä prosentissa kaikista riidoista molemmat osapuolet ovat ulkomaalaistaustaisia.

Palo-Revon aineistossa ulkomaalaistaustaisten vanhempien osuus oli 31 prosenttia eli vielä suurempi (Palo-Repo 2015, 48). Eroa selittänee osaltaan se, että Palo-Revon

Taulukko 4.1: Osapuolten ulkomaalaistaustaisuus

		Isä		Yht.
		Ei ulkomaalainen	Ulkomaalainen	
Äiti	Ei ulkomaalainen	402 (80,4%)	30 (6%)	432 (86,4%)
	Ulkomaalainen	42 (8,4%)	26 (5,2%)	68 (13,6%)
Yht.		444 (88,8%)	56 (11,2%)	500

aineisto oli kerätty ainoastaan Helsingin hovioikeuspiiristä, jossa ulkomaalaistaustaisen väestön osuus on oletettavasti suurempi kuin muualla Suomessa. Lisäksi eroa voi selittää myös se, että Palo-Revon aineistoa ei ollut rajoitettu vain asumisriitoihin kuten tässä tutkielmassa. On todennäköistä, että ulkomaalaistaustaisten vanhempien osuus on merkittävän suuri erityisesti sellaisissa riidoissa, joissa riidellään vain lasten huollosta, mutta ei asumisesta. Hyvin moni kirjaamisessa sivuutettu huoltoriita, jossa ei riideltä asumisesta, oli sellainen, jossa ainakin toinen osapuoli oli ulkomaalainen.

### 4.3 Edeltävä huolto- ja asumistilanne sekä tilanteen vakiintuminen

Edeltäväksi huolto- ja asumistilanteeksi on kirjattu se tilanne, joka on ollut vallitseva ennen valituksenalaista kärjäoikeuden päätöstä. Huolto- ja asumistilanne on kirjattu samaan muuttujaan. Lasten huolto on voinut perustua osapuolten tekemään sopimukseen, väliaikais määräykseen tai oikeuden päätökseen. Tästä on kirjattu tieto erilliseen muuttujaan.

Jos voimassa on ollut yhteishuolto, niin vallitsevaksi asumistilanteeksi on katsottu se, kenen luona lapsi on todellisuudessa asunut ennen kärjäoikeuden päätöstä, mikäli tieto on ilmennyt asiakirjoista. Lasten asumisesta on siis voinut olla voimassa oleva määräys tai sopimus, mutta asiakirjojen perusteella on yleistä, että lapset ovat tosiasiaa voineet asua päätöksen vastaisesti. Esimerkiksi jos voimassa olisi ollut päätös lasten asumisesta äidin luona, mutta lapset olisivat todellisuudessa asuneet koko ajan isänsä luona, olisi asumistilanteeksi kirjattu asuminen isän luona.

Jos vallitseva huoltotilanne on ollut yksinhuolto, niin yksinhuolto on kirjattu asumisesta riippumatta. Yksinhuollosta huolimatta lapset ovat kuitenkin voineet asua todellisuudessa koko ajan toisen vanhempansa luona.

Vuoroasuminen on kirjattu, mikäli lapset ovat tavanneet etävanhempansa vähintään 37 prosenttia ajasta.

Asumistilanteen on katsottu vakiintuneen toisen osapuolen luokse, jos lapset ovat yhtäjaksoisesti asuneet kärjäoikeuskäsittelyä edeltävässä asuinpaikassaan vähintään vuoden ajan. Asumisen vakiintuminen on kirjattu omaan muuttujaansa. Jos asumisaika ei ole ilmennyt asiakirjoista, on tilanne katsottu vakiintuneeksi osapuolen luokse, mikäli

tämä on selvästi ilmennyt esimerkiksi päätöksen perusteluista. Jos edeltävä asumistilanne on ollut vuoroasuminen, niin tilanteen ei ole katsottu vakiintuneen kummankaan osapuolen luokse.

Vakiintunut asumistilanne voi myös olla ristiriidassa edeltävän huolto- asumistilanteen kanssa, jos esimerkiksi lapset olisivat asuneet vuosia isän luona ja väliaikaismääräys olisi muuttanut tilanteen vuoroasumiseksi. Tällöin edeltävä huolto- ja asumistilanne olisi vuoroasuminen ja vakiintunut asumistilanne isän luona. Lapset ovat voineet myös asua toisen osapuolen luona vakiintuneesti, vaikka tällä ei edes olisi lasten huoltajuutta.

Taulukko 4.2: Edeltävä huolto- ja asumistilanne ja vakiintunut asumistilanne. Edeltävää huolto- ja asumistilannetta vastaavaa muuttujaa on karkeistettu.

	Vakiintuminen			Yht.
	Äidin luokse	Ei kummankaan luokse	Isän luokse	
<b>Edeltävä tilanne</b>				
Äidin yksinhuolto	32	4	4	40 (8%)
Yhteishuolto, asuminen äidillä	152	51	3	206 (41,2%)
Vuoroasuminen tai vanhempien yhteinen koti ennen eroa	3	107	7	117 (23,4%)
Yhteishuolto, asuminen isällä	3	36	80	119 (23,8%)
Isän yksinhuolto	0	4	14	18 (3,6%)
Yht.	190 (38,0%)	202 (40,4%)	108 (21,6%)	500

Lasten asuminen ennen käräjäoikeuden päätöstä on huomattavasti useammin äidin kuin isän luona. Myös äidin yksinhuolto on vallitsevana tilanteena yleisempi kuin isän yksinhuolto. Noin 60 prosentissa tapauksista tilanne on vakiintunut jommankumman osapuolen, useimmiten äidin luokse.

## 4.4 Osapuolten vaatimukset

Valittajan vaatimukseksi hovioikeudessa on kirjattu hänen ensisijainen vaatimuksensa lasten hullosta ja asumisesta. Koska aineistossa on vain asumisriitoja, on valittajan vaatimus aina käräjäoikeuden päätöksen muuttaminen vähintään siten, että ainakin yhden lapsen asuminen muutetaan tämän luokse. Myös vaatimus käräjäoikeudessa on kirjattu aineistoon.

Vastapuolen vaatimus hovioikeudessa on useimmiten suoraan sama kuin käräjäoikeuden päätös, koska vastapuoli hovioikeudessa on yleensä voittanut riidan asumisesta käräjäoikeudessa. Vaatimus saattaa kuitenkin poiketa käräjäoikeuden päätöksestä lapsen huollon osalta, mikäli käräjäoikeuden päätös lapsen hullosta ei ole ollut osapuolen vaatimuksen mukainen. Aineistoon on kirjattu myös vastapuolen vaatimus käräjäoikeudessa.

Yleisin asumisriita-asetelma on sellainen, missä riidellään ainoastaan lasten asuinpaikasta. Näissä siis kumpikin osapuoli vaatii yhteishuoltoa ja lasten asumista luonaan. Toiseksi yleisin asetelma on sellainen, missä kumpikin osapuoli vaatii lasten yksinhuol-

Taulukko 4.3: Osapuolten vaatimukset karkeistettuina. Yhdessä tapauksessa toisen osapuolen vaatimus on lapsen huostaanotto, jolloin tätä tapausta ei voi pitää varsinaisena asumisriitana.

	Isän vaatimus			Yht.
	Vuoroasuminen	Yhteishuolto, isän lähihuoltajuus	Isän yksinhuolto	
<b>Äidin vaatimus</b>				
Vuoroasuminen	4	8	4	16 (3,2%)
Yhteishuolto, äidin lähihuoltajuus	13	215	36	264 (52,9%)
Äidin yksinhuolto	6	87	126	219 (43,9%)
Yht.	23 (4,6%)	310 (62,1%)	166 (33,3%)	499

toa itselleen. Äidit vaativat isiä useammin lasten yksinhuoltoa. Vuoroasumista vaaditaan asumisriidoissa melko harvoin. (Taulukko 4.3.)

## 4.5 Oikeuden päätös ja riidan voittaja

Hovioikeuden päätökseksi on kirjattu päätös lapsen huollosta ja asumisesta samaan tapaan kuin edeltävä huolto- ja asumistilanne. Mikäli lapsen tapaamisoikeudeksi etävanhempaan on määrätty vähintään 37 prosenttia ajasta, päätös on kirjattu vuoroasumiseksi. Käräjäoikeuden päätös on kirjattu samalla tavalla.

Riidan voittajan voi määritellä usealla eri tavalla riippuen siitä, miltä kannalta riitaa tarkastelee. Voittajaksi voi esimerkiksi määritellä sen osapuolen, jonka vaatimus lasten asumisesta on hovioikeuden päätöksen mukainen. Jos hovioikeus on määrännyt vuoroasumisen, voittajaksi katsotaan tällä kriteerillä myös se osapuoli, jonka luokse lapsi on määrätty kirjoille. Tasapelejä olisivat sellaiset riidat, joissa hovioikeuden päätös on ollut joko lasten jako tai esimerkiksi huostaanotto. Päätös lasten jakamisesta tai huostaanotosta on tässä aineistossa varsin harvinainen, joten tällä tavalla määriteltäviä tasapelejä on aineistossa hyvin vähän.

Mikäli voittajaksi katsotaan se osapuoli, jonka asumisvaatimus on hovioikeuden päätöksen mukainen, niin aineistossa äidit voittavat riidoista 52,8 prosenttia ja isät 45,2 prosenttia. Vain kaksi prosenttia riidoista päättyy tällä kriteerillä tasapeliin.

Luvuissa 5 ja 6 sovitettavissa malleissa tasapeleihin lasketaan kuuluvaksi myös sellaiset riidat, joissa päätös on ollut vuoroasuminen riippumatta siitä kumman vanhemman luokse lapset on määrätty kirjoille.

Kun aineistosta poistetaan sellaiset riidat, joissa hovioikeuden päätöksenä on jokin muu kuin lasten asuminen jommankumman vanhempansa luona, esimerkiksi huostaanotto, niin jäljelle jää 494 asumisriitaa. Alla on tarkasteltu osapuolten vaatimuksia näissä riidoissa suhteessa hovioikeuden päätökseen.

Kun äiti vaatii yksinhuoltoa, niin noin 49 prosenttia päätöksistä on äidin vaatimuksen

Taulukko 4.4: Äidin vaatimus ja hovioikeuden päätös karkeistettuina. Vihreällä ovat ne tapaukset, joissa päätös on täysin äidin vaatimuksen mukainen tai parempi. Keltaisella ovat ne tapaukset, joissa päätös on ainakin osittain äidin vaatimuksen mukainen, mutta ei täysin. Punaisella ovat ne tapaukset, joissa päätös ei ole lainkaan äidin vaatimuksen mukainen.

	Äidin yksinhuolto	Yhteishuolto, äidin lähihuoltajuus	Lasten jako tai vuoroasuminen	Yhteishuolto, isän lähihuoltajuus	Isän yksinhuolto	Yht.
Vuoroasuminen	0	0	8	0	2	16
Vain lähihuoltajuus	1	124	12	107	18	262
Yksinhuolto	106	22	4	18	66	216
Yht.	107	146	24	131	86	494

Taulukko 4.5: Isän vaatimus ja hovioikeuden päätös karkeistettuina. Vihreällä ovat ne tapaukset, joissa päätös on täysin isän vaatimuksen mukainen tai parempi. Keltaisella ovat ne tapaukset, joissa päätös on ainakin osittain isän vaatimuksen mukainen, mutta ei täysin. Punaisella ovat ne tapaukset, joissa päätös ei ole lainkaan isän vaatimuksen mukainen.

	Äidin yksinhuolto	Yhteishuolto, asuminen äidin luona	Lasten jako tai vuoroasuminen	Yhteishuolto, isän lähihuoltajuus	Isän yksinhuolto	Yht.
Vuoroasuminen	5	6	12	0	0	23
Vain lähihuoltajuus	51	122	12	121	1	307
Yksinhuolto	51	18	0	10	85	164
Yht.	107	146	24	131	86	494

mukaisia. Jos isä puolestaan vaatii yksinhuoltoa, niin noin 52 prosenttia päätöksistä on isän vaatimuksen mukaisia. Kun äidin vaatimus koskee vain lähihuoltajuutta, eli lapsen asumista hänen luonaan, niin noin 48 prosenttia päätöksistä on äidin vaatimuksen mukaisia. Isällä vastaava osuus puolestaan on noin 40 prosenttia. Kun äidin vaatimus on vuoroasuminen, niin 50 prosenttia päätöksistä on äidin vaatimuksen mukaisia. Isällä vastaava osuus on noin 52 prosenttia. (Taulukot 4.4 ja 4.5.)

## 4.6 Olosuhdeselvityksen suositus

Olosuhdeselvityksiin liittyvät tiedot perustuvat aineiston keräämisessä käytettyjen oikeudenkäyntiasiakirjojen ohella lastenhuoltolakiin (Laki lapsen huollosta ja tapaamis-oikeudesta 361/1983) sekä Rovaniemen hovioikeuspiirin tuomioistuinten laatuhankeen laatimaan oppaaseen (Rovaniemen hovioikeuspiiri 2016).

Erityisesti asumisriidoissa oikeus voi tilata sosiaalilautakunnan suorittaman olosuhdeselvityksen. Selvitys on suurimmassa osassa tapauksia tehty käräjäoikeusvaiheessa, mutta myös hovioikeus voi tilata olosuhdeselvityksen, mikäli selvitystä ei ole ollut kärä-



jäoikeudessa tai mikäli olosuhteiden katsotaan muuttuneen oleellisesti. Uutta selvitystä ei välttämättä ole tilattu, mikäli osapuolet ovat riidelleet huollosta ja asumisesta oikeudessa aiemminkin ja aiempaan riitaan tilattu olosuhdeselvitys on ollut tuore.

Mikäli vanhemmat asuvat eri paikkakunnilla, olosuhdeselvityksen tekee kummankin asuinpaikkakunnan sosiaaliviranomaiset. Tieto olosuhdeselvityksestä on kirjattu, mikäli asiakirjoissa on maininta ainakin toisen vanhemman luona tehdystä selvityksestä.

Olosuhdeselvitys voi sisältää sosiaaliviranomaisten esittämän suosituksen huolto- ja asumistilanteeseen. Suosituksessa voidaan ottaa kantaa huoltoon tai asumiseen tai molempiin. Suositus ei aina ole suora ja usein suositusta ei esitetä olosuhdeselvityksessä lainkaan, vaan yhteenvedossa saatetaan esimerkiksi todeta lapsen voivan asua kumman tahansa vanhemman luona. Suositus on kirjattu aineistoon samalla asteikolla kuin oikeuden päätös. Koska olosuhdeselvityksiä voi olla kaksi, mahdollinen suositus on kirjattu sen mukaan, kumman vanhemman puolesta on enemmän suosituksia. Jos esimerkiksi toinen selvitys puoltaa äidin luona asumista ja toisessa ei esitetä suositusta, niin aineistoon on kirjattu suositus äidille. Jos taas suositukset menevät ristiin, niin suositusta ei ole kirjattu kummallekaan.

Taulukko 4.6: Olosuhdeselvitykset. Suositus asumisesta äidin tai isän luona kattaa myös sellaiset suositukset, joissa on suositeltu tämän osapuolen yksinhuoltoa.

Olosuhdeselvityksen suositus	Lkm
Suosituksena asuminen äidin luona	94
Suosituksena asuminen isän luona	101
Suosituksena kumman tahansa yksinhuolto	1
Suosituksena yhteishuolto ottamatta kantaa asumiseen	7
Suosituksena vuoroasuminen tai lasten jako	8
Ei suositusta tai suositukset menevät ristiin	231
Ei olosuhdeselvitystä	58

Äidit ja isät näyttävät pärjäävän olosuhdeselvityksissä lähes yhtä hyvin. Aineistossa isät saavat äitejä hieman enemmän olosuhdeselvitysten suosituksia. Suurimmassa osassa olosuhdeselvityksiä ei kuitenkaan esitetä suositusta tai mahdolliset kaksi suositusta menevät ristiin. Aineiston perusteella sosiaaliviranomaiset eivät siis näytä suosivan erityisemmin äitejä tai isiä. (Taulukko 4.6.)

## 4.7 Ratkaisijoiden ja avustajien tiedot

Hovioikeuden ratkaisun tekee kolmihenkinen tuomaristo. Mikäli tuomareiden ratkaisu ei ole yksimielinen, tuomarit suorittavat äänestyksen. Aineistossa päätökset olivat yksimielisiä jopa 98 prosentissa tapauksista. Mikäli joku ratkaisijoista on ollut eri mieltä, tästä on kirjattu tieto ylös.

Asiakirjoista ilmenee usein tieto osapuolen avustajasta oikeudenkäynnissä. On myös mahdollista, että osapuolella ei ole ollut avustajaa. Avustajaksi on merkitty joko julkinen oikeusavustaja, asianajaja, muu lakimies tai tieto avustajan puuttumisesta. Mikäli osapuolella on ollut avustaja, on aineistoon kirjattu myös avustajan sukupuoli. Mikäli asiakirjassa ei ole löytynyt mainintaa osapuolen avustajasta, tieto avustajasta on merkitty puuttuvaksi. Puuttuva tieto voi siis tarkoittaa sitä, että osapuolella ei ole ollut avustajaa tai sitä, että tietoa avustajasta ei ole löytynyt.

## 4.8 Lasten tiedot

Aineistoon on kirjattu tieto vain niistä lapsista, joiden asumisesta osapuolet ovat riidelleet hovioikeudessa. Vaikka osapuolilla olisi ollut käräjäoikeudessa riitaa useammasta lapsesta, näiden tietoja ei ole kirjattu mukaan. Jokaisesta lapsesta, joista osapuolet ovat riidelleet hovioikeudessa, on kirjattu aineistoon ikä ja sukupuoli.

## 4.9 Vanhempien taustatiedot

Kummastakin osapuolesta on kirjattu tieto heidän pääasiallisesta toiminnastaan sekä nettotuloista, mikäli ne ovat olleet saatavilla. Nettotulotietoja oli usein saatavilla vain sellaisista tapauksista, joissa vanhemmat riitelivät myös lasten elatusavusta. Tulotiedot löytyivät vain noin 47 prosentissa tapauksista.

## 4.10 Osapuolten esittämät syytökset

On tavallista, että jompikumpi osapuoli esittää syytöksen toista osapuolta kohtaan. Syytös ilmenee usein selvästi osapuolen valituksen tai vastauksen perusteluissa hovi- tai käräjäoikeuteen, mutta se on voinut ilmetä myös tausta-asiakirjoista.

Aineistossa syytökset on jaettu neljään kategoriaan. Ensimmäinen kategoria on syytös väkivalta-, päihde- tai mielenterveysongelmista. Syytös on kirjattu, mikäli osapuoli on esittänyt toista osapuolta tai tämän uutta puolisoa kohtaan tämän syytöksen. Toinen kategoria on syytös tapaamisten estämisestä tai lapsen vieraannuttamisesta toisesta vanhemmastaan. Kolmas kategoria on uhka lapsen kaappaamisesta ulkomaille. Neljäs kategoria on syytös lapsen seksuaalisesta hyväksikäytöstä. Tämä syytös on kirjattu, mikäli syytös on esitetty toista osapuolta tai tämän uutta puolisoa kohtaan.

Aineistoon on kirjattu myös erikseen jokaisen syytöskategorian osalta tieto siitä, onko oikeus todennut osapuolen tekemän syytöksen toteennäytetyksi. Syytös on kirjattu toteennäytetyksi, mikäli joko käräjä- tai hovioikeus on todennut asian päätöksen perusteluissa. Jos vain käräjäoikeus on katsonut syytöksen toteennäytetyksi, mutta hovioikeuden päätöksen perusteluista ilmenee, että hovioikeus on katsonut asian toisin, niin syytöstä ei ole kirjattu toteennäytetyksi.

Taulukko 4.7: Osapuolten esittämät syytökset ja niiden osuudet kaikista tapauksista. Syytös 1: Syytös toista osapuolta tai tämän uutta puolisoa kohtaan väkivaltaisuudesta, päihteidenkäytöstä tai mielenterveysongelmista. Syytös 2: Syytös toista osapuolta kohtaan tapaamisten estämisestä tai lasten vieraannuttamisesta. Syytös 3: Syytös lasten kaappausuhasta ulkomaille. Syytös 4: Syytös toista osapuolta tai tämän uutta puolisoa kohtaan lasten seksuaalisesta hyväksikäytöstä.

	Syytös 1	Syytös 2	Syytös 3	Syytös 4
Ei syytöksiä	218 (43,6%)	294 (58,8%)	457 (91,4%)	474 (94,8%)
Vain isä syyttää äitiä	104 (20,8%)	115 (23,0%)	29 (5,8%)	3 (0,6%)
Vain äiti syyttää isää	90 (18,0%)	65 (13,0%)	12 (2,4%)	22 (4,4%)
Molemmat syyttävät toisiaan	88 (17,6%)	26 (5,2%)	2 (0,4%)	1 (0,2%)

Taulukko 4.8: Oikeuden todeksi katsomat syytökset ja niiden osuudet kaikista tapauksista.

	Syytös 1	Syytös 2	Syytös 3	Syytös 4
Ei kumpaakaan kohtaan	439 (87,8%)	435 (87,0%)	490 (98,0%)	499 (99,8%)
Äitiä kohtaan	40 (8,0%)	37 (7,4%)	8 (1,6%)	1 (0,2%)
Isää kohtaan	20 (4,0%)	27 (5,4%)	2 (0,4%)	0 (0%)
Molempia kohtaan	1 (0,2%)	1 (0,2%)	0 (0%)	0 (0%)

Väkivalta-, päihteidenkäyttö- tai mielenterveys-syytöksiä esitetään useammassa kuin joka toisessa riidassa. Myös syytöksiä lasten tapaamisten estämisestä tai vieraannuttamisesta esitetään paljon. Isät esittävät yleisesti syytöksiä enemmän kuin äidit, mutta äidit esittävät huomattavasti isää useammin syytöksen lasten seksuaalisesta hyväksikäytöstä. (Taulukko 4.7.)

Oikeuden toteamia syytöksiä on huomattavasti vähemmän kuin esitettyjä syytöksiä. Mielenkiintoinen huomio on, että isää tai tämän uutta puolisoa kohtaan tehtyä syytöstä lasten seksuaalisesta hyväksikäytöstä ei kertaakaan ole katsottu oikeudessa toteennäytetyksi, vaikka äiti on esittänyt tällaisen syytöksen 22 tapauksessa. (Taulukot 4.7 ja 4.8.)

## 4.11 Esimerkki muuttujien kirjaamisesta

Vanhemmat oli tuomittu avioeroon vuonna 2001. Avioerosta huolimatta vanhemmat olivat asuneet yhdessä syyskuuhun 2008 saakka, jolloin isä oli muuttanut pois heidän yhteisestä asunnosta. Vanhempien kaksi lasta, joita kutsutaan tässä tytöksi ja pojaksi, olivat ajoittain asuneet pitkiäkin aikoja äidin vanhempien luona, mutta olivat muuttaneet isän luo syksyllä 2008.

Käräjäoikeus oli ensin antanut asiasta väliaikais määräyksen keväällä 2009. Sen mukaan tytön tuli asua toistaiseksi äidin luona ja pojan isän luona, kunnes varsinainen päätös annettaisiin. Äiti oli käräjäoikeudessa vaatinut, että molemmat lapset jäävät yh-

teishuoltoon ja asumaan hänen luonaan. Toissijaisesti hän oli vaatinut, että tyttö oli joka tapauksessa määrättävä asumaan hänen luonaan. Perusteinaan äiti oli esittänyt, että lasten isä ei tue lasten yhteydenpitoa äitiin ja että tytön muuttaminen isän luokse ei olisi millään tavalla lapsen edun mukaista. Isä puolestaan oli vaatinut käräjäoikeudessa, että kumpikin lapsi oli määrättävä yksin hänen huoltoonsa ja asumaan hänen luonaan. Isä oli perustellut vaatimustaan sillä, että kyseessä oli lasten oma tahto.

Käräjäoikeus oli katsonut vanhempien riitaisten välien vuoksi, että yhteishuollolle ei ollut asiassa edellytyksiä. Olosuhdeselvityksen mukaan asiassa tuli kummankin lapsen iän vuoksi kuulla lasten omia mielipiteitä, joka oli olosuhdeselvityksen mukaan isän luona asuminen. Käräjäoikeus oli katsonut tämän olevan kummankin lapsen harkittu mielipide ja oli määrännyt kummankin lapsen isän yksinhuoltoon ja asumaan hänen luonaan.

Hävinneenä osapuolena äiti oli valittanut asiasta hovioikeuteen. Äidin vaatimus hovioikeudessa oli, että lapset määrätään vanhempien yhteishuoltoon ja tyttö määrätään asumaan hänen luonaan. Äidin mukaan oli kyseenalaista perustaa päätös asumisesta lapsen omaan tahtoon, koska tyttö oli vasta 10-vuotias. Isä oli hovioikeudessa vaatinut äidin valituksen hylkäämistä ja perustellut tätä sillä, että tytön asuessa isän luona tapaamiset äidin kanssa olivat toteutuneet hyvin.

Hovioikeus oli käräjäoikeuden tavoin katsonut, että yhteishuolto ei ollut lasten edun mukaista. Hovioikeus oli myös katsonut, että asiassa ei ollut tullut ilmi seikkoja, joiden mukaan jompikumpi vanhemmista olisi kykenemätön toimimaan huoltajana. Hovioikeus hyväksyi käräjäoikeuden ratkaisun ja lapset jäivät siten isän yksinhuoltoon.

Esimerkkitapauksessa valittajaksi kirjattiin äiti. Koska hovioikeudessa riita koski tyttöä, edeltävä huoltotilanne kirjattiin sen mukaan, mikä se oli tytön osalta ennen käräjäoikeuden päätöstä. Asumistilanne ennen käräjäoikeuden päätöstä oli siis väliaikaismääräys, jonka mukaan tyttö asui äidin luona. Edeltäväksi huoltotilanteeksi kirjattiin siten yhteishuolto, jossa asuminen oli äidin luona. Tilannetta ei kirjattu vakiintuneeksi, koska tyttö ei ollut yhtäjaksoisesti asunut kummankaan luona vähintään vuoden ajan. Koska hovioikeuden päätös oli isän hovioikeusvaatimuksen mukainen, niin isän voi katsoa voittaneen riidan. Havaintoon kirjattiin äidin esittämä syytös isää kohtaan tapaamisten rajoittamisesta. Olosuhdeselvityksen suositukseksi kirjattiin asuminen isän luona.

## Luku 5

# Ilmiön mallintaminen

### 5.1 Mallin käyttöperuste ja oletukset

Oikeuden päätös asumisriidassa on luokiteltavissa symmetrisesti äidin voittoon, tasapeliin ja isän voittoon. Tuloluokkien määrä voi vaihdella riippuen siitä, miten tarkasti oikeuden päätös luokitellaan. Oli luokittelu miten tarkka hyvänsä, tuloluokat on huomioitava sekä äidin että isän voittotapauksissa, jolloin asteikko pysyy symmetrisenä. Tuloluokkien välillä on myös selkeä järjestys luokkien määrästä riippumatta. Jos esimerkiksi asteikon alkupäää käsittää äidin voittoluokat ja loppupäää isän voittoluokat, niin asteikon alkupäästä loppupäähän siirtyminen merkitsee isän kannalta parempaa tulosta ja päin vastoin. Vastemuuttujan luonteen kannalta järjestysregressiomallin käyttäminen ilmiön mallintamiseen on siten perusteltua. Järjestysasteikon käyttäminen tuo enemmän informaatiota selitettävästä muuttujasta kuin pelkästään voittajan määrittely dikotomiseti ja se mahdollistaa myös tasapelien ottamisen mukaan malliin.

Tavallisessa kertymä-logit-regressiomallissa oletetaan, että selittäjän arvon muuttuminen vaikuttaa jokaiseen kertymävastasuhteeseen samansuuruisesti vastemuuttujan luokasta riippumatta. Jos esimerkiksi yksittäisen selittäjän vaikutus  $\beta_k > 0$ , niin jokainen kertymävastasuhte pienenee toisiinsa nähden saman verran selittäjän arvojen kasvaessa, ja kertymävastasuhteiden välinen etäisyys eri luokkien välillä pysyy koko ajan samana selittäjän arvoista riippumatta. Tämä oletus on keskeinen mallin toimivuuden kannalta. Oletuksen realistisuuden perustelu asumisriitojen lopputuloksen mallintamisessa ei kuitenkaan ole yksinkertaista ja se riippuu tietenkin myös selittäjistä. Helpompaa on löytää yksittäisiä selittäjiä, joilla tämä oletus ei päde. Oikeuden päätös voi esimerkiksi rajoittaa merkittävästi sen mukaan, mitä riidan osapuolet vaativat, jolloin oletus on selvästi kyseenalainen, kun selittäjänä käytetään osapuolten vaatimuksia. Tämä ongelma pyritään korjaamaan skaala- ja nominaaliselittäjillä jaksossa 5.3.

## 5.2 Vastemuuttujan määrittely

Aineisto on jo keruuvaiheessa rajoitettu koostumaan riidoista, joissa osapuolet ovat eri mieltä vähintään lapsen asumisesta, mutta mahdollisesti myös huollosta. Aineiston muuttuja, joka kertoo hovioikeuden päätöksen, ei sellaisenaan voi toimia selitettävänä muuttujana, sillä hovioikeuden päätös on voinut olla asian palauttaminen käräjäoikeuteen, lapsen huostaanotto tai muu päätös, jota ei varsinaisesti voi pitää asumisriidan lopputuloksena. Tällaisia päätöksiä ei myöskään kannata luokitella tasapeleiksi, jos tasapeleihin halutaan luokitella kuitenkin sellaiset päätökset, joissa päätös on lasten jako tai vuoroasuminen. Kun havainnoista karsitaan pois sellaiset riidat, joissa hovioikeuden päätös on ollut niin sanottu muu päätös, niin jäljelle jää 494 asumisriitaa.

Kun hovioikeuden päätös luokitellaan mahdollisimman tarkasti, niin päätös voi olla äidin yksinhuolto (1), äidin yksinhuolto isän tiedonsaantioikeudella (2), yhteishuolto tehtävänjakomääräyksin siten, että lasten asuminen on äidin luona (3), normaali yhteishuolto siten, että lasten asuminen on äidin luona (4), lasten jako (5), vuoroasuminen (6), normaali yhteishuolto siten, että lasten asuminen on isällä (7), yhteishuolto tehtävänjakomääräyksin siten, että lasten asuminen on isällä (8), isän yksinhuolto äidin tiedonsaantioikeudella (9) tai isän yksinhuolto (10). Kun selitettävän muuttujan tasapeliluokaksi valitaan lasten jako ja vuoroasuminen, eli alkuperäisen muuttujan luokat 5 ja 6, niin selitettävässä muuttujassa voi käytännössä olla joko kolme, viisi, seitsemän tai yhdeksän luokkaa siten, että tasapeliluokka on keskimäinen.

Kolmiluokkaisessa jaossa toinen ääriluokka edustaa kaikkia äidin voittoluokkia eli alkuperäisiä luokkia 1, 2, 3 ja 4. Toinen ääriluokka puolestaan edustaa kaikkia isän voittoluokkia eli alkuperäisiä luokkia 7, 8, 9 ja 10.

Viisiluokkaisessa jaossa tehdään asumisriidan voittajan erottelun lisäksi ero yksin- ja yhteishuollon välillä. Tällaisessa jaossa ensimmäinen luokka kattaa alkuperäiset luokat 1 ja 2, toinen luokka alkuperäiset luokat 3 ja 4, kolmas luokka alkuperäiset luokat 5 ja 6, neljäs luokka alkuperäiset luokat 7 ja 8 ja viides luokka alkuperäiset luokat 9 ja 10. Toinen vaihtoehto on ottaa yksinhuoltoluokkaan myös yhteishuolto tehtävänjakomääräyksin, eli rajoitettu yhteishuolto. Tällöin alkuperäisen muuttujan luokat 1, 2 ja 3 yhdistetään ensimmäiseksi luokaksi, luokka 4 toiseksi, luokat 5 ja 6 kolmanneksi, luokka 7 neljänneksi ja luokat 8, 9 ja 10 viidenneksi.

Seitsenluokkaisessa jaossa on kaksi tulkinallisesti järkevää jaottelutapaa. Toisessa tavassa ensimmäinen luokka kattaa alkuperäisen luokan 1, toinen luokka kattaa alkuperäiset luokat 2 ja 3, kolmas luokka kattaa alkuperäisen luokan 4, neljäs luokka kattaa alkuperäiset luokat 5 ja 6, viides luokka kattaa alkuperäisen luokan 7, kuudes luokka kattaa alkuperäiset luokat 8 ja 9 ja seitsemäs luokka kattaa alkuperäisen luokan 10. Vaihtoehtoisesti seitsenluokkainen muuttuja voi olla myös sellainen, että ensimmäinen luokka kattaa alkuperäiset luokat 1 ja 2, toinen luokka alkuperäisen luokan 3, kuudes luokka alkuperäisen luokan 8, seitsemäs luokka alkuperäiset luokat 9 ja 10 ja muut luokat

ovat kuten edellä.

Kaikista hienojakoisimmalla yhdeksän luokan jaolla ainoastaan alkuperäiset luokat 5 ja 6 yhdistetään tasapeliluokaksi ja muita luokkia ei yhdistetä lainkaan.

Taulukko 5.1: Hovioikeuden päätösten jakautuminen

	Lukumäärä
Äidin yksinhuolto (1)	61
Äidin yksinhuolto isän tiedonsaantioikeudella (2)	27
Yhteishuolto tehtävänjakomääräyksin, asuminen äidillä (3)	19
Normaali yhteishuolto, asuminen äidillä (4)	146
Lasten jako (5)	7
Vuoroasuminen (6)	17
Normaali yhteishuolto, asuminen isällä (7)	131
Yhteishuolto tehtävänjakomääräyksin, asuminen isällä (8)	12
Isän yksinhuolto äidin tiedonsaantioikeudella (9)	33
Isän yksinhuolto (10)	41

Havaintojen lukumäärä on keskeinen rajoittava tekijä vastemuuttujan luokkien lukumäärälle. Jos jossakin vastemuuttujan luokassa on liian vähän havaintoja, niin malli voi toimia huonosti. Mallin on kuitenkin oltava sellainen, että sen luokilla on selkeä käytännön tulkinta.

Tasapeliluokka jää havaintomäärältään pieneksi luokkien lukumäärästä riippumatta, mutta tähän liittyvä ongelma pyritään korjaamaan jaksossa 5.3. Seitseen- tai yhdeksänluokkaisella vastemuuttujalla muutkin luokat voivat jäädä havaintomäärältään pieneksi, joten viisi- tai kolmiluokkainen vastemuuttuja on tässä mielessä parempi. Viisiluokkainen vastemuuttuja johtaa tulkinnaltaan varsin selkeään malliin, jossa asumisriidan voittajan lisäksi tehdään myös ero yhteis- ja yksinhuollon välillä. Kolmiluokkainen jako ei lainkaan huomioi eroa yhteis- ja yksinhuollon välillä, jolloin tällainen aineiston tuoma mahdollisesti arvokas informaatio voisi mennä hukkaan. Viisiluokkaisella jaolla havaintojen lukumäärät eri luokissa pysyvät myös kohtalaisen suurina.

Koska alkuperäisessä muuttujassa on varsin paljon havaintoja luokissa 4 ja 7 (normaali yhteishuolto), niin näitä ei välttämättä kannata yhdistää luokkiin 3 ja 8 (rajoitettu yhteishuolto), sillä tällöin yhdistetty luokka voi kasvaa turhan suureksi suhteessa muihin luokkiin. Käytännön tulkinnan kannalta keskeinen kysymys on se, että onko rajoitettu yhteishuolto lähempänä yksinhuoltoa vai yhteishuoltoa. Rajoitettu yhteishuolto voi tarkoittaa esimerkiksi sitä, että toinen vanhempi saa yksin päättää lapsen koulunkäynnistä, mutta kaikesta muusta vanhemmat päättävät yhdessä. Toisaalta rajoitettu yhteishuolto voi tarkoittaa myös sitä, että toinen vanhempi päättää yksin kaikista lapseen liittyvistä asioista, mutta vanhemmat päättävät yhdessä esimerkiksi lapsen passista. Tehtävänjakomääräysten määrän vaihtelun perusteella rajoitetun yhteishuollon voisi

luokitella kumpaan tahansa luokkaan. Koska rajoitetussa yhteishuollossa toinen vanhemmista asetetaan kuitenkin jossain määrin toisen edelle ja koska yhteishuoltoon luokittelu voi kasvattaa yhteishuoltoluokan havaintomäärän liian suureksi suhteessa muihin luokkiin, niin rajoitettu yhteishuolto päädytään luokittelemaan vastemuuttujassa yksinhuoltoluokkaan.

Jatkossa esitettävien mallien luokkia merkitään järjestysasteikolla luvuin 1–5, missä 1 tarkoittaa äidin yksinhuoltoa, 2 tarkoittaa normaalia yhteishuoltoa ja äidin lähihuoltajuutta, 3 tarkoittaa lasten jakoa tai vuoroasumista, 4 tarkoittaa normaalia yhteishuoltoa ja isän lähihuoltajuutta ja 5 tarkoittaa isän yksinhuoltoa. On huomattava, että yksinhuolto kattaa jatkossa myös yksinhuollon toisen vanhemman tiedonsaantioikeudella sekä yhteishuollon tehtävänjakomääräyksin.

### 5.3 Osapuolten vaatimusten huomioiminen

On syytä olettaa, että oikeuden päätös riippuu merkittäväällä tavalla osapuolten esittämistä vaatimuksista. Jos esimerkiksi kumpikaan osapuoli ei vaadi lasten yksinhuoltoa, niin on hyvin epätodennäköistä, että oikeus määräisi lapset kummankaan vanhemman yksinhuoltoon. Toisaalta mikäli jompikumpi osapuolista vaatisi vuoroasumista, niin tasapelin todennäköisyys oletettavasti kasvaisi. Osapuolten vaatimukset on siis jollakin tavalla huomioitava mallissa, jotta malli voisi toimia järkevällä tavalla.

Vaatimusten ottaminen mukaan tavallisiksi selittäjiksi malliin ei vaikuta järkevältä, sillä kertymä-logit-malli olettaa tavallisten selittäjän vaikutuksen yhtä suureksi jokaisen luokan välillä. Sen sijaan tällaisessa tilanteessa taustalla olevan piilomuuttujan skaalan mallintaminen osapuolten vaatimuksilla voi olla järkevää, koska tällöin esimerkiksi äärimmäiset yksinhuoltoluokat voidaan saada todennäköisemmiksi, kun yksinhuoltoa vaaditaan. Jos taas kumpikaan osapuoli ei vaadi yksinhuoltoa, niin selitettävän muuttujan keskimmäiset luokat voidaan saada todennäköisemmiksi ja ääri luokat käytännössä mahdolltomiksi. Tällä perusteella yksinhuoltovaatimus voisi toimia hyvänä skaalaselittäjänä.

Aineiston asumisriidoissa vuoroasumista tai lasten jakoa määrätään harvoin. Osapuolen vuoroasumisvaatimus on niin ikään aineistossa melko harvinainen, mutta mikäli sitä jompikumpi osapuoli vaatii, niin vuoroasuminen määrätään paljon useammin kuin muissa tapauksissa.

Taulukko 5.2: Tasapelin yleisyys eri vaatimuksilla

	Päätös lasten jako tai vuoroasuminen	Muu päätös	Yht.
Ei vuoroasumisvaatimusta	8 (1,7%)	451 (98,3%)	459
Vuoroasumisvaatimus	16 (45,7%)	19 (54,3%)	35
Yht.	24 (4,9%)	470 (95,1%)	494



Taulukon 5.2 perusteella tasapeliluokan yleisyys on huomioitava mallissa, kun vuoroasumista vaaditaan. Kun jompikumpi osapuoli vaatii vuoroasumista, niin tasapeliluokan suhteellinen osuus kasvaa lähes 27-kertaiseksi verrattuna tilanteeseen, jossa kumpikaan osapuoli ei vaadi vuoroasumista. On syytä olettaa, että vuoroasumisvaatimuksen ottaminen skaalaselittäjäksi ei merkittävästi parantaisi mallia, sillä tasapeliluokka jäisi joka tapauksessa pieneksi luokaksi. Vaikka piilomuuttujan varianssi olisi hyvin pieni, niin kynnyksparametrit olisivat joka tapauksessa lähellä toisiaan. Tällaisessa tilanteessa voisi olla perusteltua sallia kynnyksparametrien riippua siitä, onko vuoroasumista vaadittu, eli vuoroasumisvaatimus voisi toimia hyvänä nominaaliselittäjänä.

Sopivia skaala- ja nominaaliselittäjiä voidaan tutkia uskottavuusosamäärätestillä vertaamalla skaala- tai nominaaliselittäjillä varustettua mallia sellaiseen malliin, jossa nämä selittäjät toimivat tavallisina selittäjinä. Vertailtavissa malleissa käytetään viisi-luokkaista vastemuuttujaa. Kumpikin selittäjistä on indikaattorimuuttuja, eli joko 0 tai 1 sen mukaan, onko jompikumpi osapuolista vaatinut yksinhuoltoa tai vuoroasumista

Taulukko 5.3: Uskottavuusosamäärätestit skaalaselittäjille

Selittäjä	df	LRT	$p$ -arvo	$P(\chi_{df}^2 \geq \text{LRT})$
vaat_yh	1	371,22		$< 10^{-15}$
vaat_vuoroas	1	14,08		0,0002

Taulukko 5.4: Uskottavuusosamäärätestit nominaaliselittäjille

Selittäjä	df	LRT	$p$ -arvo	$P(\chi_{df}^2 \geq \text{LRT})$
vaat_yh	3	375,34		$< 10^{-15}$
vaat_vuoroas	3	63,57		$< 10^{-12}$

Uskottavuusosamäärätestit vahvistavat sen, että osapuolten vaatimukset kannattaa huomioida skaala- tai nominaaliselittäjien muodossa, sillä nämä parantavat mallia (taulukot 5.3 ja 5.4).

Nominaaliselittäjiksi ei kannata valita liikaa selittäjiä, sillä tämä johtaa helposti estimointiongelmiiin (Christensen 2018a, 12). Liian suuri nominaaliselittäjien määrä myös hankaloittaa mallin tulkintaa. On siis luultavasti järkevämpää käyttää joko molempia skaalaselittäjinä tai toista skaalaselittäjänä ja toista nominaaliselittäjänä.

Malli, jossa vuoroasumisvaatimusta käytetään skaalaselittäjänä ja yksinhuoltovaatimusta nominaaliselittäjänä, johtaa singulaariseen Hessen matriisiin, joten tällainen malli on hylättävä. Intuitiivisestikaan tällainen malli ei olisi ollut järkevä. Ensimmäinen vaihtoehto on siis käyttää mallia, jossa molemmat toimivat skaalaselittäjiä. Toinen vaihtoehto on käyttää mallia, jossa yksinhuoltovaatimus toimii skaalaselittäjänä ja vuoroasumisvaatimus nominaaliselittäjänä. Toinen vaihtoehto on siis alkuperäisen intuition mukainen malli. Verrataan näitä kahta mallia uskottavuusosamäärätestillä.

Taulukko 5.5: Kahden mallin vertailu uskottavuusosamäärätestillä

	Skaalaselittäjä	Nominaaliselittäjä	logLik	df	LRT	$p$ -arvo
Malli A	vaat_yh+vaat_vuoroas	-	-538,72			
Malli B	vaat_yh	vaat_vuoroas	-523,29	3	30,86	$< 10^{-6}$

Uskottavuusosamäärätestin perusteella malli B, eli malli jossa skaalaselittäjä on yksinhuoltovaatimus ja nominaaliselittäjä on vuoroasumisvaatimus, on näistä kahdesta parempi (taulukko 5.5).

Taulukko 5.6: Skaalaselittäjän estimaatti ja Waldin testitulos.

	$\hat{\zeta}$	Keskivirhe	Wald	$p$ -arvo
vaat_yh	2,72	0,29	9,55	$< 10^{-15}$

Kun testataan tällä mallilla nollahypoteesia, jonka mukaan skaalaselittäjä  $\zeta$  on nolla, saadaan Waldin testissä hyvin pieni  $p$ -arvo (taulukko 5.6). Yksinhuoltovaatimus on siis tilastollisesti erittäin merkitsevä skaalaselittäjä tällaisessa mallissa. Kun mukana on vain nominaaliselittäjä eikä yhtään varsinaisia selittäjiä, niin skaalaparametrin estimoitu arvo on selvästi positiivinen. Koska skaalaselittäjä on indikaattorimuuttuja, niin se tarkoittaa, että vastemuuttujan hajonta kasvaa silloin, kun yksinhuoltovaatimus on esitetty.

Mallin kynnysparametrit  $\theta_j$  määräävät kunkin luokan  $j$  todennäköisyydet, kun mallin muiden parametrien arvot on kiinnitetty. Mitä kauempana kaksi kynnysparametria toisistaan ovat, sitä suurempi on näiden väliin jäävän luokan todennäköisyys. Kun mallissa on nominaaliselittäjänä vuoroasumisvaatimus, niin mallin kynnysparametreja tarkastellaan laajennettuina kynnysparametreina  $\tilde{\theta}_j$  siten, että  $\tilde{\theta}_j = \theta_j$  silloin, kun vuoroasumisvaatimusta ei ole esitetty ja  $\tilde{\theta}_j = \theta_j - \tilde{\beta}_j$  silloin, kun vuoroasumisvaatimus on esitetty. Tässä  $\tilde{\beta}_j$  tulkitaan vuoroasumisvaatimuksen vaikutukseksi luokan  $j$  kertymätodennäköisyyteen.

Kynnysparametrien sekä vuoroasumisvaatimuksen vaikutuksen estimaatit mallissa, jossa on vain skaala- ja nominaaliselittäjät on esitetty taulukoissa 5.7 ja 5.8. Kun mallissa on mukana vain skaala- ja nominaaliselittäjä, niin estimoitu luokan  $j$  todennäköisyys havainnolle, jossa vuoroasumisvaatimus on  $w$  (0 tai 1) ja yksinhuoltovaatimus on  $z$  (0 tai 1), saadaan kaavalla

$$\hat{\pi}_j = \frac{\exp(\hat{\eta}_j)}{1 + \exp(\hat{\eta}_j)} - \frac{\exp(\hat{\eta}_{j-1})}{1 + \exp(\hat{\eta}_{j-1})}, \quad j = 1, 2, 3, 4, 5,$$

missä

$$\hat{\eta}_j = \frac{\hat{\theta}_j - w\tilde{\beta}_j}{\exp(z\hat{\zeta})}.$$

Estimoidut todennäköisyydet kullekin luokalle eri vaatimussyhdistelmällä on esitetty taulukossa 5.9. Kun yksinhuoltovaatimusta tai vuoroasumisvaatimusta ei esitetä, niin

luokat 2 ja 4, eli yhteishuoltoluokat ovat todennäköisimpiä ja yksinhuoltoluokat 1 ja 5 käytännössä mahdottomia. Tämä on toivottavaa, sillä oikeus ei käytännössä koskaan määrää yksinhuoltoa, jos kumpikaan osapuoli ei sitä vaadi. Kun jompikumpi osapuoli esittää yksinhuoltovaatimuksen ja vuoroasumisvaatimusta ei esitetä, niin luokat 1 ja 5 muuttuvat todennäköisemmiksi ja luokkien 2 ja 4 todennäköisyydet pienenevät. Myös tämä on toivottava ominaisuus mallilta. Kun jompikumpi osapuoli esittää vuoroasumisvaatimuksen ja yksinhuoltovaatimusta ei esitetä niin tasapeliluokan 3 todennäköisyys kasvaa merkittävästi. Kun toinen osapuoli vaatii yksinhuoltoa ja toinen vuoroasumista, niin yksinhuoltoluokat pysyvät todennäköisimpinä ja tasapeliluokka on vuoroasumisvaatimuksesta huolimatta kaikista epätodennäköisin.

Taulukko 5.7: Kynnysparametrien estimaatit mallissa, jossa on vain skaalaselittäjä ja nominaaliselittäjä.

Kynnys	$\hat{\theta}_j$	Keskivirhe
1 2	-8,10	2,43
2 3	-0,04	0,14
3 4	0,10	0,14
4 5	6,70	1,76

Taulukko 5.8: Estimoidut vuoroasumisvaatimuksen vaikutukset kynnysparametreihin.

Kynnys	$\hat{\beta}_j$	Keskivirhe
1 2	3,18	3,13
2 3	-1,14	0,49
3 4	1,61	0,57
4 5	11,21	10,60

Taulukko 5.9: Pyöristetyt estimoidut todennäköisyydet eri vaatimussyhdteillä mallissa, jossa on vain skaalaselittäjä ja nominaaliselittäjä.

Yksinhuoltovaatimus	Vuoroasumisvaatimus	$\hat{\pi}_1$	$\hat{\pi}_2$	$\hat{\pi}_3$	$\hat{\pi}_4$	$\hat{\pi}_5$
Ei	Ei	0	0,49	0,034	0,474	0,001
Kyllä	Ei	0,37	0,129	0,002	0,107	0,392
Ei	Kyllä	0,007	0,227	0,612	0,153	0
Kyllä	Kyllä	0,42	0,061	0,047	0,236	0,236

Malli vaikuttaa siis ottavan osapuolten vaatimukset halutulla tavalla huomioon skaala- ja nominaaliselittäjien muodossa. Huomattavaa kuitenkin on, että nominaaliselittäjällä keskivirheet ovat suuria luokkien 1 ja 5 kynnyksillä. Tämä tarkoittaa, että vuoroasumisvaatimus nominaaliselittäjänä ei toimi yhtä hyvin kaikkien luokkien todennäköisyyksien mallintamisessa. Vuoroasumisvaatimus nominaaliselittäjänä vaikuttaa kuitenkin huomioivan onnistuneesti vaikutuksen tasapeliluokan todennäköisyyteen, eli tältä

kannalta nominaaliselittäjän valintaa voi pitää perusteltuna. Koska tämä on välttämätöntä mallin käytännön toimivuuden kannalta, niin nominaaliselittäjä pidetään mallissa.

## 5.4 Mallien tulkitsemisesta

Määritellään mallien tulkintaa varten käsitteet

$$\text{äidin vastasuhde} = \frac{P(Y \leq j)}{1 - P(Y \leq j)} \text{ ja}$$

$$\text{isän vastasuhde} = \frac{P(Y \geq j)}{1 - P(Y \geq j)}$$

kaikilla  $j = 1, 2, 3, 4, 5$ , missä  $Y$  on oikeuden päätöstä kuvaava viisiluokkainen vastemuuttuja. Koska vastemuuttujan pienet arvot (luokat 1 ja 2) ovat äidin voittoluokkia ja suuret arvot (luokat 4 ja 5) isän voittoluokkia, niin äidin vastasuhteen kasvaminen merkitsee kertymätodennäköisyyksien  $P(Y \leq j)$  kasvamista kaikilla indekseillä  $j$ , mikä puolestaan merkitsee äidin voittoluokkien todennäköisyyksien kasvamista suhteessa isän voittoluokkiin. Vastaavasti isän vastasuhteen kasvaminen kaikilla indekseillä  $j$  merkitsee todennäköisyyksien  $P(Y \geq j)$  kasvamista kaikilla indekseillä  $j$ , mikä puolestaan merkitsee isän voittoluokkien todennäköisyyksien kasvamista suhteessa äidin voittoluokkiin.

Kun mallissa on mukana tavallinen selittäjä  $x_k$ , niin tätä vastaavan estimaatin  $\hat{\beta}_k$  etumerkki kertoo estimoidun vaikutuksen suunnasta. Jatkuvilla selittäjillä jos estimaatti on positiivinen, niin selittäjän arvojen kasvaessa kertymätodennäköisyydet  $P(Y \leq j)$  pienenevät kaikilla indekseillä  $j$ , eli isän vastasuhde kasvaa. Vastaavasti jos estimaatti on negatiivinen, niin selittäjän arvojen kasvaessa äidin vastasuhde kasvaa. Järjestyssasteikollisilla selittäjillä tulkinta on samanlainen. Indikaattoriselittäjillä estimaatin positiivisuus merkitsee isän estimoitujen voittotodennäköisyyksien kasvamista, kun indikaattorin arvo vaihtuu nolasta yhteen. Yhden yksikön muutos selittäjässä merkitsee  $\exp(\beta)$ -kertaista muutosta äidin tai isän vastasuhteessa.

## 5.5 Selittäjien tarkastelu yhden selittäjän malleilla

Skaala- ja nominaaliselittäjien määrä kannattaa pitää pienenä, jotta malli pysyy riittävän yksinkertaisena. Näin ollen mallin otettavat muut selittäjät oletetaan niin sanotuiksi tavallisiksi selittäjiksi, eli niiden yhteys kertymävastasuhteisiin oletetaan jokaisen luokan välillä yhtä suureksi. Näitä selittäjiä tarkastellaan ensin yhden varsinaisen selittäjän malleilla, joissa on kaikissa mukana edellisessä jaksossa käsitellyt yksinhuolto- ja vuoroasumisvaatimukset skaala- ja nominaaliselittäjinä. Luvussa 6 selittäjistä muodostetaan usean selittäjän malleja jaksossa 3.4 esitellyin mallinvalintamenetelmin.

Kaikissa malleissa käytetään vastemuuttujana viisiluokkaista päätösmuuttujaa, skaalaselittäjänä yksinhuoltovaatimusta ja nominaaliselittäjänä vuoroasumisvaatimusta. Selittävien muuttujien tarkemmat kuvaukset ja asteikot on dokumentoitu tarkemmin liitteessä B.

Oletettavasti sosiaaliviranomaisten suositus, edeltävä huolto- ja asumistilanne sekä vakiintunut asumistilanne selittävät oikeuden päätöstä. Myös osapuolten toisistaan tekemät syytökset saattavat vaikuttaa oikeuden päätökseen, tai ainakin osapuolet esittävät syytöksiä tarkoituksenaan parantaa omia voittomahdollisuuksiaan. Aluksi tarkastellaan jaksossa 4.10 esiteltyä neljää syytösmuuttujaa järjestysasteikollisina. Järjestysasteikollisissa muuttujissa on tieto siitä, onko syytöstä esitetty äitiä kohtaan, ei kumpaakaan tai molempia kohtaan, vai isää kohtaan. Lisäksi tarkastellaan vastaavassa muodossa olevia syytösmuuttujia, joissa on tieto siitä, onko oikeus katsonut vastaavan syytöksen toteennäytetyksi.

Vanhempien taustatiedoista on kiinnostavaa tutkia ulkomaalaistaustaisuuden yhteyttä päätökseen. Tarkasteluun otetaan myös vanhempien välinen ikäero, vaikka tässä muuttujassa on jonkin verran puuttuvia arvoja. Vanhempien tulotietoja ei oteta tarkasteluun suuren puuttuvuuden vuoksi. Hieman vastaavanlaista informaatiota voi kuitenkin tarjota tieto vanhemman työttömyydestä, jossa puuttuvuutta ei ole yhtä paljon.

On myös kiinnostavaa tutkia, eroavatko päätökset hovioikeuksien välillä. Lisäksi on mielenkiintoista tutkia, onko aiemmalla oikeuskäsittelyllä tai voimassa olevalla väliaikais määräyksellä merkitystä. Otetaan siis yhden selittäjän tarkasteluun tieto hovioikeudesta, mahdollisesta aiemmasta oikeuskäsittelystä sekä mahdollisesta voimassa olevasta väliaikais määräyksestä.

Pere, Lahti ja Sutela (2017) tutkivat sukupuolivaikutuksia törkeissä rattijuopumus tuomioissa käräjäoikeuksissa. He havaitsivat, että tuomarin, syyttäjän ja syytetyn sukupuoli yhdessä ovat yhteydessä annettuun tuomioon. Tämän perusteella on kiinnostavaa tutkia ainakin tuomarikokoonpanon sukupuolivaikutusta annettuun päätökseen. Syyttäjälle ja syytetylle ei ole vastaavaa roolia riita-asiassa, mutta esimerkiksi osapuolten avustajien sukupuolilla voi olla merkitystä, erityisesti yhteisvaikutuksessa tuomariston sukupuolien kanssa.

Sukupuolivaikutusten ohella myös tuomarien iällä voi olla merkitystä, joten valitaan alustavaan yhden selittäjän tarkasteluun myös tuomarikokoonpanon ikien keskiarvo ja varianssi.

Koska hovioikeuden päätöksiä on vuosituhanen vaihteesta aina vuoteen 2016 asti, on oletettavaa, että myös ajankohta voi tuoda oman vaikutuksensa oikeuden päätöksiin. Aikatrendin sovittamista varten käytetään keskistettyä aikamuuttujaa (liite B).

Palo-Revon (2015) tutkielmassa havaittiin, että kun lapsi on alle kouluikäinen, niin äidin voittomahdollisuudet paranevat. Otetaan siis tarkasteluun lasten ikä. Tutkitaan yhden selittäjän malleissa kaikkien lasten ikien keskiarvoa, nuorimman lapsen ikää sekä vanhimman lapsen ikää. Otetaan tarkasteluun myös nuorimman ja vanhimman lapsen sukupuoli.

Selittäjien testaamisessa käytetään Waldin testiä uskottavuusosamäärätestin sijasta jakson 3.1 lopussa kerrotusta syystä. Ennen yhden selittäjän analyyseja määritellään parametreja koskevilla testeillä tilastollinen merkitsevyys seuraavasti. Asteriskien tar-

koitus on tehostaa taulukoissa esitettyjen tulosten tulkintaa.

$p$ -arvo	Tulkinta	Merkintä
$< 0,05$	Melkein merkitsevä	*
$< 0,01$	Merkitsevä	**
$< 0,001$	Erittäin merkitsevä	***

Taulukossa 5.10 on esitetty kullekin selittäjälle testitulokset yhden selittäjän malleissa. Kaikista parhaimpia selittäjiä Waldin testin tilastollisen merkitsevyyden perusteella ovat valittajan sukupuoli, edeltävä huolto- ja asumistilanne, olosuhdeselvityksen suositus sekä syytös tapaamisten estämisestä tai vieraannuttamisesta. Tilastollisesti erittäin merkitsevillä selittäjillä nollassa nollahypoteesi, jonka mukaan vastaava regressioparametri on nolla, voidaan hylätä 0,1 prosentin merkitsevyydellä.

Asetelma, jonka perusteella aineisto on muodostunut, johtaa siihen, että valittajan sukupuoli on suoraan yhteydessä valituksenalaiseen käräjäoikeuden päätökseen. Mikäli käräjäoikeuden päätös on ollut asuminen äidin luona, niin isä on hovioikeudessa valittaja ja päin vastoin. Käytännössä valittajan sukupuoli siis kertoo käräjäoikeuden päätöksen lapsen asumisen osalta. Koska tässä aineistossa päätökset muuttuvat hovioikeudessa varsin harvoin, niin valittajan sukupuoli selittää myös hovioikeuden päätöstä erittäin hyvin, minkä vuoksi tätä vastaava regressiokertoimen estimaatti on itseisarvoltaan suuri. Usean selittäjän malleissa tämän selittäjän käyttäminen voi johtaa siihen, että selittäjä vie selitysvaimaa muilta kiinnostavilta selittäjiltä. Jos riidan lopputulosta halutaan kuitenkin mallintaa ottaen lähtökohdaksi käräjäoikeuden päätös, niin valittajan sukupuoli tai käräjäoikeuden päätös on mielenkiintoinen ja tärkeä selittäjä usean selittäjän mallissa.

Vakiintunut asumistilanne sekä edeltävä huoltotilanne ovat odotetusti hyvin oikeuden päätöstä selittäviä tekijöitä. Vakiintunutta asumistilannetta vastaavan selittäjän regressiokertoimen estimaatti on näistä itseisarvoltaan suurempi ja selittäjä myös tilastollisesti merkitsevämpi. Nämä selittäjät riippuvat toisistaan voimakkaasti ja sisältävät osittain päällekkäistä informaatiota (jakso 4.3). Näin ollen ne voivat toimia usean selittäjän malleissa huonosti yhdessä.

Selittäjäehdokkaista nousevat yhden selittäjän malleissa esiin tilastollisen merkitsevyyden perusteella myös oikeuden toteama syytös väkivaltaisuudesta, päihteistä tai mielenterveysongelmista, äidin työttömyys, nuorimman lapsen ikä, vanhempien välinen ikäero sekä tuomareiden keski-ikä.

Vanhempien välinen ikäero näyttää olevan positiivisessa yhteydessä nuoremman osapuolen vastasuhteeseen. Jos esimerkiksi isä on äitiä kymmenen vuotta vanhempi, niin äidin vastasuhte on  $\exp(10 \cdot 0,06) \approx 1,8$ -kertainen verrattuna tilanteeseen, jossa vanhemmat olisivat samanikäisiä. Ikäero jatkuvana muuttujana on tilastollisesti melkein merkitsevä selittäjä yhden selittäjän mallissa ( $p$ -arvo 0,023).

Ikäerolla voi myös olla interaktio sen kanssa, kumpi osapuolista on nuorempi. Tar-

Taulukko 5.10: Selittäjäehdokkaiden regressiokertoimien estimaatit ja Waldin testitulokset yhden selittäjän malleissa, joissa skaalaselittäjä on yksinhuoltovaatimus ja nominaaliselittäjä vuoroasumisvaatimus.

Selittäjä	Asteikko	$\hat{\beta}$	Keskivirhe	Wald	$p$ -arvo	Merkitsevyys
valittajaisa	indikaattori	-4,23	0,45	-9,46	$< 2,2 \cdot 10^{-16}$	***
oiktodsyyt4	järjestys	4,13	5,67	-0,728	0,466	
suositus	järjestys	2,49	0,35	7,04	$1,9 \cdot 10^{-12}$	***
oiktodsyyt1	järjestys	-1,61	0,50	-3,24	0,001	**
vakiintunut	järjestys	1,42	0,22	6,36	$2 \cdot 10^{-10}$	***
syyt4	järjestys	1,19	1,32	0,90	0,366	
syyt2	järjestys	1,11	0,28	3,99	$6,5 \cdot 10^{-5}$	***
huoltoennen	järjestys	0,95	0,17	5,55	$2,9 \cdot 10^{-8}$	***
aitityoton	indikaattori	0,70	0,32	2,23	0,026	*
aitiedsp	indikaattori	0,50	0,27	1,86	0,0629	
syyt3	järjestys	0,43	0,71	0,60	0,549	
syyt1	järjestys	-0,42	0,22	-1,87	0,062	
turku	indikaattori	0,40	0,29	1,38	0,168	
oiktodsyyt2	järjestys	-0,36	0,52	-0,70	0,486	
helsinki	indikaattori	-0,33	0,30	-1,12	0,265	
vam	indikaattori	0,33	0,30	1,11	0,269	
oikeudessa	indikaattori	0,33	0,37	0,89	0,372	
ulkom	järjestys	-0,33	0,46	-0,72	0,472	
isatyoton	indikaattori	-0,31	0,38	-0,84	0,402	
nuorinlapsisp	indikaattori	0,30	0,26	1,14	0,253	
ita	indikaattori	-0,26	0,33	-0,81	0,42	
vanhinlapsisp	indikaattori	0,24	0,26	0,94	0,346	
isaedsp	indikaattori	-0,14	0,28	-0,49	0,624	
vaasa	indikaattori	0,12	0,35	0,34	0,737	
naisenemmistö	indikaattori	-0,10	0,26	-0,37	0,71	
nuorlapsika	jatkuva	0,09	0,04	1,99	0,047	*
oiktodsyyt3	järjestys	-0,09	1,41	-0,06	0,951	
rovaniemi	indikaattori	0,08	0,38	0,2	0,841	
lapsika	jatkuva	0,07	0,04	1,70	0,090	
ikaero	jatkuva	-0,06	0,02	-2,28	0,0228	*
tuomika	jatkuva	0,06	0,03	2,04	0,041	*
vanhlapsika	jatkuva	0,05	0,04	1,30	0,193	
aika	jatkuva	0,04	0,04	1,10	0,273	
tuokavar	jatkuva	-0,002	0,002	-1,15	0,249	

kastellaan indikaattorimuuttujaa, joka kertoo onko äiti nuorempi kuin isä. Jos indikaattorimuuttujan ja ikäeron välistä interaktiota käytetään yksin selittäjänä, niin tätä vastaava regressiokertoimen estimaatti on noin -0,07 ja  $p$ -arvo noin 0,023, eli interaktio on selittäjänä lähes yhtä merkitsevä kuin ikäero sellaisenaan. Koska estimaatti on negatiivinen, niin se tarkoittaa, että ikäeron kasvaessa äidin vastasuhde kasvaa, kun äiti on isää nuorempi.

Myös nuorimman lapsen ikä jatkuvana selittäjänä on tilastollisesti melkein merkitsevä ( $p$ -arvo 0,047). Mitä vanhempi nuorin lapsi on, sitä paremmat näyttävät olevan isän voittomahdollisuudet. Myös kaikkien lasten ikien keskiarvon kasvaessa isän voittomahdollisuudet näyttävät kasvavan. Selittäjä ei kuitenkaan ole tilastollisesti merkitsevä ( $p$ -arvo 0,090).

Hovioikeuden tuomareiden ikien keskiarvolla on myös tilastollisesti melkein merkitse-

vä yhteys oikeuden päätökseen yhden selittäjän mallissa ( $p$ -arvo 0,041). Mitä vanhempia tuomarit ovat, sitä parempi on isän vastasuhde, ja vastaavasti mitä nuorempia tuomarit ovat, sitä parempi on äidin vastasuhde. Jos esimerkiksi tuomarien ikien keskiarvo on 60 vuotta, niin isän vastasuhde on  $\exp(6 \cdot 0,06) \approx 1,4$ -kertainen verrattuna tilanteeseen, jossa tuomarien ikien keskiarvo on 54 vuotta.

Isät näyttävät hyötyvän äitejä enemmän siitä, jos vastapuoli on työtön. Kun äidin työttömyyttä käytetään selittäjänä, niin selittäjän  $p$ -arvo on 0,026, eli selittäjä on tilastollisesti melkein merkitsevä. Äidin työttömyys nostaa isän vastasuhteen  $\exp(0,70) \approx 2,0$ -kertaiseksi verrattuna tilanteeseen, jossa äiti ei ole työtön. Isän työttömyys puolestaan nostaa äidin vastasuhteen  $\exp(0,31) \approx 1,4$ -kertaiseksi verrattuna tilanteeseen, jossa isä ei ole työtön. Isän työttömyys ei ole tilastollisesti merkitsevä selittäjä ( $p$ -arvo 0,402).

Tuomarikokoonpanon sukupuolilla ei näytä yhden selittäjän tarkastelun perusteella olevan yhteyttä oikeuden päätökseen, mutta mielenkiintoisena yhteytenä esiin nousee äidin avustajan sukupuoli. Jos äidin avustaja on mies, niin isän vastasuhde kasvaa  $\exp(0,50) \approx 1,6$ -kertaiseksi verrattuna tilanteeseen, jossa äidin avustaja on nainen. Äidin avustajan sukupuoli ei kuitenkaan ole tilastollisesti merkitsevä selittäjä ( $p$ -arvo 0,063).

Taulukossa 5.11 on erikseen tarkasteltu indikaattorimuotoisia syytösmuuttujia yhden selittäjän malleissa sen mukaan, kumpaa osapuolta kohtaan syytös on tehty. Tulosten perusteella kummankaan tekemä syytös toisen väkivaltaisuudesta, päihteidenkäytöstä tai mielenterveysongelmista (syytös 1) ei ole tilastollisesti merkitsevä selittäjä yhden selittäjän mallissa. Isän tekemä syytös äidistä ( $p$ -arvo 0,061) nostaa isän vastasuhteen  $\exp(0,50) \approx 1,6$ -kertaiseksi verrattuna tilanteeseen, jossa syytöstä ei esitetä. Äidin tekemä syytös isästä ( $p$ -arvo 0,67) nostaa äidin vastasuhteen  $\exp(0,12) \approx 1,1$ -kertaiseksi verrattuna tilanteeseen, jossa syytöstä ei esitetä. Sen sijaan jos oikeus on katsonut tämän syytöksen äidistä todeksi, niin selittäjä on tilastollisesti merkitsevä ( $p$ -arvo 0,002). Kun oikeus katsoo tämän syytöksen äidistä todeksi, niin isän vastasuhde kasvaa  $\exp(1,89) \approx 6,6$ -kertaiseksi. Oikeuden todeksi katsoma syytös isästä puolestaan ei ole tilastollisesti merkitsevä selittäjä ( $p$ -arvo 0,269). Tämä ei kuitenkaan tarkoita, että todetun syytöksen merkityksellä olisi ero äidin ja isän välillä. Eroa voidaan testata usean selittäjän mallissa.

Pelkkä syytös toista osapuolta kohtaan tapaamisten estämisestä tai vieraannuttamisesta (syytös 2) vaikuttaa sekä isän että äidin kohdalla olevan negatiivisessa yhteydessä tämän voittomahdollisuuksiin. Kun isä syyttää äitiä, niin se kasvattaa äidin vastasuhteen  $\exp(0,94) \approx 2,6$ -kertaiseksi. Kun äiti syyttää isää, niin se kasvattaa isän vastasuhteen  $\exp(1,25) \approx 3,5$ -kertaiseksi. Sekä isän tekemä syytös äidistä ( $p$ -arvo 0,006) että äidin tekemä syytös isästä ( $p$ -arvo 0,002) ovat tilastollisesti merkitseviä selittäjiä. Kuitenkin oikeuden toteamalla tapaamisten estämis- tai vieraannuttamissytyksellä näyttää olevan positiivinen yhteys vain äidin voittomahdollisuuksiin. Kun syytös on todettu isää kohtaan, niin äidin vastasuhde kasvaa  $\exp(1,73) \approx 5,6$ -kertaiseksi. Selittäjä ei kuitenkaan ole tilastollisesti merkitsevä ( $p$ -arvo 0,117). Jos puolestaan oikeus on todennut tämän



Taulukko 5.11: Syytösmuuttujat selittäjinä yhden selittäjän malleissa.

Selittäjä	$\hat{\beta}$	Keskivirhe	Wald	$p$ -arvo	Merkitsevyys
syytaitia1	0,50	0,27	1,87	0,061	
syytisaa1	-0,12	0,28	-0,43	0,67	
oiktodaitia1	1,89	0,61	3,09	0,002	**
oiktodisaa1	-0,83	0,75	-1,11	0,269	
syytaitia2	-0,94	0,35	-2,73	0,006	**
syytisaa2	1,25	0,41	3,03	0,002	**
oiktodaitia2	-0,34	0,67	-0,51	0,611	
oiktodisaa2	-1,73	1,10	-1,57	0,117	
syytaitia3	-0,97	0,91	-1,07	0,285	
syytisaa3	-0,89	1,39	-0,64	0,522	
oiktodaitia3	-0,08	1,41	-0,06	0,956	
syytaitia4	-2,29	3,30	-0,69	0,488	
syytisaa4	0,85	1,51	0,57	0,572	

syytteen äitiä kohtaan toteennäytetyksi, niin tämä ei näytä parantavan isän vastasuhdetta. Sen sijaan äidin vastasuhde näyttää tällöin jopa kasvavan, mutta selittäjä ei ole tilastollisesti merkitsevä ( $p$ -arvo 0,611).

Kaappausuhkasyytökset eivät nouse tilastollisesti merkitsevinä esiin yhden selittäjän malleissa ja tällaisia havaintoja on aineistossa hyvin vähän (jakso 4.10). Sama pätee syytöksiin seksuaalisesta hyväksikäytöstä. Pienen havaintomäärän vuoksi osaa näihin syytöksiin liittyvistä muuttujista ei voi käyttää selittäjinä, sillä tämä johtaa estimointiongelmiin.

Yhden selittäjän mallit tarjoavat vain alustavaa näkemystä hyvistä selittäjistä. Mallit kertovat selittäjän yhteydestä kummankin voittomahdollisuuksiin silloin, kun mitään muita tekijöitä ei oteta huomioon. On kiinnostavampaa muodostaa usean selittäjän malleja, jotka kertovat siitä, millainen yhteys selittäjillä on osapuolten voittomahdollisuuksiin, kun muitakin tekijöitä otetaan huomioon. Usean selittäjän malleissa voidaan myös testata sitä, onko jonkin selittäjän, esimerkiksi todetun syytöksen, yhteys oikeuden päätökseen erilainen äidin ja isän välillä. Usean selittäjän mallit ovat myös lähtökohtaisesti parempia, kun mallilla halutaan ennustaa.

## Luku 6

# Usean selittäjän mallit

Usean selittäjän malleja varten asetetaan kolme erilaista lähtökohtaa tavoitteena muodostaa kuhunkin lähtökohtaan sopiva mali. Ensimmäisessä lähtökohdassa pyritään kartoittamaan, mitkä tekijät yhdessä selittävät parhaiten asumisriitojen lopputuloksia. Toisessa lähtökohdassa pyritään ennustamaan asumisriidan lopputulos ennen käräjäoikeuden päätöstä. Kolmannessa lähtökohdassa pyritään ennustamaan lopputulos asumisriidassa, josta on käräjäoikeuden päätös, mutta josta ollaan valittamassa hovioikeuteen. Kolmannella mallilla pyritään siis myös tutkimaan sitä, onko olemassa tekijöitä, jotka selittävät päätöksen muuttumista hovioikeudessa.

Mallien pohjaksi muodostetaan aluksi niin sanottu juurimalli. Tätä muokataan edellä mainittujen lähtökohtien mukaiseksi ottaen huomioon eri lähtökohtien tuomat rajoitteet. Esimerkiksi toiseen malliin ei voi ottaa selittäjäksi valittajan sukupuolta tai käräjäoikeuden päätöstä, koska mallin on tarkoitus ennustaa oikeuden päätös jo ennen käräjäoikeuden päätöstä.

Myös mallin käyttötarkoitus määrää sen, miten juurimallia kannattaa muokata. Mikäli mallin tarkoitus ei ole uusien havaintojen ennustaminen, vaan aiempien riitojen lopputuloksiin vaikuttavien tekijöiden kartoittaminen, niin selittäjien suuri määrä ja mallin ylisopivuus käytettyyn aineistoon ei ole niin suuri ongelma kuin ennustavilla malleilla. Jos päämääränä on uusien riitojen lopputulosten ennustaminen, niin malli ei saa olla ylisopiva, vaan sen on oltava riittävän yksinkertainen, mutta kuitenkin ennustekyvyltään hyvä.

Kaikissa malleissa käytetään samaa skaala- ja nominaaliselittäjää kuin aiemminkin. Jatkossa näitä malleja kutsutaan nimillä juurimalli, malli 1, malli 2 ja malli 3.

### 6.1 Juurimalli

Koska aineistossa on paljon muuttujia ja potentiaalisia selittäjäehdokkaita, voi parhaan osajoukon valinta olla menetelmänä liian hidas juurimallin muodostamiseen. Myös eteen-

päin askeltava valinta on liian hidas samasta syystä. Suuren muuttujamäärän vuoksi tuntuu järkevältä hyödyntää määrätietoista mallinvalintamenetelmää taaksepäin askeltavalla valinnalla, jotka kuvailtiin jaksossa 3.4. Valitaan siis aluksi täyteen malliin sellaiset muuttujat, joiden ennestään tiedetään olevan tärkeitä selittäjiä sekä sellaiset muuttujat, jotka yhden selittäjän mallissa osoittautuivat hyviksi selittäjiksi.

Ennen täyttää mallia karsitaan sellaiset selittäjät, joiden Waldin testin  $p$ -arvo yhden selittäjän tarkastelussa on suurempaa kuin 0,3. Edellisessä luvussa esitettyjen estimointitulosten perusteella täydestä mallista karsiutuvat tällöin Itä-Suomen, Vaasan ja Rovaniemen hovioikeutta kuvaavat indikaattoriselittäjät. Myös tieto aiemmasta oikeuskäsittelystä, vanhimman lapsen sukupuoli, isän avustajan sukupuoli ja tuomarien sukupuolikokoonpano karsiutuvat täydestä mallista. Koska valittajan sukupuoli voi viedä selitysvoimaa muilta selittäjiltä, niin se otetaan myös pois täydestä mallista. Se huomioidaan kuitenkin kolmannessa mallissa, jossa hovioikeuden päätöstä selitetään käräjäoikeuden päätöksen pohjalta. Määrätietoisessa mallinvalinnassa myös karsittuja selittäjiä voidaan ottaa vielä tarkasteluun, kun muodostetaan varsinainen malli. Karsitut selittäjät otetaan uudelleen tarkasteluun jaksossa 6.2, kun juurimallia muokataan eri lähtökohtiin.

Osa järjestysasteikollisista selittäjistä on mahdollista hajottaa vastaaviin sukupuolitettuihin indikaattorimuuttujiin (äiti/isä). Tällaisia ovat muuttujia ovat syytösmuuttujat, olosuhdeselvityksen suositus, lapsen vakiintunut asuinpaikka sekä osapuolten ulkomaalaistaustaisuus. Nämä selittäjät otetaan kaikki täyteen malliin, vaikka yhden selittäjän malli ei olisi antanut minkäänlaista viitettä siitä, että kyseessä voisi olla hyvä selittäjä. Muuttujat pidetään ensin alkuperäisessä muodossaan järjestysasteikollisena ja ne hajotetaan vastaaviin sukupuolitettuihin indikaattorimuuttujiin, mikäli vastaava alkuperäisessä muodossa oleva muuttuja päädytään poistamaan taaksepäin askeltavassa valinnassa. Tämän tarkoitus on tuoda esiin mahdolliset eriävät sukupuolivaikutukset. Poikkeuksellisesti oikeuden toteamaa syytöstä seksuaalisesta hyväksikäytöstä ei oteta täyteen malliin, koska yhdessäkään tapauksessa oikeus ei ole katsonut toteennäytetyksi isää kohtaan esitettyä syytöstä seksuaalisesta hyväksikäytöstä ja äitiäkin kohtaan vain yhdessä tapauksessa.

Taaksepäin askeltavassa valinnassa mallista poistetaan selittäjä, jota vastaava Waldin testin  $p$ -arvo on suurin. Poistamista jatketaan, kunnes kaikkien jäljelle jäävien parametrien kohdalla voidaan hylätä viiden prosentin merkitsevyystasolla nollahypoteesi, jonka mukaan kyseinen parametri on nolla, eli tällöin vastaava  $p$ -arvo on pienempää kuin 0,05. Jos malliin jää tällöin vielä hajotettava selittäjä, joita ei ole vielä hajotettu sukupuolitettuihin muuttujiin, niin nämä muuttujat hajotetaan ja taaksepäin askeltavaa valintaa jatketaan edelleen, kunnes kaikkien selittäjien kohdalla alitetaan viiden prosentin merkitsevyystaso.

Taaksepäin askeltava valinta johtaa malliin, jossa selittäjiksi jäävät taulukossa 6.1 luetellut 12 selittäjää. Muodostuneessa mallissa osapuolen vastasuhdetta nostavat jälleen odotetusti lasten vakiintunut asuinpaikka tämän luona, osapuolen voittoa puoltava

Taulukko 6.1: Juurimallin selittäjien, skaalaselittäjän, kynnyksparametrien sekä nominaaliselittäjän estimaatit ja Waldin testitulokset. Skaalaselittäjä on yksinhuoltovaatimus, nominaaliselittäjä on vuoroasumisvaatimus ja  $n = 387$ .

Selittäjä	$\hat{\beta}_k$	Keskivirhe	Wald	$p$ -arvo	Merkitsevyys
syttisaa4	3,58	1,66	2,15	0,031	*
suosituisalle	3,27	0,65	5,06	$< 10^{-6}$	***
vakiintunut_isalle	2,73	0,61	4,48	$< 10^{-5}$	***
vakiintunut_aidille	-2,55	0,60	-4,27	$< 10^{-4}$	***
suositusaidille	-2,41	0,87	-2,77	0,006	**
oiktodisaa1	-2,35	0,95	-2,48	0,013	*
oiktodaitia1	1,93	0,72	2,67	0,008	**
aitiedsp	1,20	0,43	2,81	0,005	**
aitityoton	1,15	0,51	2,25	0,024	*
helsinki	1,07	0,49	2,21	0,027	*
aika	0,14	0,07	2,10	0,036	*
ikaero	-0,12	0,03	-3,47	0,0005	***
Skaalaselittäjä	$\hat{\zeta}$	Keskivirhe	Wald	$p$ -arvo	Merkitsevyys
vaat_yh	2,58	0,20	12,7	$< 10^{-15}$	***
Kynnys	$\hat{\theta}_j$	Keskivirhe	$\hat{\beta}_j$	Keskivirhe	
1   2	-6,67	1,52	2,06	3,58	
2   3	2,21	0,76	-1,79	0,69	
3   4	2,53	0,77	2,00	0,78	
4   5	11,80	1,73	6,70	12,42	

sosiaaliviranomaisten suositus sekä toista osapuolta kohtaan todeksi näytetty syytös väkivaltaisuudesta, päihteistä tai mielenterveysongelmista. Malliin tilastollisesti melkein merkitsevänä jäävä aikatrendi vastaa käsitystä, jonka mukaan ajan kuluessa isien voittomahdollisuudet huolto- ja asumisriidoissa ovat kasvaneet. Lisäksi osapuolten välisen ikäeron kasvaessa nuoremman osapuolen vastasuhde kasvaa kuten yhden selittäjän mallissakin.

Erittäimen mielenkiintoinen huomio on se, että usean selittäjän malliin jää selittäjäksi äidin tekemä syytös isää kohtaan lasten seksuaalisesta hyväksikäytöstä, joka yhden selittäjän mallissa ei noussut esiin. Yhteys äidin vastasuhteeseen on erittäin negatiivinen. Äidin tekemä syytös isää kohtaan nostaa isän vastasuhteen peräti  $\exp(3,58) \approx 35,9$ -kertaiseksi verrattuna tilanteeseen, jossa syytöstä ei esitetä. Selittäjän keskivirhe on suuri, mutta selittäjä on usean selittäjän mallissa tilastollisesti melkein merkitsevä ( $p$ -arvo 0,031).

Yhden selittäjän malleissa esiin nousseet äidin työttömyys ja äidin miesavustaja näkyvät samanlaisena isän vastasuhdetta parantavina selittäjinä myös usean selittäjän mallissa. Mielenkiintoinen huomio on, että usean selittäjän mallissa Helsingin hovioikeus näyttää tilastollisesti melkein merkitseväenä selittäjänä parantavan isän vastasuhdetta.

Mallin havaintomäärä on 387, mikä johtuu puuttuvista arvoista vanhempien ikätiedoissa sekä avustajien tiedoissa. Skaalaselittäjän Waldin testin  $p$ -arvo on hyvin lähellä nollaa, mikä tarkoittaa, että skaalaselittäjä on myös tässä mallissa tilastollisesti erittäin merkitsevä (taulukko 6.1). Mallin yhteensopivuusindeksi on noin 0,802, mitä voi esimerkiksi Harrellin (2015, 257) mukaan pitää hyvänä. Karkeasti malli osaa siis laittaa yli 80 prosenttia aineiston vastemuuttujan havaintopareista oikeaan suuruusjärjestykseen järjestysasteikolla. Mallin yhteiskorrelaatiokerroin on noin 0,534.

Tutkitaan juurimallin ennustekykä myös ennustamalla oikeuden päätös jokaisessa 387 havainnossa, joilla malli sovitettiin. Ennusteena kullekin havainnolle toimii se luokka, jonka estimoitu todennäköisyys on suurin. Ennusteen laskeminen on kuvailtu tarkemmin jaksossa 6.3. Ennusteet on taulukoitu ristiin oikeuden päätöksen kanssa taulukossa 6.2

Taulukko 6.2: Juurimallin ennusteet ristiintaulukoituna oikeuden päätöksen kanssa. 1=äidin yksinhuolto, 2=yhteishuolto, äidin lähihuoltajuus, 3=lasten jako tai vuoroasuminen, 4=yhteishuolto, isän lähihuoltajuus, 5=isän yksinhuolto. Vihreällä ovat täysin oikein ennustetut riidat. Vihreällä ja keltaisella ovat oikein ennustetut voittajat asumisriidassa.

		Päätös					Yht.	% lopputulos oikein	% asumisriidan voittaja oikein
		1	2	3	4	5			
Ennuste	1	70	18	2	8	15	113	61,9	77,9
	2	0	75	2	12	0	89	84,3	84,3
	3	0	3	10	2	0	15	66,7	66,7
	4	0	10	4	77	0	91	84,6	84,6
	5	9	5	1	13	51	79	64,6	81,0
Yht.		79	111	19	112	66	387		

Kaikista parhaiten malli näyttää toimivan luokkien 2 ja 4 ennustamisessa. Malli ennustaa yli 84 prosenttia havainnoista oikein silloin, kun mallin ennusteena on yhteishuolto ja jommankumman vanhemman lähihuoltajuus. Oikeuden päätös ei esimerkiksi koskaan ole kummankaan yksinhuolto silloin, kun malli ennustaa yhteishuoltoa. Vaikuttaa siis siltä, että yksinhuoltovaatimus skaalaselittäjänä näyttää toimivan myös ennustamisen kannalta toivotulla tavalla.

Tasapelejä malli ei ennusta yhtä hyvin. Vain kaksi kolmasosaa tasapeliennusteista on todellisuudessa myös tasapelejä. Aineistossa on kuitenkin tasapelejä vain vähän, mikä osittain voi selittää myös niiden heikompa ennustamista.

Kun malli ennustaa äidin yksinhuoltoa, niin vain noin 62 prosenttia ennusteista on täysin oikein. Isän yksinhuoltoennusteista noin 65 prosenttia on täysin oikein. Jos tarkastellaan vain asumisriidan voittajaa, niin äidin yksinhuoltoennusteista noin 78 prosenttia on äidin voittoja asumisriidassa (luokka 1 tai 2). Isän yksinhuoltoennusteista noin 81 prosenttia on isän voittoja asumisriidassa (luokka 4 tai 5). Asumisriidan voittajan ennustamisessa malli toimii siis kohtalaisen hyvin. Sekä yhteisesti että yksinhuollon tapauksissa malli näyttää ennustavan isän voiton hieman paremmin kuin äidin voiton. Yksinhuoltoluokkien ennusteista kuitenkin nähdään, että malli hieman yliennustaa yksinhuoltoa. Tämä voi johtua siitä, että skaalavaikutus tulee mukaan aina, kun jompikumpi vanhemmista vaatii yksinhuoltoa. Ongelman voisi mahdollisesti korjata huomioimalla yksinhuoltovaatimuksen mallissa jollakin toisella tavalla, esimerkiksi käyttämällä kummankin osapuolen yksinhuoltovaatimukselle omaa skaalaselittäjää. Tätä kokeiltiin, mutta useampi skaalaselittäjä aiheutti toistuvasti ongelmia mallin parametrien estimoinnissa erityisesti silloin, kun muiden selittäjien määrä oli suuri.

Joissain tapauksissa päätös saattaa olla täysin päinvastainen ennusteen kanssa, mitä voi pitää pahana ennustevirheenä. Malli saattaa esimerkiksi ennustaa todennäköisimpänä äidin yksinhuoltoa, mutta päätös onkin isän yksinhuolto. Kun näitä katsotaan tarkemmin, niin kaikissa näissä tapauksissa mallin ennustama todennäköisyys isän yksinhuollolle on kuitenkin vähintään 20 prosenttia ja lähes aina tämä on mallin mukaan toiseksi todennäköisin luokka. Vastaavasti jos malli ennustaa isän yksinhuoltoa ja päätös onkin äidin yksinhuolto, niin äidin yksinhuolto on kuitenkin aina toiseksi todennäköisin luokka ja erot näiden kahden luokan välisissä todennäköisyyksissä ovat varsin pieniä. Kyse ei siis varsinaisesti ole pahasta ennustevirheestä, vaan siitä, että malli pitää kumpaakin ääriluokkaa melko todennäköisenä. Käytännössäkin tilanne, jossa keskimmäiset päätösluokat on suljettu pois, on varsin tavallinen. Vanhemmat saattavat esimerkiksi olla lähtökohtaisesti erittäin tasavertaisissa asemissa, mutta vanhempien riitaisten välien vuoksi yhteishuollolle ei ole edellytyksiä.

Ennustekykynsä puolesta malli vaikuttaa kokonaisuudessaan kohtalaisen hyvältä erityisesti asumisriidan voittajan osalta. On kuitenkin huomattava, että ennusteeksi katsottiin todennäköisin luokka eikä luokkien välisten estimoitujen todennäköisyyksien välisiä eroja tarkasteltu tarkemmin. On lisäksi mahdollista, että suurehkon selittäjämäärän vuoksi juurimalli on liian riippuvainen juuri kyseisestä aineistosta eikä se välttämättä toimi uusien havaintojen ennustamisessa yhtä hyvin.

Koska puuttuvuuden vuoksi juurimalli sovitettiin vain 387 havainnolla, niin mallia voi mahdollisesti parantaa poistamalla selittäjiä, joissa on puuttuvuutta, eli vanhempien ikäero, äidin avustajan sukupuoli ja tieto äidin työttömyydestä. Lisäksi mallia voidaan yrittää parantaa ottamalla siihen mukaan sellaisia selittäjiä, joita ei otettu täyteen malliin.

## 6.2 Eri lähtökohtiin sovitetut mallit

### 6.2.1 Ensimmäinen malli

Ensimmäisen mallin tavoitteena on etsiä tekijät, jotka parhaiten selittävät yhdessä tähänastisten asumisriitojen lopputuloksia. Malli muodostetaan juurimallin pohjalta.

Tarkastellaan juurimallia ja tutkitaan ensin selittäjien poistamisen vaikutusta Akaiken informaatiokriteeriin (AIC). Yhdenkään selittäjän poistaminen ei laske mallin AIC:tä, eli informaatiokriteerin perusteella mallista ei tulisi poistaa selittäjiä.

Koska mallissa on jo hyvin paljon selittäjiä, niitä ei kannata lisätä liikaa. Määrätietoisessa mallinvalinnassa tulisi kuitenkin kokeilla malliin vielä sellaisia selittäjiä, joita ei otettu mukaan täyteen malliin. Tutkitaan täydestä mallista pois jätettyjen selittäjien lisäämisen vaikutusta.

Täydestä mallista jätettiin pois selittäjät, joissa on tieto siitä onko päätös Itä-Suomen, Vaasan tai Rovaniemen hovioikeudesta. Yhdenkään näistä ottaminen mukaan malliin ei tee lisättävästä selittäjästä tilastollisesti merkitsevää. Aiempi oikeuskäsittely, voimassa oleva väliaikais määräys tai ratkaisijoiden sukupuolikokoonpano, jotka niin ikään jäivät täyden mallin ulkopuolelle, eivät myöskään ole tilastollisesti merkitseviä, jos niitä yritetään lisätä malliin.

Kun malliin lisätään selittäjäksi vanhimman lapsen sukupuoli, niin Waldin testin perusteella selittäjä on tilastollisesti melkein merkitsevä ( $p$ -arvo 0,034). Selittäjän lisääminen aiheuttaa kuitenkin sen, että äidin työttömyys ja aikatrendi eivät ole mallissa tilastollisen merkitsevyyden perusteella enää yhtä hyviä selittäjiä kuin aiemmin ( $p$ -arvot 0,065 ja 0,058). Mallin AIC kuitenkin pienenee.

Kun malliin lisätään selittäjäksi vanhimman lapsen sukupuolen ohella myös vanhimman lapsen ikä, niin Waldin testin perusteella vanhimman lapsen ikä ei enää ole tilastollisesti merkitsevä selittäjä ( $p$ -arvo 0,379). Kokeillaan vaihtaa lapsen ikätarkastelu koskemaan sitä, onko lapsi alle kouluikäinen (alle 7-vuotias). Vaihtoehtoina on tarkastella joko nuorinta tai vanhinta lasta. Näistä vanhimman lapsen kouluikäisyys toimii Waldin testin perusteella mallissa paremmin. Lisätty selittäjä ei ole tilastollisesti merkitsevä ( $p$ -arvo 0,105), mutta AIC pienenee edelleen.

Vanhimman lapsen sukupuoli ja kouluikäisyys päädytään lisäämään malliin, koska kokonaisuutena näiden voi katsoa parantaneen mallia. Millään muulla selittäjällä tai selittäjäyhdistelmällä malli ei parane parametrien tilastollisen merkitsevyyden tai informaatiokriteerin perusteella. Uuden mallin yhteensopivuusindeksi on noin 0,802, eli lähes sama kuin juurimallilla. Mallin yhteiskorrelaatiokerroin noin 0,542. Parametrien estimaatit ja näiden 95 prosentin uskottavuusosamäärään perustuvat luottamusvälit on esitetty taulukossa 6.3. Estimointitulokset ovat kokonaisuudessaan liitteessä C.2.

Mallin perusteella äidin vastasuhde näyttää kasvavan, kun vanhin lapsi on alle kouluikäinen. Isän vastasuhde puolestaan kasvaa silloin, kun vanhin lapsi on poika.

On kiinnostavaa tutkia mahdollista interaktiota vanhimman lapsen sukupuolen ja

Taulukko 6.3: Parametrien estimaatit ja uskottavuusosamäärään perustuvat luottamusvälit ensimmäisessä mallissa,  $n = 384$ .

Selittäjä	$\hat{\beta}_k$	95 %	
		luottamusväli	
syytisa4	4,29	0,98	8,46
suositusisalle	3,29	2,10	4,67
vakiintunut_aidille	-2,90	-4,25	-1,72
vakiintunut_isalle	2,65	1,51	3,97
suositusaidille	-2,54	-4,53	-0,96
oiktodisaal	-2,46	-4,49	-0,53
oiktodaitial	2,38	0,86	4,06
aitiedsp	1,15	0,30	2,04
helsinki	1,09	0,12	2,13
aitityoton	0,95	-0,08	2,02
vanhinlapsipoika	0,84	-0,01	1,72
vanhlapsa7	-0,70	-1,57	0,14
aika	0,13	0,00	0,27
ikaero	-0,12	-0,19	-0,05
Skaalaselittäjä	$\hat{\zeta}$		
vaat_yh	2,59	2,20	3,01

vanhimman lapsen kouluikäisyyden välillä. Kun malliin kokeillaan lisätä selittäjäksi näiden välinen interaktio, niin tämä ei kuitenkaan paranna parametrien tilastollista merkitsevyyttä, ja interaktiota vastaava  $p$ -arvo on noin 0,813.

Lasten vakiintunut asuinpaikka on kummankin osapuolen kohdalla tilastollisesti erittäin merkitsevässä yhteydessä osapuolen vastasuhteeseen (liite C.2). Äidin vastasuhte kasvaa  $\exp(2,90) \approx 18,2$ -kertaiseksi silloin, kun lasten vakiintunut asuinpaikka on tämän luona. Isän vastasuhte puolestaan kasvaa  $\exp(2,65) \approx 14,2$ -kertaiseksi vastaavassa tilanteessa. Jos mallissa testataan uskottavuusosamäärätestillä nollahypoteesia, jonka mukaan vakiintuneen asuinpaikan merkitys on sama äidin ja isän kohdalla, niin  $p$ -arvoksi saadaan 0,792 (liite C.2). Nollahypoteesia ei siis ole syytä hylätä, eli tilastollisesti merkitsevää eroa vakiintuneen asuinpaikan merkityksessä äidin ja isän välillä ei tässä mallissa ole.

Osapuolta kohtaan esitetty sosiaaliviranomaisten suositus on sekä isän että äidin kohdalla tilastollisesti merkitsevässä yhteydessä tämän vastasuhteeseen (liite C.2). Isällä vastasuhteen kasvu on  $\exp(3,29) \approx 26,8$ -kertainen, kun äidillä se on  $\exp(2,54) \approx 12,7$ -kertainen verrattuna tilanteeseen, jossa suositusta ei esitetä tätä osapuolta kohtaan. Kun testataan uskottavuusosamäärätestillä nollahypoteesia, jonka mukaan olosuhdeselvityk-



Taulukko 6.4: Kynnysparametrien ja nominaalivaikutusten estimaatit ensimmäisessä mallissa.

Kynnys	$\hat{\theta}_j$	Keskivirhe	$\hat{\beta}_j$	Keskivirhe
1 2	-6,20	1,72	2,30	3,59
2 3	3,00	1,09	-1,71	0,72
3 4	3,33	1,09	2,53	0,87
4 5	12,70	1,94	7,00	12,40

sen suosituksen merkitys on sama äidin ja isän kohdalla, niin  $p$ -arvoksi saadaan 0,505 (liite C.2). Nollahypoteesia ei siis ole syytä hylätä, eli tilastollisesti merkitsevää eroa olosuhdeselvityksen suosituksen merkityksessä äidin ja isän välillä ei ole.

Oikeuden toteama väkivalta-, päihteidenkäyttö- tai mielenterveys-syytös (syytös 1) toisesta osapuolesta on myös kummankin osapuolen kohdalla positiivisessa yhteydessä tämän vastasuhteeseen. Kun testataan uskottavuusosamäärätestillä nollahypoteesia, jonka mukaan todeksi katsotun syytöksen merkitys on sama äidin ja isän kohdalla, niin  $p$ -arvoksi saadaan 0,949 (liite C.2). Nollahypoteesia ei siis ole syytä hylätä, eli tilastollisesti merkitsevää eroa todeksi katsotun syytöksen merkityksessä äidin ja isän välillä ei ole.

Isää kohtaan tehty syytös lasten seksuaalisesta hyväksikäytöstä (syytös 4) näyttää juurimalliin verrattuna olevan entistä suuremmassa yhteydessä isän voittomahdollisuuksiin. Selittäjä on tilastollisesti melkein merkitsevä (liite C.2).

Muilla selittäjillä vaikutukset ovat suurin piirtein samansuuruisia kuin juurimallissa. Malli ei ole merkittävästi juurimallia parempi, mutta se ottaa huomioon myös lapsen iän, joka Palo-Revon (2015) tutkimuksessa todettiin keskeiseksi tekijäksi selittämään äidin voittoa. Mukana mallissa on myös vanhimman lapsen sukupuoli, joka yhdessä vanhimman lapsen kouluikäisyyden kanssa paransi juurimallia paremmin kuin kumpikaan näistä yksin.

Mallissa on varsin paljon selittäjiä, mikä voi merkitä ylisopivuutta. Tämän mallin tarkoitus ei kuitenkaan ole uusien havaintojen ennustaminen, vaan keskeisimpien osapuolen voittoa selittävien tekijöiden löytäminen tähänastisissa huolto- ja asumisriidoissa. Taulukoissa 6.3 ja 6.4 sekä liitteessä C.2 esitetty malli on ensimmäisen lähtökohdan mukainen malli.

## 6.2.2 Toinen malli

Toisen mallin tavoite on oikeuden päätöksen ennustaminen huolto- ja asumisriidassa. Tällöin mallin tulisi olla ennustekyvyltään hyvä eikä se saa riippua liikaa vain tietystä aineistosta. Tämän vuoksi pyritään aiempaa mallia pienempään selittäjämäärään, koska suuri selittäjämäärä merkitsee usein ylisopivuutta.

Koska juurimallissa on paljon selittäjiä, lähdetään liikkeelle yksinkertaisesta mallista

ja sovelletaan eteenpäin askeltavaa valintaa, eli lisätään selittäjä, joka on Waldin testin perusteella merkitsevin. Rajataan tarkastelu koskemaan vain asetelmaan sopivia selittäjiä. Koska tavoitteena on luoda malli, joka ennustaa riidan lopputuloksen ennen koko oikeudenkäyntiprosessia, selittäjiksi ei voida ottaa oikeuden toteamia syytöksiä, tietoja ratkaisijoista tai avustajista eikä sosiaaliviranomaisten suosituksia. Sen sijaan huomioon voidaan ottaa edeltävä huoltotilanne, vakiintunut asumistilanne, lasten ja vanhempien ikätiedot sekä vanhempien työttömyystiedot, jotka lähtökohtaisesti tiedetään jo ennen kuin riita on viety oikeuteen. Oletetaan myös osapuolten toisistaan tekemät syytökset tiedetyiksi jo ennen ensimmäistä oikeudenkäyntiä. Hyödynnetään kuitenkin juurimallia ja yhden selittäjän malleja siten, että valitaan lähtötilanteeseen vain sellaiset selittäjät, jotka ovat antaneet jonkinlaista viitettä siitä, että kyseessä voisi olla hyvä selittäjä.

Rajaamisen jälkeen selittäjäehdokkaiksi jäävät edeltävä huoltotilanne, vakiintunut asumistilanne, äitiä kohtaan tehty syytös väkivaltaisuudesta, päihteenkäytöstä tai mielenterveysongelmista, kummankin tekemät syytökset tapaamisten estämisestä tai vieraannuttamisesta, isää kohtaan tehty syytös lasten seksuaalisesta hyväksikäytöstä, vanhempien välinen ikäero, vanhimman lapsen kouluikäisyys, vanhimman lapsen sukupuoli, äidin työttömyys, Helsingin hovioikeus ja aikatrendi. Selittäjien lisääminen lopetetaan, kunnes yhdenkään uuden selittäjän Waldin testin  $p$ -arvo ei ole alle 0,05. Samalla tarkkaillaan selittäjien lisäämisen vaikutusta mallin yhteensopivuusindeksiin. Yhteensopivuusindeksi on käytetyistä diagnostiikkatyökaluista paras mittaamaan mallin ennustekykä, joka tällä mallilla on tarkoitus saada mahdollisimman hyväksi.

Kaikista yksinkertaisimman mallin lähtökohtana on vain edeltävä huolto- ja asumistilanne, joka yhden selittäjän tarkastelussa osoittautui tilastollisesti erittäin merkitseväksi selittäjäksi (taulukko 5.10). Kun selittäjänä on vain tämä muuttuja, niin mallin yhteensopivuusindeksi on noin 0,67. Karkeasti malli osaa siis laittaa kaksi kolmasosaa vastemuuttujan havaintopareista oikeaan suuruusjärjestykseen järjestyksasteikolla, kun tiedetään vain edeltävä huoltotilanne sekä skaala- ja nominaaliselittäjien muodossa huomioitavat osapuolten vaatimukset.

Eteenpäin askeltava valinta johtaa seitsemän varsinaisen selittäjän malliin, jossa selittäjinä ovat edeltävä huolto- ja asumistilanne, vakiintunut asumistilanne isän luona, vanhempien välinen ikäero, äidin työttömyys, äitiä kohtaan tehty väkivalta-, päihde- tai mielenterveys-syytös, äitiä kohtaan tehty tapaamisten estämis- tai vieraannuttamis-syytös sekä isää kohtaan tehty syytös lasten seksuaalisesta hyväksikäytöstä. Selittäjät lisättiin malliin tässä järjestyksessä tilastollisen merkitsevyyden perusteella. Mallin AIC on noin 830,18, yhteensopivuusindeksi noin 0,731 ja yhteiskorrelaatiokerroin noin 0,390.

Kun malliin lisätään selittäjäksi tieto siitä, onko vanhin lapsi alle kouluikäinen, niin tämä ei ole tilastollisesti merkitsevä ( $p$ -arvo 0,065). Selittäjän lisääminen parantaa kuitenkin mallin AIC:tä (820,66). Yhteensopivuusindeksiä se nostaa vain hieman (0,732) ja yhteiskorrelaatiokerroin nousee myös hieman (0,399). Koska selittäjän voi katsoa parantavan mallia hieman, niin se päädytään lisäämään malliin.

Kun tähän lisätään vielä selittäjäksi vakiintunut asuinpaikka äidin luona, niin lisätty selittäjä on tilastollisesti merkitsevä ( $p$ -arvo 0,007), mutta tämä aiheuttaa sen, että edeltävä huoltotilanne ei ole enää tilastollisesti merkitsevä selittäjä ( $p$ -arvo 0,071). Vakiintunut asuinpaikka äidin luona parantaa entisestään mallin AIC:tä (815,01), mutta erityisesti yhteensopivuusindeksiä, joka nousee selittäjän lisäämisen myötä 0,759:ään. Ennustekyvyn kannalta kyseessä saattaa siis olla tärkeä selittäjä, joten myös tämä selittäjä otetaan vielä mukaan malliin. Jos selittäjistä poistetaan edeltävä huoltotilanne, joka ei ole enää tilastollisesti merkitsevä, niin mallin AIC tai yhteensopivuusindeksi ei kuitenkaan parane. Myös edeltävä huoltotilanne pidetään siis mallissa.

Muiden selittäjien lisääminen ei paranna mallia merkittävästi, joten malliksi valitaan taulukon 6.5 mukainen malli. Mallin estimointitulokset ovat kokonaisuudessaan liitteessä C.3. Vaikka mallin yhteensopivuusindeksi on pienempi kuin aiemmillä malleilla, niin yhteensopivuusindeksi on kuitenkin kohtalainen. Malli osaa laittaa hieman yli kolme neljäsosaa aineiston vastemuuttujan havaintopareista oikeaan suuruusjärjestykseen. Yhteensopivuusindeksi mittaa ennustekykä vain siinä aineistossa, jolla malli sovitettiin, joten siitä ei tule tehdä liian vahvoja johtopäätöksiä uusien havaintojen ennustamisessa. Koska tässä mallissa ei ole enää selittäjänä äidin avustajan sukupuoli, jossa oli paljon puuttuvia arvoja, niin malli on estimoitu suuremmalla havaintomäärällä kuin aiemmat mallit.

Taulukko 6.5: Selittäjien ja skaalaselittäjän estimaatit ja uskottavuusosamäärään perustuvat luottamusvälit toisessa mallissa,  $n=414$ .

Selittäjä	$\hat{\beta}_k$	95 %	
		luottamusväli	
syytisaa4	4,49	1,32	8,38
vakiintunut_isalle	1,63	0,62	2,73
vakiintunut_aidille	-1,45	-2,53	-0,42
syytaitia2	-1,29	-2,35	-0,31
syytaitia1	1,11	0,40	1,86
aitityoton	0,91	0,12	1,71
vanhlapsa7	-0,75	-1,46	-0,07
huoltoennen	0,55	-0,05	1,15
ikaero	-0,11	-0,17	-0,05
Skaalaselittäjä	$\hat{\zeta}$		
vaat_yh	2,39	2,02	2,84

Edeltävä huolto- ja asumistilanne selittäjänä heikentää vakiintuneen asuinpaikan merkitystä selittäjänä, kun tuloksia verrataan aiempiin malleihin, joissa edeltävä huoltotilanne ei ollut selittäjänä. Nämä eivät samassa mallissa kuitenkaan syö kokonaan tois-

Taulukko 6.6: Kynnysparametrien ja nominaalivaikutusten estimaatit toisessa mallissa.

Kynnys	$\hat{\theta}_j$	Keskivirhe	$\hat{\beta}_j$	Keskivirhe
1 2	-5,41	1,45	0,68	2,17
2 3	1,35	0,99	-1,64	0,65
3 4	1,56	0,99	1,57	0,64
4 5	8,14	1,47	6,48	10,21

tensa selitysvoimaa, vaikka ne sisältävät paljon päällekkäistä informaatiota. Koska edeltävä huoltotilanne on samalla järjestysasteikolla kuin vastemuuttujakin, niin estimaatin positiivinen arvo merkitsee isän vastasuhteen kasvamista, kun selittäjä saa suurempia arvoja järjestysasteikolla (edeltävä huolto tai asuminen isällä). Vastaavasti äidin vastasuhte kasvaa, kun selittäjä saa pienempiä arvoja järjestysasteikolla (edeltävä huolto tai asuminen äidillä).

Ikäeron, äidin työttömyyden, vanhimman lapsen kouluikäisyyden ja isää kohtaa tehdyn syytöksen lasten seksuaalisesta hyväksikäytöstä estimoitu vaikutus on suurin piirtein saman suuruista kuin aiemmassa mallissa. Jos isä syyttää äitiä väkivaltaisuudesta, päihteidenkäytöstä tai mielenterveysongelmista, niin tämä näyttää nostavan isän vastasuhteen  $\exp(1,11)$ -kertaiseksi eli noin kolminkertaiseksi verrattuna tilanteeseen, jossa syytöstä ei tehdä. Kuitenkin jos isä syyttää äitiä tapaamisten estämisestä tai vieraannuttamisesta, niin tämä nostaakin äidin vastasuhteen  $\exp(1,29) \approx 3,6$ -kertaiseksi verrattuna tilanteeseen, jossa syytöstä ei tehdä.

### 6.2.3 Kolmas malli

Kolmannen mallin muodostamisessa lähtökohtana on hovioikeuden päätöksen ennustaminen käräjäoikeuden päätöksen ja muiden taustatietojen pohjalta. Tähän malliin kannattaa siis ottaa selittäjäksi valittajan sukupuoli, joka yhden selittäjän mallissa osoittautui erittäin tärkeäksi selittäjäksi. Syy selittäjän tärkeydelle on kuitenkin se, että valittajan sukupuoli sisältää paljon informaatiota käräjäoikeuden päätöksestä, johon hovioikeuden päätös pohjautuu. Asetelma, jonka perusteella aineisto on muodostunut, aiheuttaa itse asiassa sen, että tieto valittajan sukupuolesta tuo myös suoraan tiedon käräjäoikeuden päätöksestä asumisriidan osalta. Koska käräjäoikeuden päätös huomioi varmasti jo monia merkityksellisiä seikkoja, niin valittajan sukupuoli oletettavasti vie selitysvoimaa muilta selittäjiltä.

Kun selittäjänä on pelkästään valittajan sukupuoli, niin mallin yhteensovitusindeksi on noin 0,843, eli parempi kuin yhdelläkään aiemmalla mallilla. Jos siis tiedetään kumpi osapuoli on valittaja, niin malli osaa laittaa yli viisi kuudesosaa aineiston vastemuuttujan havaintopareista oikeaan suuruusjärjestykseen. Näin hyvän tuloksen selittää kuitenkin pitkälti se, että aineiston riidoista noin 82 prosenttia on sellaisia, joissa päätös ei muutu hovioikeudessa, jolloin käräjäoikeuden päätös ja siten valittajan sukupuoli se-

littävät jo itsessään hyvin suuren osan hovioikeuden päätöksistä. Kolmannessa mallissa pyritäänkin löytämään sellaiset selittäjät, jotka parhaiten selittävät päätöksen muuttumista hovioikeudessa.

Koska lähtökohtana on ennustaa oikeudenkäynnin lopputulos ennen valittamista hovioikeuteen, niin selittäjiksi ei voida ottaa hovioikeuden ratkaisijoiden tietoja. Sen sijaan olosuhdeselvityksen tilaa usein jo käräjäoikeus, joskin myös hovioikeus voi tilata sen siinä tapauksessa, että selvitystä ei ole tehty käräjäoikeusvaiheessa ja se katsotaan tarpeelliseksi. Näin ollen olosuhdeselvityksen suositus on useimmiten tiedossa jo käräjäoikeuskäsittelyn jälkeen. Myös osapuolten avustajien tiedot ovat yleensä tiedossa käräjäoikeuskäsittelyn jälkeen, joskin avustajaa voi myös vaihtaa. Käräjäoikeus voi todeta osapuolten esittämiä syytöksiä, mutta hovioikeus voi katsoa asian eri tavalla kuin käräjäoikeus. Näin ollen oikeuden toteamia syytöksiä ei voida ottaa kolmannen mallin muodostamiseen, sillä hovioikeuden näkemys syytöksistä tiedetään vasta hovioikeuden päätöksen jälkeen.

Koska kolmanteen malliin otetaan selittäjäehdokkaaksi valittajan sukupuoli, joka sisältää käräjäoikeuden päätöksen asumisesta ja jota ei käytetty juurimallissa, niin mallin muodostamisessa ei käytetä lähtökohtana juurimallia. On mahdollista, että valittajan sukupuoli toimii selittäjänä sellaisten selittäjien kanssa, jotka eivät päässeet alkuperäiseen malliin. Tämän vuoksi kolmannen mallin muodostaminen aloitetaan jälleen täydestä mallista. Malli muodostetaan taaksepäin askeltavalla valinnalla, kunnes kaikki selittäjät ovat tilastollisesti merkitseviä.

Täydestä mallista karsitaan kuitenkin heti sellaiset selittäjät, jotka eivät sovi lähtökohtaan ja jotka yhden selittäjän malleissa eivät antaneet minkäänlaista viitettä siitä, että kyseessä voisi olla hyvä selittäjä. Syytösmuuttujia, vakiintunutta asumistilannetta, olosuhdeselvityksen suositusta ja osapuolten ulkomaalaistaustaisuutta käytetään ensin järjestysasteikollisina ja ne hajotetaan samalla periaatteella kuin juurimallin muodostamisessa.

Taaksepäin askeltava valinta johtaa aluksi viiden selittäjän malliin, jossa olosuhdeselvityksen suositusta ei ole vielä hajotettu vastaaviin sukupuolitettuihin indikaattorimuuttujiin. Mallin muut selittäjät ovat valittajan sukupuoli (joka sisältää myös käräjäoikeuden päätöksen asumisriidassa), edeltävä huoltotilanne, isää kohtaan tehty syytös väkivaltaisuudesta, päihteenkäytöstä tai mielenterveysongelmista sekä vanhempien välinen ikäero. Mallin AIC noin 738,06 ja yhteensopivuusindeksi noin 0,833. Jos tilastollisesti erittäin merkitsevä suositusselittäjä ( $p$ -arvo 0,000003) hajotetaan, niin mallin AIC ja yhteensopivuusindeksi heikkenevät hieman eivätkä äitiä ( $p$ -arvo 0,038) ja isää ( $p$ -arvo 0,00004) vastaavat sukupuolitettut suositusselittäjät ole yhtä merkitseviä kuin suositusselittäjä ilman hajottamista. Yhteiskorrelaatiokerroin puolestaan kasvaa, aivan kuten selittäjien määrän kasvaessa kuuluukin, mutta kasvu on merkityksettömän pieni. Selittäjän hajottamisen voi siis katsoa heikentävän mallia, joten olosuhdeselvityksen suositus päädytään pitämään mallissa järjestysasteikollisena.

Yhdenkään selittäjän poistaminen ei paranna mallia selittäjien tilastollisen merkitsevyyden perusteella. Paras yhteensopivuusindeksi saadaan kuitenkin mallilla, josta olosuhdeselvityksen suositus poistetaan kokonaan. Tällä mallilla yhteensopivuusindeksi on noin 0,851, eli toistaiseksi kaikista malleista korkein. Mallin yhteiskorrelaatiokerroin on noin 0,571, eli myös kaikista malleista korkein. Mallin AIC kuitenkin heikkenee hieman. Selittäjän poistaminen aiheuttaa sen, että äidin tekemä syytös isästä ei ole enää tilastollisesti merkitsevä selittäjä ( $p$ -arvo 0,087). Yhdenkään sellaisen selittäjän, jota ei otettu täyteen malliin, lisääminen ei merkittävästi paranna mallia selittäjien tilastollisen merkitsevyyden tai yhteensopivuusindeksin perusteella.

Jos ikäeron kohdalla testataan interaktiota, jossa huomioidaan se onko äiti nuorempi, niin interaktiota vastaava  $p$ -arvo on noin 0,044, eli hieman suurempi kuin ikäerolla sellaisenaan ja lisäksi tämä heikentää myös isää kohtaa tehdyn syytöksen tilastollista merkitsevyyttä ( $p$ -arvo 0,100). Interaktio ei myöskään paranna mallin yhteensopivuusindeksiä.

Koska tavoitteena on suhteellisen yksinkertainen malli, joka on ennustekyvyltään hyvä, niin kolmanneksi malliksi valitaan taulukon 6.7 mukainen malli, jossa ei ole olosuhdeselvityksen suositusta mukana ja ikäero pidetään ilman interaktiota. Mallin estimointitulokset ovat kokonaisuudessaan liitteessä C.4.

Taulukko 6.7: Selittäjien ja skaalaselittäjän regressioparametrien estimaatit ja uskottavuusosamäärään perustuvat luottamusvälit kolmannessa mallissa,  $n=438$ .

Selittäjä	$\hat{\beta}_k$	95 % luottamusväli	
		valittajaisa	-4,88
huoltoennen	0,95	0,46	1,46
syytisaal	-0,72	-1,56	0,10
ikäero	-0,07	-0,14	-0,01
Skaalaselittäjä	$\hat{\zeta}$		
vaat_yh	2,26	1,92	2,64

Taulukko 6.8: Kynnysparametrien ja nominaalivaikutusten estimaatit kolmannessa mallissa.

Kynnys	$\hat{\theta}_j$	Keskivirhe	$\hat{\beta}_j$	Keskivirhe
1 2	-11,90	1,45	0,63	1,95
2 3	-4,33	0,99	-2,90	0,87
3 4	-4,01	0,98	2,23	0,66
4 5	2,11	1,27	6,13	6,39

Kolmannessa mallissa on vähemmän selittäjiä kuin aiemmissa, mikä johtuu siitä, että valittajan sukupuoli sisältää käräjäoikeuden päätöksen asumisriidan voittajasta, joka puolestaan vie muiden selittäjien selitysvoimaa. Tässä mallissa lasten edeltävä huolto- ja asumistilanne toimii paremmin selittäjänä kuin lasten vakiintunut asumistilanne, joka taas toimi tärkeänä selittäjänä muissa malleissa. Kyse voi olla siitä, että vakiintunut asumistilanne on tekijä, jota käräjä- ja hovioikeus harvoin katsovat eri tavalla, jolloin päätöksen muutos hovioikeudessa selittyy enemmän muilla tekijöillä.

Toisin kuin aiemmissa malleissa, kolmanteen malliin jäi selittäjäksi äidin tekemä syytös isän väkivaltaisuudesta, päihteidenkäytöstä tai mielenterveysongelmista. Tämä johtuu luultavasti siitä, että kolmas malli ei huomioi oikeuden todeksi katsomia syytöksiä, jotka aiemmissa malleissa havaittiin tärkeiksi selittäjiksi. Pelkästään esitetyt syytökset pitävät sisällään myös todeksi katsotut syytökset, jolloin tämä näkyy vaikutuksena myös esitettyä syytöstä vastaavan selittäjän kohdalla. Estimoitu vaikutus on äidin vastasuhteen kannalta positiivinen, mutta vastaava regressiokerroin ei kuitenkaan ole tilastollisesti merkitsevästi nolasta poikkeava ( $p$ -arvo 0,087).

Vanhempien välinen ikäero jää myös malliin selittäjäksi ja vaikutus on samanlainen kuin aiemmissakin malleissa, eli nuorempi osapuoli hyötty ikäerosta. Tässä mallissa ikäeroa vastaava regressiokerroin on kuitenkin hieman lähempänä nolaa kuin aiemmin, mutta edelleen tilastollisesti melkein merkitsevästi nolasta poikkeava ( $p$ -arvo 0,030).

### 6.3 Malleilla ennustaminen

Tarkastellaan ennustamiseen tarkoitettuja malleja, eli malleja 2 ja 3, ja tutkitaan millaisia ennusteita ne antavat vastemuuttujan eri luokille eri tapauksissa.

Skaala- ja nominaalivaikutuksilla laajennetun kertymä-logit-mallin laskema estimaatti  $\hat{\pi}_j$  luokan  $j$  todennäköisyydelle  $\pi_j$  jokaiselle  $j = 1, 2, 3, 4, 5$  saadaan kaavasta

$$\hat{\pi}_j = \frac{\exp(\hat{\eta}_j)}{1 + \exp(\hat{\eta}_j)} - \frac{\exp(\hat{\eta}_{j-1})}{1 + \exp(\hat{\eta}_{j-1})}, \quad j = 1, 2, 3, 4, 5,$$

missä

$$\hat{\eta}_j = \frac{\hat{\theta}_j - \mathbf{x}^\top \hat{\boldsymbol{\beta}} - w \hat{\beta}_j}{\exp(z \hat{\zeta})}.$$

Kaavassa  $\hat{\theta}_j$  on  $j$ :nnen kynnyksparametrin estimaatti,  $\mathbf{x}$  on tavallisten selittäjien tunnetuista arvoista muodostettu vektori,  $\hat{\boldsymbol{\beta}}$  on selittäjiä vastaavien regressiokertoimien estimaateista muodostettu vektori,  $w$  on nominaaliselittäjä eli vuoroasumisvaatimus (indikaattori),  $\hat{\beta}_j$  on nominaaliselittäjän vaikutuksen estimaatti  $j$ :nessä kynnyksessä,  $z$  on skaalaselittäjä eli yksinhuoltovaatimus (indikaattori) ja  $\hat{\zeta}$  on skaalaparametrin estimaatti.

### 6.3.1 Uuden riidan lopputuloksen ennustaminen

Oletetaan, että vanhemmat ovat juuri eronneet ja heillä on yksi yhteinen lapsi, joka on kouluikäinen eli yli 7-vuotias. Oletetaan, että äiti on jäänyt lapsen kanssa asumaan heidän entiseen yhteiseen kotiinsa. Vanhemmat ovat samanikäisiä ja kumpikin vanhempi käy töissä. Oletetaan, että kumpikin vanhempi hyväksyy yhteishuollon, mutta heillä on erimielisyys lapsen asumisesta. Vanhemmat eivät esitä toisiaan kohtaan syytöksiä.

Kun ennustamiseen käytetään mallia 2, niin edeltävä huoltotilanne on yhteishuolto ja asuminen äidin luona eli järjestysasteikon luokka 2 (huoltoennen=2), asuminen ei ole vakiintunut isälle (vakiintunut\_isalle=0), ikäeroa ei ole (ikäero=0), äiti ei ole työtön (aitityoton=0), syytöksiä ei ole tehty (syytaitia1=0, syytaitia2=0, syytisaa4=0), asuminen ei ole vielä vakiintunut äidille (vakiintunut\_aidille=0) ja lapsi ei ole alle 7-vuotias (vanhlapsa7=0). Kumpikaan osapuoli ei vaadi yksinhuoltoa eikä vuoroasumista, jolloin skaala- ja nominaaliselittäjät katoavat mallista. Käyttämällä taulukkojen 6.5 ja 6.6 estimaatteja saadaan

$$\hat{\eta}_j \approx \hat{\theta}_j - 2 \cdot 0,55.$$

Sijoittamalla mallin kynnyksparametrien estimaatit  $\hat{\theta}_1 \approx -5,41$ ,  $\hat{\theta}_2 \approx 1,35$ ,  $\hat{\theta}_3 \approx 1,56$  ja  $\hat{\theta}_4 \approx 8,14$  saadaan  $\hat{\eta}_1 \approx -6,51$ ,  $\hat{\eta}_2 \approx 0,25$ ,  $\hat{\eta}_3 \approx 0,46$  ja  $\hat{\eta}_4 \approx 7,04$ . Koska  $\theta_0 = -\infty$  ja  $\theta_5 = \infty$ , niin mallin antamat estimaatit eri luokkien todennäköisyyksille ovat

$$\begin{aligned}\hat{\pi}_1 &\approx \frac{\exp(-6,51)}{1 + \exp(-6,51)} - \underbrace{\lim_{x \rightarrow -\infty} \frac{\exp(x)}{1 + \exp(x)}}_{=0} \approx 0,0015 \\ \hat{\pi}_2 &\approx \frac{\exp(0,25)}{1 + \exp(0,25)} - \frac{\exp(-6,51)}{1 + \exp(-6,51)} \approx 0,5602 \\ \hat{\pi}_3 &\approx \frac{\exp(0,46)}{1 + \exp(0,46)} - \frac{\exp(0,25)}{1 + \exp(0,25)} \approx 0,0508 \\ \hat{\pi}_4 &\approx \frac{\exp(7,04)}{1 + \exp(7,04)} - \frac{\exp(0,46)}{1 + \exp(0,46)} \approx 0,3861 \\ \hat{\pi}_5 &\approx \underbrace{\lim_{x \rightarrow \infty} \frac{\exp(x)}{1 + \exp(x)}}_{=1} - \frac{\exp(-6,51)}{1 + \exp(-6,51)} \approx 0,0009.\end{aligned}$$

Käytännössä malli ennustaa, että oikeuden päätös lapsen huollon osalta on yhteishuolto, sillä yksinhuoltoluokkien estimoidut todennäköisyydet ovat mitättömän pieniä. Tämä on toivottavaa juuri silloin, kun kumpikaan osapuoli ei vaadi yksinhuoltoa. Äiti voittaa asumisriidan ennusteen mukaan noin 56 prosentin todennäköisyydellä ja isä noin 39 prosentin todennäköisyydellä. Tasapelille, eli tässä tapauksessa vuoroasumiselle malli ennustaa noin viiden prosentin todennäköisyyden. Jos lapsen asuminen olisi ehtinyt jo vakiintua äidille, niin malli ennustaisi äidin voittavan asumisriidan noin 84 prosentin todennäköisyydellä ja isän noin 13 prosentin todennäköisyydellä.

Mikäli lapsi olisi eron jälkeen jäänyt asumaan isän luokse, niin samalla asetelmalla isä olisi mallin perusteella voittanut asumisriidan noin 65 prosentin todennäköisyydellä



ja äiti noin 30 prosentin todennäköisyydellä. Jos asuminen olisi ehtinyt jo vakiintua isän luokse, niin mallin perusteella isä voittaisi asumisriidan noin 89 prosentin todennäköisyydellä ja äiti noin 8 prosentin todennäköisyydellä.

Mikäli asetelma olisi muuten samanlainen kuin alkuperäinen, mutta jompikumpi osapuoli vaatisi yksinhuoltoa, niin malliin tulee mukaan skaalavaikutus. Tällöin

$$\hat{\eta}_j = \frac{\hat{\theta}_j - 2 \cdot 0,55}{\exp(2,39)},$$

jolloin kynnyksparametrien estimaatit sijoittamalla estimoiduiksi todennäköisyyksiksi saataisiin  $\hat{\pi}_1 \approx 0,3548$ ,  $\hat{\pi}_2 \approx 0,1509$ ,  $\hat{\pi}_3 \approx 0,0048$ ,  $\hat{\pi}_4 \approx 0,1458$  ja  $\hat{\pi}_5 \approx 0,3436$ . Nyt yksinhuoltoluokat muuttuvat todennäköisimmiksi ja tasapeli äärimmäisen epätodennäköiseksi. Tässä asetelmassa osapuolet olisivat lähes tasapuolisissa asemissa.

Osapuolten voittomahdollisuuksia parantaisi lapsen vakiintunut asuinpaikka osapuolen luona. Äidin voittomahdollisuuksia parantaisi se, jos lapsi olisi alle kouluikäinen tai jos isä syyttäisi äitiä tapaamisten estämisestä tai vieraannuttamisesta. Isän voittomahdollisuuksia puolestaan parantaisi äidin työttömyys, isän tekemä syytös äidin väkivaltaisuudesta, päihteidenkäytöstä tai mielenterveysongelmista tai isää kohtaan tehty syytös lapsen seksuaalisesta hyväksikäytöstä. Ikäero parantaisi nuoremman osapuolen voittomahdollisuuksia.

Malli ei huomioi sitä, miten syytösten todeksi katsominen oikeudessa vaikuttaisi osapuolten voittomahdollisuuksiin. Toisen mallin (malli 1) perusteella voitaisiin kuitenkin ottaa huomioon se, että mikäli oikeus toteaa kumpaa tahansa osapuolta kohtaan syytöksen väkivaltaisuudesta, päihteidenkäytöstä tai mielenterveysongelmista todeksi, niin tämä parantaisi merkittävästi vastakkaisen osapuolen voittomahdollisuuksia. Lisäksi oikeuden toteama syytös tapaamisten estämisestä tai vieraannuttamisesta ei yhdessäkään mallissa noussut tilastollisesti merkitseväksi selittäjäksi, eli isän ei luultavasti kannata esittää tätä syytöstä äidistä ainakaan kevein perustein, sillä tämä ei näytä parantavan isän voittomahdollisuuksia, vaan ennemminkin heikentävän niitä.

### 6.3.2 Lopputuloksen ennustaminen käräjäoikeuden ratkaisemassa riidassa, josta ollaan valittamassa hovioikeuteen

Oletetaan, että jaksossa 6.3.1 tarkasteltu esimerkkitapaus ratkaistaan käräjäoikeudessa äidin hyväksi, eli lapsi jää vanhempien yhteishuoltoon ja asumaan äidin luokse. Käytetään mallia 3 ennustamaan lopputulos, kun isä valittaa päätöksestä hovioikeuteen.

Mallin selittäjillä isä on valittaja (valittajaisa=1) edeltävä huoltotilanne on yhteishuolto ja asuminen äidin luona (huoltoennen=2), syytöstä isää kohtaan ei ole esitetty (syytisaa1=0) ja ikäeroa ei ole (ikäero=0). Käyttämällä taulukkojen 6.7 ja 6.8 estimaatteja saadaan

$$\hat{\eta}_j \approx \hat{\theta}_j - (-4,88) \cdot 1 - 2 \cdot 0,95.$$

Estimoiduiksi todennäköisyyksiksi eri luokille saadaan  $\hat{\pi}_1 \approx 0,0173$ ,  $\hat{\pi}_2 \approx 0,9542$ ,  $\hat{\pi}_3 \approx 0,0076$ ,  $\hat{\pi}_4 \approx 0,0208$  ja  $\hat{\pi}_5 \approx 0,00005$ .

Malli ennustaa, että yli 95 prosentin todennäköisyydellä hovioikeuden päätös on yhteishuolto ja lapsen asuminen äidin luona. Isän voitolle malli ennustaa vain noin kahden prosentin todennäköisyyttä. Tällä mallilla isän mahdollisuudet paranisivat, jos hän olisi äitiä reilusti nuorempi tai mikäli lapsen edeltävä asuminen olisi ollut isän luona. Jos lapsi olisi eron jälkeen asunut isän luona ja käräjäoikeus olisi muuttanut asumisen äidille, niin malli ennustaisi isän voitolle hovioikeudessa noin 12 prosentin todennäköisyyden ja tasapelillekin noin neljän prosentin todennäköisyyden. Edelleen äiti olisi kuitenkin selvä voittajasuosikki voitettuaan riidan käräjäoikeudessa.

Jos isä vaatisikin hovioikeudessa vuoroasumista ja asetelma olisi muuten kuten alkuuperäinen, niin malliin tulee mukaan nominaalivaikutus. Tällöin

$$\hat{\eta}_j = \hat{\theta}_j - (-4,88) \cdot 1 - 2 \cdot 0,95 - \hat{\beta}_j.$$

Sijoittamalla mallin kynnyks- ja nominaaliparametrien estimaatit, niin estimoiduiksi todennäköisyyksiksi saataisiin  $\hat{\pi}_1 \approx 0,0321$ ,  $\hat{\pi}_2 \approx 0,6196$ ,  $\hat{\pi}_3 \approx 0,3460$ ,  $\hat{\pi}_4 \approx 0,0023$  ja  $\hat{\pi}_5 \approx 0$ . Malli siis ennustaa yhä todennäköisimmäksi vaihtoehdoksi yhteishuoltoa ja asumista äidin luona, jolle se ennustaa noin 62 prosentin todennäköisyyden. Tasapelille, eli yhden lapsen tapauksessa vuoroasumiselle malli ennustaa noin 35 prosentin todennäköisyyden. Äidin yksinhuollollekin malli ennustaa noin 3 prosentin todennäköisyyden ja isän voittoluokat ovat äärimmäisen epätodennäköisiä.

### 6.3.3 Ennustekyvyn arviointi

Koska käytössä ei ole testiaineistoa, joilla mallien 2 ja 3 ennustekykyä voitaisiin arvioida uusille huolto- ja asumisriidoille, niin ennustekykyä voidaan mitata ainoastaan sillä aineistolla, jolla mallit estimoitii. Mallin 2 yhteensopivuusindeksi oli noin 0,759 ja mallin 3 noin 0,851, eli malli 3 ennustaa lähtökohtaisesti paremmin. Mallin 3 ennustekyvyn hyvyttä selittää kuitenkin se, että se hyödyntää käräjäoikeuden päätöstä, kun malli 2 pyrkii ennustamaan lopputulosta ennen käräjäoikeuden päätöstä.

Taulukoissa 6.9 ja 6.10 on kummankin mallin antamat ennusteet ristiintaulukoituina oikeuden päätöksen kanssa. Ennusteena käytetään mallin perusteella todennäköisintä luokkaa.

Kumpikin malli osaa ennustaa yhteishuollon paremmin kuin yksinhuollon. Tämä havaittiin jo juurimallissa. Kummankin mallin tasapeliennusteista noin kaksi kolmasosaa osuu oikeaan.

Malli 2 osaa parhaiten ennustaa luokan 4, eli yhteishuollon ja isän lähihuoltajuuden. Kun malli ennustaa kyseistä luokkaa todennäköisimmäksi, niin noin 76 prosenttia vastaavista päätöksistä on ennusteen mukaisia. Kun malli ennustaa yhteishuoltoa ja äidin lähihuoltajuutta, niin noin 73 prosenttia vastaavista päätöksistä on ennusteen mukaisia. Yksinhuoltoluokilla oikein osuneiden ennusteiden osuudet jäävät alle 60 prosentin,

Taulukko 6.9: Mallin 2 ennusteet ristiintaulukoituna oikeuden päätöksen kanssa. 1=äidin yksinhuolto, 2=yhteishuolto, äidin lähihuoltajuus, 3=lasten jako tai vuoroasuminen, 4=yhteishuolto, isän lähihuoltajuus, 5=isän yksinhuolto. Vihreällä ovat täysin oikein ennustetut riidat. Vihreällä ja keltaisella ovat oikein ennustetut voittajat asumisriidassa.

		Päätös					Yht.	% lopputulos oikein	% asumisriidan voittaja oikein
		1	2	3	4	5			
<b>Ennuste</b>	1	73	24	2	5	18	122	59,8	79,5
	2	0	70	4	22	0	96	72,9	72,9
	3	0	2	11	3	0	16	68,8	68,8
	4	0	20	1	68	0	89	76,4	76,4
	5	17	4	1	16	53	91	58,2	75,8
	Yht.	90	120	19	114	71	414		

Taulukko 6.10: Mallin 3 ennusteet ristiintaulukoituna oikeuden päätöksen kanssa kuten taulukossa 6.9.

		Päätös					Yht.	% lopputulos oikein	% asumisriidan voittaja oikein
		1	2	3	4	5			
<b>Ennuste</b>	1	86	25	0	6	12	129	66,7	86,0
	2	0	81	2	5	0	88	92,0	92,0
	3	0	1	11	4	0	16	68,8	68,8
	4	0	15	3	91	0	109	83,5	83,5
	5	6	5	3	16	66	96	68,8	85,4
	Yht.	92	127	19	122	78	438		

mutta asumisriidan voittajan ennustamisessa ennusteet osuvat oikeaan paremmin.

Malli 3 osaa parhaiten ennustaa yhteishuollon ja äidin lähihuoltajuuden. Kun malli ennustaa kyseistä luokkaa todennäköisimmäksi, niin noin 92 prosenttia vastaavista päätöksistä on ennusteen mukaisia. Kun malli ennustaa yhteishuoltoa ja isän lähihuoltajuutta, niin noin 84 prosenttia vastaavista päätöksistä on ennusteen mukaisia. Yksinhuoltoluokilla sekä tasapeliluokilla noin kaksi kolmasosaa ennusteista osuu oikeaan. Asumisriidan voittajan osalta ennusteet kuitenkin osuvat oikeaan kaikissa muissa paitsi tasapeliluokassa reilusti yli 80 prosentissa vastaavista päätöksistä.

Juurimallin kohdalla huomattiin, että väärin ennustettujen yksinhuoltojen kohdalla päätös onkin usein vastapuolen yksinhuolto, eli täysin päinvastainen lopputulos. Kyse ei varsinaisesti ollut vakavasta ennustevirheestä vaan siitä, että näissä tapauksissa usein molempien yksinhuoltoluokkien todennäköisyydet olivat suurehkoja. Myös mallien 2 ja 3 kohdalla suuri osa väärin ennustetuista yksinhuolloista on tällaisia. Mallia soveltaessa tulisikin ottaa huomioon eri luokille estimoitujen todennäköisyyksien väliset erot eikä pelkästään katsoa sitä, mikä päätösluokka on todennäköisin.

Ennustavien mallien kohdalla havaitaan myös, että mallit näyttävät jonkin verran yliennustavan yksinhuoltoa, aivan kuten juurimallinkin kohdalla. Kuten juurimallin kohdalla todettiin, niin ongelmaan voisi mahdollisesti auttaa skaalaselittäjän muodostaminen toisella tavalla. Yksinhuoltoluokkien huonompaa ennustamista voi selittää myös se, että suurimmassa osassa aineiston riidoista riidellään vain asumisesta ja näistä luokista on siten enemmän havaintoja kuin muista. Tällöin kynnyksparametrien estimaattien keskivirheet ovat pienempiä vastaavien luokkien kynnyksissä. Mikäli aineistossa olisi ollut tasaisemmin havaintoja kaikista luokista, myös ennustekyky voisi olla tasaisempaa luokkien välillä.

## Luku 7

# Tulosten pohdintaa

Tässä tutkielmassa pyrittiin tutkimaan, mitkä tekijät selittävät hovioikeuden päätöstä huolto- ja asumisriidoissa, kun oikeuden päätös luokitellaan äidin voittoluokkiin, tasapeliin ja isän voittoluokkiin. Tämän pohjalta pyrittiin luomaan myös malleja, jotka ennustavat oikeuden päätöksen ja estimoivat päätösluokkien todennäköisyydet erilaisissa tilanteissa.

Oikeuden päätös katsottiin parhaaksi luokitella viiteen luokkaan siten, että luokittelussa tehtiin ero paitsi äidin voiton, tasapelin ja isän voiton, myös yhteishuollon ja yksinhuollon välillä. Malleissa tuli ottaa huomioon osapuolten vaatimukset, jotka käytännössä voivat sulkea jotkin päätösluokat pois. Tämän vuoksi kertymä-logit-regressiomallia laajennettiin skaala- ja nominaalivaikutuksilla. Nämä osoittautuivat toimivaksi tavaksi huomioida osapuolten vaatimukset. Mallien sopivuus parani verrattuna tilanteeseen, jossa skaala- ja nominaalivaikutuksia ei käytetty, ja ennustavissa malleissa nämä sulki halutut tulosluokat käytännössä mahdottomiksi.

Äitien ja isien vaatimukset asumisriidoissa eivät ole samanlaisia. Äidit vaativat asumisriidoissa isiä useammin lasten yksinhuoltoa. Asumisriidoissa vaaditaan vuoroasumista melko harvoin ja tämän vuoksi oikeus myös määrää asumisriidoissa vuoroasumista vain harvoin. Isät vaativat vuoroasumista hieman enemmän kuin äidit. Selvästi yleisin asumisriita-asetelma on kuitenkin sellainen, missä kumpikin osapuoli hyväksyy yhteishuollon ja riita koskee vain lasten asumista.

Kun katsotaan vain hovioikeuden päätöksiä, niin asumisriidoista noin 51,2 prosenttia päättyy äidin voittoon, 43,9 prosenttia isän voittoon ja 4,9 prosenttia tasapeliin, eli lasten jakoon tai vuoroasumiseen. Jos voittajaksi katsotaan se osapuoli, jonka asumisvaatimus on oikeuden päätöksen mukainen, niin asumisriidoista noin 52,8 prosenttia on äidin voittoja, 45,2 prosenttia isän voittoja ja 2 prosenttia tasapelejä. Äidit ja isät ovat valittajana yhtä usein.

## 7.1 Oikeuden päätöstä selittävät tekijät

### 7.1.1 Keskeisimmät taustatekijät

Kun selittäjiä tarkasteltiin yksittäin, niin kaikista parhaiten hovioikeuden päätöstä selitti se, kumpi osapuolista valitti käräjäoikeuden päätöksestä hovioikeuteen. Tämä johtui siitä, että tieto valittajasta sisälsi myös tiedon käräjäoikeuden päätöksestä asumisriidan osalta. Tämä tieto tuo yksinään jo paljon informaatiota, koska hovioikeus ottaa päätöksessään huomioon samoja seikkoja kuin käräjäoikeuskin. Yhden selittäjän mallissa tämä kertoo lähinnä sen, että käräjäoikeudessa riidan hävinnyt osapuoli on lähtökohtaisesti erittäin heikossa asemassa. Aineiston 500 asumisriidasta hovioikeus muutti asumisriidan lopputulosta 71 riidassa eli noin 14 prosentissa valituksista. Kun mukaan otetaan myös riidat, joissa asumisriidan voittaja ei muuttunut, niin päätöksistä 83 eli noin 16 prosenttia muuttui hovioikeudessa.

Parhaita yksittäisiä selittäjiä tilastollisen merkitsevyyden perusteella olivat sosiaaliviranomaisten suositus, edeltävä huolto- ja asumistilanne ja lasten vakiintunut asumistilanne, jotka jo ennalta voi olettaa merkityksellisiksi selittäjiksi.

Sosiaaliviranomaisten suosituksen osalta havaittiin usean selittäjän malleissa, että kummankaan osapuolen kohdalla suosituksen merkitys ei ollut merkittävästi toista osapuolta suurempi. Usean selittäjän mallissa sosiaaliviranomaisten suosituksen tuoma positiivinen yhteys osapuolen voittomahdollisuuksiin näyttäytyi numeerisesti suurempana isillä kuin äideillä, mutta tilastollisesti merkitsevää eroa ei havaittu. Yli puolet aineiston riidoista oli sellaisia, joissa suositusta ei ollut esitetty tai mahdolliset kaksi suositusta menivät tasan. Koska äidit tai isät eivät myöskään näytä saavan toista osapuolta useammin sosiaaliviranomaisten suosituksia ja koska suositus ei näytä painavan oikeudessa kummankaan kohdalla toista enemmän, niin äitien ja isien voisi sanoa olevan tämän tekijän osalta varsin tasavertaisissa asemassa.

Vakiintunut asumistilanne näytti usean selittäjän malleissa hyödyttävän kumpaakin osapuolta eikä varsinaisia eroja sukupuolten välillä havaittu. Lasten vakiintunut asuinpaikka oli huomattavasti useammin äidin kuin isän luona, mikä voi osittain selittää myös sitä, miksi äidit voittavat riitoja hieman useammin kuin isät. Vakiintuneen asuinpaikan merkitys oikeuden päätöksen kannalta ei kuitenkaan näytä olevan äideillä yhtään suurempi kuin isillä.

### 7.1.2 Syytökset

Äitiä kohtaan tehty syytös väkivaltaisuudesta, päihteistä tai mielenterveysongelmista (syytös 1) näytti yhden selittäjän mallin perusteella heikosti nostavan isän voittomahdollisuuksia. Myös usean selittäjän mallissa, jossa oikeuden todeksi katsomia syytöksiä ei otettu huomioon, tällä selittäjällä oli positiivinen ja tilastollisesti merkitsevä yhteys isän voittomahdollisuuksiin. Vastaavaa ei havaittu äidin kohdalla. Jos katsotaan kuitenkin

kin oikeuden todeksi katsomia syytöksiä, niin kummankin osapuolen kohdalla vastapuolesta todeksi katsottu syytös paransi osapuolen voittomahdollisuuksia. Äidistä todeksi katsottu syytös näyttäytyi malleissa numeerisesti suurempana kuin isästä todeksi katsottu syytös, mutta tilastollisesti merkitsevää eroa näiden merkityksessä äitien ja isien voittomahdollisuuksiin ei havaittu. Oikeuden toteamista syytöksistä tämä oli myös ainoa syytöskategoria, joka nousi malleissa esiin. Kyseistä syytöstä esitettiin useammassa kuin joka toisessa riidassa ja isät esittivät tätä syytöstä jonkin verran äitejä enemmän. Vaikuttaa siltä, että tätä syytöstä esitetään usein kevein perustein, mutta aiheellisena tämä syytös on hyvin keskeinen tekijä selittämään oikeuden päätöstä.

Syytös toista osapuolta kohtaan lasten tapaamisten estämisestä tai vieraannuttamisesta (syytös 2) oli yhden selittäjän mallin perusteella sekä isän että äidin kohdalla tämän voittomahdollisuuksia heikentävä tekijä. Selittäjä oli kummankin osapuolen kohdalla tilastollisesti merkitsevä. Isän tekemä syytös äidistä nousi esiin myös usean selittäjän mallissa, jossa oikeuden toteamat syytökset eivät olleet mukana. Myös tässä mallissa havaittiin, että tämän syytöksen tekeminen äidistä heikentää isän voittomahdollisuuksia. Oikeuden toteama syytös tapaamisten estämisestä tai vieraannuttamisesta ei noussut yhdessäkään mallissa esiin kummankaan voittomahdollisuuksia muuttavana tekijänä. Näin ollen vaikuttaa siltä, että tämän syytöksen esittäminen ei ole kummankaan kohdalla erityisen kannattavaa, sillä pelkästään syytöksen esittämisellä on negatiivinen yhteys osapuolen voittomahdollisuuksiin eikä edes se, että oikeus katsoo syytöksen toteennäytetyksi, näytä muuttavan tilannetta.

Kaappausuhkasyytökset (syytös 3) eivät nousseet yhdessäkään mallissa esiin ja näitä syytöksiä esitettiin ylipäättään melko vähän. Myös oikeuden todeksi katsomia kaappausuhkasyytöksiä oli vähän, minkä vuoksi tämän syytöksen merkityksestä ja aiheellisuudesta on vaikea sanoa aineiston perusteella juuri mitään.

Syytöksiä lasten seksuaalisesta hyväksikäytöstä (syytös 4) oli aineistossa vain vähän ja äiti esitti tätä syytöstä isästä huomattavasti useammin kuin isä äidistä. Usean selittäjän malleissa äidin tekemä syytös isästä näytti voimakkaasti heikentävän äidin voittomahdollisuuksia. Oikeuden todeksi katsomia syytöksiä isää kohtaan ei tässä aineistossa ollut lainkaan, mikä viittaa siihen, että äidit esittävät tämän syytöksen usein aiheetomasti. Negatiivinen yhteys äidin voittomahdollisuuksiin voi olla seurausta siitä, että näissä tapauksissa äiti on jo erittäin heikoissa asemissa ja tällä erittäin painavalla syytöksellä tilanne yritetään kääntää. Toisaalta aiheeton vakava syytös toisesta osapuolesta on varmasti myös tekijä, joka heikentää osapuolen asemia oikeuden edessä.

### 7.1.3 Ikätekijät

Vanhempien ikätietojen osalta havaittiin sekä yhden että useamman selittäjän malleissa, että osapuolten välisen ikäeron kasvaminen on positiivisessa yhteydessä nuoremman osapuolen voittomahdollisuuksiin. Interaktiotarkastelulla havaittiin myös, että äidit hyötyvät isää enemmän siitä, jos he ovat vastapuolta nuorempia. Ikäero oli mukana tilas-

tollisesti merkitseväenä tai melkein merkitseväenä selittäjänä jokaisessa usean selittäjän mallissa.

Myös lasten iällä havaittiin olevan merkitystä. Vanhimman lapsen ikä osoittautui paremmaksi selittäjäksi kuin nuorimman lapsen ikä tai lasten ikien keskiarvo, ja pelkkä tieto vanhimman lapsen kouluikäisyydestä toimi paremmin kuin ikä jatkuvana selittäjänä. Usean selittäjän malleissa se, että vanhin lapsi oli alle kouluikäinen, paransi äidin voittomahdollisuuksia. Tulos on samansuuntainen kuin Palo-Revolla (2015), joka havaitsi, että vanhimman lapsen iän kasvaessa isän voittomahdollisuudet paranevat.

Tuomarikokoonpanon ikäkeskiarvo nousi heikosti esiin yhden selittäjän mallissa, mutta ei yhdessäkään usean selittäjän mallissa. Yhden selittäjän mallin perusteella tuomarien ikäkeskiarvon kasvaessa isän voittomahdollisuudet näyttivät lievästi paranevan. Yhteyttä tuomarien ikäkeskiarvon ja oikeuden päätöksen välillä katsottiin myös aineistossa läpi tarkemmin, mutta mitään helposti havaittavaa yhteyttä näiden väliltä ei kuitenkaan löytynyt. Kyse voi olla sattumalta ilmenneestä heikosta yhden selittäjän yhteydestä, joka katoaa, kun muita tekijöitä otetaan huomioon. Asiaa olisi kiinnostavaa tutkia tarkemmin.

#### 7.1.4 Muut tekijät

Äidin työttömyys nousi esiin sekä yhden että usean selittäjän malleissa tilastollisesti melkein merkitseväenä parantaen hieman isän voittomahdollisuuksia. Mielenkiintoisesti isän työttömyys selittäjänä ei noussut vastaavalla tavalla esiin. Näissä muuttujissa oli jonkin verran puuttuvuutta ja ne veivät siten mallien kokonaishavaintomääriä alas.

Myös äidin avustajan sukupuolella havaittiin olevan tilastollisesti merkitsevä yhteys osapuolten voittomahdollisuuksiin. Usean selittäjän malleissa havaittiin, että jos äidin avustaja on mies, niin tällä on positiivinen yhteys isän voittomahdollisuuksiin. Tämä oli ainoa vastaavankaltainen sukupuolivaikutus kuin mitä Pere, Lahti ja Sutela (2017) havaitsivat. Tuomarikokoonpanon osalta sukupuolivaikutuksia ei havaittu. Tuomarien, osapuolten ja näiden avustajien sukupuolivaikutuksia ja näiden välisiä interaktioita olisi kuitenkin mielenkiintoista tutkia tarkemmin.

Mallien perusteella oikeuden päätökset eivät näytä poikkeavan toisistaan hovioikeuksien välillä. Helsingin hovioikeus nousi ainoana hovioikeutena esiin usean selittäjän malleissa. Tällä havaittiin olevan positiivinen yhteys isän voittomahdollisuuksiin ja selittäjä oli tilastollisesti melkein merkitsevä.

Aikatrendi sopi tilastollisesti merkitseväenä usean selittäjän malleihin. Aikatrendin perusteella isien voittomahdollisuudet ovat kasvaneet vuodesta 2000 vuoteen 2016.



## 7.2 Käräjäoikeuden päätöksen muuttumista selittävät tekijät

Kun usean selittäjän mallilla 3 pyrittiin osittain tutkimaan myös sitä, mitkä tekijät selittävät parhaiten päätöksen muuttumista hovioikeudessa, niin selittäjiä jäi malliin vain vähän. Äidin voittomahdollisuuksia parantavaksi selittäjäksi nousi tässä kohtaa vain isää kohtaan tehty syytös väkivaltaisuudesta, päihteistä tai mielenterveysongelmista, joka ei kuitenkaan ollut tilastollisesti merkitsevä. Vastaavia selittäjiä parantamaan isän voittomahdollisuuksia ei noussut esiin. Malliin jäi kuitenkin tilastollisesti melkein merkitseväksi selittäjäksi ikäero, joka näytti parantavan nuoremman osapuolen voittomahdollisuuksia kuten muissakin malleissa. Interaktiotarkastelulla huomattiin myös, että äidit näyttävät hyötävän isiä enemmän siitä, jos he ovat nuorempia, mutta interaktiota ei otettu mukaan malliin. Tärkein selittäjä tässä mallissa valittajan sukupuolen ohella oli edeltävä huolto- ja asumistilanne, joka oli tilastollisesti erittäin merkitsevä. Selittäjänä tämä on kuitenkin varsin ilmeinen.

Tästä voi päätellä, että ei välttämättä ole sellaisia helposti havaittavia tekijöitä, jotka systemaattisesti selittäisivät käräjäoikeuden päätöksen muuttumista. Kukin riita on omanlaisensa eikä niissä riidoissa, joissa päätös muuttuu, ole selviä yhteisiä piirteitä. Ikäero ja isää kohtaan tehty syytös väkivaltaisuudesta, päihteistä tai mielenterveysongelmista nousivat mallissa marginaalisesti esiin, mutta suurin osa hovioikeuden päätöksestä selittyy valituksenalaisella käräjäoikeuden päätöksellä ja edeltävällä huolto- ja asumistilanteella.

Koska olisi myös mielenkiintoista tutkia sitä, onko jommankumman osapuolen kannattavampaa valittaa käräjäoikeuden päätöksestä hovioikeuteen ja koska valittajan sukupuoli selittäjänä ei sellaisenaan vielä kerro tätä, niin yksi vaihtoehto olisi ottaa tarkempi käräjäoikeuden päätös mukaan malliin interaktion valittajan sukupuolen kanssa. Tällöin malliin tulisi sovitetuksi käräjäoikeuden päätöksen vaikutus erikseen äidille ja isälle. Tätä kokeiltiin myös tässä tutkielmassa, mutta ilmeisesti interaktioiden määrää skaala- ja nominaaliselittäjien kanssa aiheutti sen, että mallia ei voinut estimoida ja tällainen malli jouduttiin sivuuttamaan.

## 7.3 Lopuksi

Äitien ja isien vaatimukset asumisriidoissa eivät jakaudu samalla tavalla, minkä vuoksi pelkästään riitojen lopputulosten jakautuminen eri luokissa ei yksin paljasta sitä, pärjäävätkö äidit vai isät asumisriidoissa paremmin. Äidit kuitenkin saavat isiä hieman useammin vaatimuksiansa vastaavia päätöksiä. Eräs syy tälle on kuitenkin se, että lasten asuminen ennen riitaa on useammin äidin kuin isän luona ja monesti asumistilanne on ehtinyt vakiintua. Lasten vakiintuneen asumistilanteen muuttaminen vaatii usein painavia perusteluja. Tämän tutkielman perusteella tällainen painava perustelu on ainakin

vanhemman väkivaltaisuus, päihteidenkäyttö tai mielenterveysongelmat.

Koska myös olosuhdeselvityksen suositukset nousivat erittäin tärkeinä esiin, niin epäilemättä asuinolosuhteet ja vakiintunut asuinpaikka painavat oikeuden päätöksessä eniten. Selvästi heikompina, mutta kuitenkin merkityksellisinä seikkoina esiin nousivat myös vanhimman lapsen ikä sekä vanhempien välinen ikäero. Nuorempi osapuoli on lähtökohtaisesti hieman paremmissa asemissa ja mitä suurempi ikäero on, sitä enemmän tällä näyttäisi olevan merkitystä. Äiti saa parhaan asetelman, kun vanhin lapsi on alle kouluikäinen. Äitien asema voi myös hieman parantua silloin, kun vanhin lapsi on tyttö. Isät puolestaan näyttävät pärjäävän paremmin silloin, kun äiti on työtön ja kun äidin avustaja oikeudessa on mies.

Mahdollisia selittäjiä oli aineistossa niin paljon, että kaikkia mahdollisia yhteyksiä on käytännössä mahdoton havaita. Analyyseja voisi kuitenkin tarkentaa esimerkiksi tutkimalla tarkemmin selittäjien välisiä interaktioita tai kokeilemalla myös muita selittäjiä nominaaliselittäjiksi, koska oletus tavallisten selittäjien yhtäläisestä vaikutuksesta jokaisen päätösluokan välillä ei välttämättä päde.

Ennustavilla malleilla saatiin kohtuullisen hyviä tuloksia erityisesti, jos niillä ennustettiin asumisriidan voittajaa. Ennusteet myös olivat lähtökohtaisesti järkeviä eivätkä ne olleet pahasti ristiriidassa osapuolten vaatimusten kanssa. Yksinhuollon osalta mallien ennustekykyä olisi mahdollisesti voinut parantaa toisenlaisen skaalaselittäjän avulla. Tämä jouduttiin kuitenkin sivuuttamaan, koska esimerkiksi useamman skaalaselittäjän käyttäminen aiheutti malleille muita ongelmia. Ennustaville malleille olisi myös ollut mielenkiintoista soveltaa testiaineistoa tai ristiinvalidointia, jolloin ennustekykyä olisi saatu arvioitua paremmin uusille havainnoille. Ennustamisessa voitaisiin käyttää myös suuremman selittäjämäärän malleja, joita epäiltiin ylisopiviksi. Mahdollinen ylisopivuus paljastuisi kuitenkin vasta sitten, kun mallia olisi sovellettu uusiin havaintoihin.

# Lähteet

- Agresti, Alan 2018: An Introduction to Categorical Data Analysis, 3. laitos. Wiley. Hoboken.
- Agresti, Alan 2010: Analysis of Ordinal Categorical Data, 2. laitos. Wiley. Hoboken.
- Agresti, Alan 2015: Foundations of Linear and Generalized Linear Models. Wiley. Hoboken.
- Agresti, Alan ja Tarantola, Claudia 2018: Simple Ways to Interpret Effects in Modeling Ordinal Categorical Data. *Statistica Neerlandica*, 72(3), 210–223.
- Christensen, Rune Haubo B. 2012: Sensometrics: Thurstonian and Statistical Models. Väitöskirja, Tanskan teknillinen yliopisto. IMM-PHD-2012, No. 271.
- Christensen, Rune Haubo B. 2018a: Cumulative Link Models for Ordinal Regression with the R Package ordinal. [https://cran.r-project.org/web/packages/ordinal/vignettes/clm\\_article.pdf](https://cran.r-project.org/web/packages/ordinal/vignettes/clm_article.pdf) (viitattu 18.11.2019).
- Christensen, Rune Haubo B. 2018b: Regression Models for Ordinal Data. Paketin versio 2019.4-25. <https://cran.r-project.org/web/packages/ordinal/index.html> (viitattu 18.11.2019).
- Cox, Christopher 1995: Location-Scale Cumulative Odds Models for Ordinal Data, *Statistics in Medicine*, 14, 1191–1203.
- Harrell, Frank E. 2015: Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, 2. laitos. Springer. New York.
- Held, Leonhard ja Sabanés Bové, Daniel 2014: Applied Statistical Inference. Springer. Berlin.
- Hosmer, David W., Lemeshow, Stanley ja Sturdivant, Rodney X. 2013: Applied Logistic Regression, 3. laitos. Wiley. Hoboken.
- Laki lapsen huollosta ja tapaamisoikeudesta 361/1983: <https://www.finlex.fi/fi/laki/ajantasa/1983/19830361> (viitattu 31.10.2019).

Laki oikeudenkäymiskaaren muuttamisesta 381/2003: <https://www.finlex.fi/fi/laki/alkup/2003/20030381> (viitattu 18.11.2019).

Laki oikeudenkäymiskaaren muuttamisesta 650/2010: <https://www.finlex.fi/fi/laki/alkup/2010/20100650> (viitattu 18.11.2019).

Laki oikeudenkäymiskaaren muuttamisesta 386/2015: <https://www.finlex.fi/fi/laki/alkup/2015/20150386> (viitattu 18.11.2019).

Laurila, Petri 2011: Suomen Lakimiehet 2011. Talentum Media Oy. Helsinki.

McCullagh, Peter 1980: Regression Models for Ordinal Data. *Journal of the Royal Statistical Society*. Series B (Methodological), 42, 109–142.

Palo-Repo, Mari 2015: Lasten huolto- ja asumisriidat Helsingin hovioikeudessa 2003–2006. Pro gradu -tutkielma. Helsingin yliopisto, Valtiotieteellinen tiedekunta.

Pere, Pekka; Lahti, Tuomas ja Sutela, Mika 2017: Sentences and Prosecutors' Demands for Aggravated Drunk Driving in Finland. *International Journal of Comparative and Applied Criminal Justice*, 42(2–3), 157–175.

Pratt, John W 1981: Concavity of the Log Likelihood. *Journal of the American Statistical Association*, 76(373), 103–106.

Rovaniemen hovioikeuspiiri 2016: Oikeudenkäynti lapsiasioissa. Käsittelijän opas. Rovaniemen hovioikeuspiirin tuomioistuinten laatuhanke. Laatutyöryhmän 1/2014 raportti lapsiasioiden työprosessien yhtenäistämiseksi ja kehittämiseksi. Toimittanut: Helevä-Vuoti, Kaisa ja Leikas, Minna.

Tutz, Gerhard ja Berger, Moritz 2016: Separating Location and Dispersion in Ordinal Regression Models. *Econometrics and Statistics*, 2, 131–148.

Valkama, Elisa ja Laasola, Marjukka 2008: Lasten huoltoriidat hovioikeuksissa. Oikeuspoliittinen tutkimuslaitos. Helsinki.

## Liite A

# Oikeustapausaineiston muuttujat

Alla on kuvaus aineiston muuttujista sellaisena kuin ne ovat alkuperäisessä aineistossa. Analyseissa käytetyt muuttujat on kuvailtu erikseen liitteessä B.

### **nro**

Päätöksen taltionumero hovioikeuden taltiokirjassa.

### **diaarinro**

Päätöksen diaarinumero.

### **hovioikeus**

Hovioikeus, jossa asia on käsitelty.

- Helsinki
- Itä-Suomi
- Rovaniemi
- Turku
- Vaasa

### **valpvm**

Päivämäärä, jolloin valitus on saapunut hovioikeuteen.

## **antopvm**

Hovioikeuden päätöksen antamispäivämäärä.

## **valittajasp**

Valittajan sukupuoli.

1 Nainen

2 Mies

## **valittajasynt**

Valittajan syntymäaika.

## **valittajatyo**

Valittajan pääasiallinen toiminta.

0 Ei tietoa

1 Töissä

2 Opiskelija

3 Hoitaa lapsia kotona

4 Työtön tai työkyvytön

5 Eläkeläinen

## **valittajaulkom**

Tieto valittajan ulkomaalaistaustaisuudesta. Ulkomaalaisuus on jaettu alla oleviin kategorioihin inhimillisen kehityksen indeksin perusteella.

0 Suomi

1 EU-maat, Sveitsi, Norja, Anglo-Amerikka, Japani, Etelä-Korea

2 Latinalainen Amerikka, Turkki, Venäjä, Kazakstan, Albania, Bosnia ja Hertsegovina, Ukraina

3 Lähi-Itä, Kaakkois-Aasia, Intia, Pohjois-Afrikka

4 Saharan etelänpuoleinen Afrikka, Bangladesh, Pakistan

## **vpsp**

Vastapuolen sukupuoli. Katso valittajasp.

## **vpsynt**

Vastapuolen syntymäaika.

## **vptyo**

Vastapuolen pääasiallinen toiminta. Katso valittajatyo.

## **vpulkom**

Tieto vastapuolen ulkomaalaistaustaisuudesta. Katso valittajaulkom.

## **aitihetu**

Äidin henkilötunnus.

## **isahetu**

Isän henkilötunnus.

## **vakiintunut**

Tieto siitä, onko lapsen asuinpaikka vakiintunut jommankumman vanhemman luokse ennen oikeuden päätöstä. Katso jakso 4.3.

- 1 Vakiintunut äidin luokse
- 0 Ei vakiintunut kummankaan luokse
- 1 Vakiintunut isän luokse

## **huoltoennen**

Huolto- ja asumistilanne ennen käräjäoikeuden päätöstä. Katso jakso 4.3.

- 1 Äidin yksinhuolto
- 2 Äidin yksinhuolto siten, että isällä on tiedonsaantioikeus

- 3 Yhteishuolto tehtävänjakomääräyksin siten, että lapset asuvat äidin luona
- 4 Yhteishuolto siten, että lapset asuvat äidin luona
- 5 Lasten jako
- 6 Yhteishuolto ja vuoroasuminen, ei tietoa kumman luona lapset ovat kirjoilla
- 61 Yhteishuolto ja vuoroasuminen siten, että lapset ovat kirjoilla äidin luona
- 62 Yhteishuolto ja vuoroasuminen siten, että lapset ovat jaetusti kirjoilla molempien vanhempien luona
- 63 Yhteishuolto ja vuoroasuminen siten, että lapset ovat kirjoilla isän luona
- 7 Yhteishuolto siten, että lapset asuvat isän luona
- 8 Yhteishuolto tehtävänjakomääräyksin siten, että lapset asuvat isän luona
- 9 Isän yksinhuolto siten, että äidillä on tiedonsaantioikeus
- 10 Isän yksinhuolto
- 99 Muu tilanne, esimerkiksi huostaanotto

## **valittajavaat**

Valittajan vaatimus hovioikeudessa. Katso huoltoennen 1–10.

## **vpvaat**

Vastapuolen vaatimus hovioikeudessa. Katso huoltoennen 1–10.

## **paatos**

Hovioikeuden päätös. Katso huoltoennen.

## **jako**

Sanallinen tieto siitä, kuinka lasten huolto ja asuminen on jaettu, jos hovioikeuden päätös on ollut lasten jako

## **valittajavaatko**

Valittajan vaatimus käräjäoikeudessa. Katso huoltoennen 1–10.



## **vpvaatko**

Vastapuolen vaatimus käräjäoikeudessa. Katso huoltoennen 1–10.

## **kopaatos**

Käräjäoikeuden päätös, josta on valitettu hovioikeuteen.

1–99 Katso huoltoennen

999 Käräjäoikeus ei ole katsonut olevansa toimivaltainen käsittelemään asiaa

## **muutoshovi**

Tieto siitä, onko hovioikeus muuttanut käräjäoikeuden päätöstä lasten huollon tai asu-  
misen osalta

0 Ei

1 Kyllä

## **vsukup**

Voittajan sukupuoli. Katso voittajan määrittely jaksosta 4.5.

0 Ei voittajaa

1 Äiti

2 Isä

## **syotpvm**

Päivämäärä, jolloin havainto on kirjattu aineistoon.

## **tuomari1, tuomari2 ja tuomari3**

Hovioikeuden ratkaisijoiden nimet.

## **tuosyntv1, tuosyntv2 ja tuosyntv3**

Hovioikeuden ratkaisijoiden syntymävuodet.

## **tuosp1, tuosp2 ja tuosp3**

Hovioikeuden ratkaisijoiden sukupuoli.

- 1 Nainen
- 2 Mies

## **titteli**

Päätöksen ratkaisijoiden tittleistä korkein.

- Hovioikeuden presidentti
- Hovioikeudenlaamanni
- Hovioikeudenneuvos

## **erimielta**

Tieto hovioikeuden päätöksen yksimielisyydestä. Jos kolmihenkisen tuomariston päätös ei ole yksimielinen, tuomarit suorittavat äänestyksen.

- 0 Päätös on yksimielinen
- 1 Tuomari 1 eri mieltä
- 2 Tuomari 2 eri mieltä
- 3 Tuomari 3 eri mieltä
- 4 Päätös ei ole yksimielinen, mutta eri mieltä oleva tuomari ei selviä asiakirjoista

## **aitied**

Tieto äidin avustajasta hovioikeudessa.

- 0 Ei avustajaa tai tieto puuttuu
- 1 Julkinen oikeusavustaja
- 2 OTM/OTK/Lakimies/Varatuomari tms.
- 3 Asianajaja

## **aitiedsp**

Äidin avustajan sukupuoli

- 0 Ei tietoa
- 1 Nainen
- 2 Mies

## **isaed**

Tieto isän avustajasta hovioikeudessa. Katso aited.

## **isaedsp**

Isän avustajan sukupuoli. Katso aitedsp.

## **lapsilkm**

Tieto siitä, kuinka monen lapsen huollosta tai asumisesta osapuolet riitelevät hovioikeudessa.

## **lapsi1synt, lapsi2synt, lapsi3synt, lapsi4synt, lapsi5synt ja lapsi6synt**

Lasten syntymäajat järjestyksessä vanhimmasta lapsesta alkaen.

## **lapsi1ika, lapsi2ika, lapsi3ika, lapsi4ika, lapsi5ika ja lapsi6ika**

Lasten iät järjestyksessä vanhimmasta lapsesta alkaen.

## **lapsi1sp, lapsi2sp, lapsi3sp, lapsi4sp, lapsi5sp ja lapsi6sp**

Lasten sukupuolet järjestyksessä vanhimmasta lapsesta alkaen.

- 1 Tyttö
- 2 Poika

## **aititulot**

Äidin nettotulot euroina kuukaudessa.

## **isatulot**

Isän nettotulot euroina kuukaudessa.

## **syyt1**

Osapuolten esittämät syytökset väkivallasta, päihteistä tai mielenterveysongelmista.

- 0 Ei syytöksiä
- 1 Isä syyttää äitiä tai tämän uutta puolisoa
- 2 Äiti syyttää isää tai tämän uutta puolisoa
- 3 Molemmat syyttävät toisiaan

## **oiktodsyyt1**

Tieto siitä, onko oikeus katsonut syytöksen väkivallasta, päihteistä tai mielenterveysongelmista toteennäytetyksi.

- 0 Ei toteennäytettyjä syytöksiä
- 1 Syytös äitiä kohtaan katsottu toteennäytetyksi
- 2 Syytös isää kohtaan katsottu toteennäytetyksi
- 3 Syytös molempia kohtaan katsottu toteennäytetyksi

## **syyt2**

Osapuolten esittämät syytökset tapaamisten estämisestä tai lapsen vieraannuttamisesta toisesta vanhemmastaan. Katso syytos1.

## **oiktodsyyt2**

Tieto siitä, onko oikeus katsonut syytöksen tapaamisten estämisestä tai vieraannuttamisesta toteennäytetyksi. Katso oiktodsyytos1.

## **syyt3**

Osapuolten esittämät syytökset lapsen kaappausuhasta ulkomaille. Katso syytos1.

## **oiktodsyyt3**

Tieto siitä, onko oikeus katsonut syytöksen kaappausuhasta toteennäytetyksi. Katso oiktodsyytos1.

## **syyt4**

Osapuolten esittämät syytökset lapsen seksuaalisesta hyväksikäytöstä. Katso syytos1.

## **oiktodsyyt4**

Tieto siitä, onko oikeus katsonut syytöksen lapsen seksuaalisesta hyväksikäytöstä toteennäytetyksi. Katso oiktodsyytos1.

## **suositus**

Sosiaaliviranomaisten tekemän olosuhdeselvityksen suositus asian ratkaisemisesta.

0 Ei olosuhdeselvitystä tai ei tietoa olosuhdeselvityksestä

1–10 Katso huoltoennen

999 Olosuhdeselvitys on tehty, mutta suositusta ei ole annettu

1234 Suositus ainakin lapsen asumisesta äidin luona

4567 Suositus yhteishuollosta

12910 Suositus jommankumman vanhemman yksinhuollosta

78910 Suositus ainakin lapsen asumisesta isän luona

## **viimpaat**

Viimeisin päätös tai sopimus lasten huollosta tai asumisesta ennen käräjäoikeuden päätöstä.

0 Ei voimassa olevaa päätöstä/sopimusta tai tieto puuttuu

1 Osapuolten tekemä sopimus

2 Käräjäoikeuden tekemä väliaikais määräys

3 Oikeuden päätös

## **oikeudessa**

Tieto siitä, ovatko osapuolet riidelleet asiasta oikeudessa aiemmin.

0 Ei

1 Kyllä

## **esittsp**

Hovioikeuden esittelijän sukupuoli.

1 Nainen

2 Mies

## Liite B

# Analyyseissa käytetyt muuttujat

Alla on tiedot kaikista niistä muuttujista, joita on käytetty malleissa.

### **paa (vastemuuttuja)**

Hovioikeuden päätös viisiportaisella asteikolla.

- 1 Äidin yksinhuolto, äidin yksinhuolto isän tiedonsaantioikeudella tai rajoitettu yhteishuolto siten, että lasten asuminen on äidillä.
- 2 Yhteishuolto siten, että lasten asuminen on äidillä
- 3 Lasten jako tai vuoroasuminen
- 4 Yhteishuolto siten, että lasten asuminen on isällä
- 5 Isän yksinhuolto, isän yksinhuolto äidin tiedonsaantioikeudella tai rajoitettu yhteishuolto siten, että lasten asuminen on isällä.

### **vaat\_yh**

Tieto siitä, vaatiiko kumpikaan osapuoli yksinhuoltoa. Tämä kattaa myös yhteishuollon tehtävänjakomääräyksen.

- 0 Ei
- 1 Kyllä

### **vaat\_vuoroas**

Tieto siitä, vaatiiko kumpikaan osapuoli vuoroasumista. Katso vaat\_yh.

## **valittajaisa**

Tieto siitä, onko isä valittaja. Jos isä on valittaja, niin isä on hävinnyt asumisriidan käräjäoikeudessa. Jos isä ei ole valittaja, niin äiti on hävinnyt asumisriidan käräjäoikeudessa.

0 Ei

1 Kyllä

## **suositus**

Sosiaaliviranomaisten tekemän olosuhdeselvityksen suositus asian ratkaisemisesta järjestysasteikollisena muuttujana.

-1 Suositus äidin yksinhuollolle tai äidin lähivanhemmuudelle.

0 Ei suositusta

1 Suositus isän yksinhuollolle tai isän lähivanhemmuudelle.

## **vakiintunut**

Kuten vastaava aineiston alkuperäinen muuttuja. Katso liite A.

## **huoltoennen**

Edeltävä huolto- ja asumistilanne samalla viisiportaisella asteikolla kuin hovioikeuden päätös. Katso paa.

## **aitityoton**

Tieto siitä, onko äiti työtön.

0 Ei

1 Kyllä

## **isatyoton**

Tieto siitä, onko isä työtön. Katso aityoton.



## **aitiedsp**

Äidin avustajan sukupuoli. Kuten vastaava muuttuja aineistossa. Katso liite A.

## **isaedsp**

Isän avustajan sukupuoli. Kuten vastaava muuttuja aineistossa. Katso liite A.

## **turku**

Tieto siitä, onko päätös tehty Turun hovioikeudessa.

0 Ei

1 Kyllä

## **helsinki**

Tieto siitä, onko päätös tehty Helsingin hovioikeudessa. Katso turku.

## **rovaniemi**

Tieto siitä, onko päätös tehty Rovaniemen hovioikeudessa. Katso turku.

## **vaasa**

Tieto siitä, onko päätös tehty Vaasan hovioikeudessa. Katso turku.

## **ita**

Tieto siitä, onko päätös tehty Itä-Suomen hovioikeudessa. Katso turku.

## **vam**

Tieto siitä, onko viimeisin päätös väliaikais määräys.

0 Ei

1 Kyllä

## **oikeudessa**

Kuten vastaava muuttuja aineistossa. Katso liite A.

## **ulkom**

Ulkomaalaistaustaisuus järjestysasteikollisena muuttujana.

-1 Äiti on ulkomaalaistaustainen

0 Kumpikaan ei ole ulkomaalaistaustainen tai molemmat ovat ulkomaalaistaustaisia

1 Isä on ulkomaalaistaustainen

## **aitiulkom**

Tieto siitä, onko äiti ulkomaalaistaustainen.

0 Ei

1 Kyllä

## **isaulkom**

Tieto siitä, onko isä ulkomaalaistaustainen. Katso aitiulkom.

## **ikaero**

Isän ja äidin ikäero vuosina. Kummankin ikä on muodostettu laskemalla päivien lukumäärä hovioikeuden päätöksen antamispäivämäärän ja valittajan syntymäajan välillä ja jakamalla tämä luvulla 365.

## **nuorlapsika**

Nuorimman lapsen ikä. Ikä on muodostettu laskemalla päivien lukumäärä hovioikeuden päätöksen antamispäivämäärän ja lapsen syntymäajan välillä ja jakamalla tämä luvulla 365.

## **vanhlapsika**

Vanhimman lapsen ikä.

## **lapsika**

Lasten ikien keskiarvo.

## **nuorinlapsipoika**

Tieto siitä, onko nuorin lapsi poika.

0 Ei

1 Kyllä

## **vanhinlapsipoika**

Tieto siitä, onko vanhin lapsi poika. Katso nuorinlapsipoika.

## **vanhlapsa7**

Tieto siitä, onko vanhin lapsi alle 7-vuotias.

0 Ei

1 Kyllä

## **naisenemmisto**

Tieto siitä, onko hovioikeuden kolmesta ratkaisijasta suurin osa naisia.

0 Ei

1 Kyllä

## **tuomika**

Hovioikeuden ratkaisijoiden ikien keskiarvo. Kunkin ratkaisijan ikä on laskettu vähentämällä ratkaisijan syntymävuosi siitä vuodesta, jolloin hovioikeuden päätös on annettu.

## **tuokavar**

Hovioikeuden ratkaisijoiden ikien varianssi.

## **syyt1, syyt2, syyt3 ja syyt4**

Osapuolten esittämät syytökset koodattuna järjestysasteikolliseksi muuttujaksi. Jos molemmat osapuolet syyttävät toisiaan, syytökset kumoavat toisensa.

-1 Isä syyttää äitiä

0 Ei syytöksiä tai molemmat syyttävät toisiaan

1 Äiti syyttää isää

### **syytaitia1, syytaitia2, syytaitia3 ja syytaitia4**

Tieto siitä, onko äitiä kohtaan esitetty syytös. Syytös on kirjattu myös, jos molemmat ovat syyttäneet toisiaan.

0 Ei

1 Kyllä

### **syytisaa1, syytisaa2, syytisaa3 ja syytisaa4**

Tieto siitä, onko isää kohtaan esitetty syytös. Katso syytaitia1.

### **oiktodsytyt1, oiktodsytyt2, oiktodsytyt3 ja oiktodsytyt4**

Tieto siitä, onko oikeus katsonut syytöksen toteennäytetyksi. Jos syytökset on todettu molempia osapuolia kohtaan toteennäytetyksi, nämä kumoavat toisensa.

-1 Syytös äitiä kohtaan katsottu toteennäytetyksi

0 Ei toteennäytettyjä syytöksiä tai syytökset molempia kohtaan katsottu toteennäytetyksi

1 Syytös isää kohtaan katsottu toteennäytetyksi

### **oiktodaitia1, oiktodaitia2, oiktodaitia3 ja oiktodaitia4**

Tieto siitä, onko syytös äitiä kohtaan katsottu toteennäytetyksi.

0 Ei

1 Kyllä

### **oiktodisaa1, oiktodisaa2, oiktodisaa3 ja oiktodisaa4**

Tieto siitä, onko syytös isää kohtaan katsottu toteennäytetyksi. Katso oiktodaitia1.

## **aika**

Keskistetty aikamuuttuja vuoden tarkkuudella. Koska ajallisesti keskimäinen havainto on vuodelta 2012, on aikamuuttuja laskettu vähentämällä päätöksen antamisvuodesta luku 2012.

## Liite C

# R-koodit ja tulosteet

### C.1 Juurimalli

```
> library(ordinal)
> fit <- clm(paa~vakiintunut_aidille+vakiintunut_isalle+
            suositusaidille+suositusisalle+
            oiktodaitia1+oiktodisaa1+
            syytisaa4+
            ikaero+
            aitityoton+
            helsinki+
            aitiedsp+
            aika,
            scale=~vaat_yh,nominal=~vaat_vuoroas)

summary(fit)

formula:
paa ~ vakiintunut_aidille + vakiintunut_isalle +
suositusaidille + suositusisalle +
oiktodaitia1 + oiktodisaa1 +
syytisaa4 + ikaero + aitityoton + helsinki + aitiedsp + aika
scale:   ~vaat_yh
nominal: ~vaat_vuoroas

link threshold nobs logLik AIC   niter max.grad cond.H
logit flexible 387  -327.23 696.47 14(2) 3.76e-14 2.1e+05

Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z )	
vakiintunut_aidille	-2.54811	0.59596	-4.276	1.91e-05	***
vakiintunut_isalle	2.73094	0.61011	4.476	7.60e-06	***
suositusaidille	-2.40919	0.86940	-2.771	0.005587	**
suositusisalle	3.27271	0.64685	5.059	4.20e-07	***
oiktodaitia1	1.93418	0.72474	2.669	0.007613	**
oiktodisaa1	-2.34631	0.94728	-2.477	0.013253	*
syytisaa4	3.57820	1.66205	2.153	0.031327	*
ikaero	-0.11755	0.03391	-3.467	0.000526	***
aitityoton	1.15416	0.51282	2.251	0.024410	*
helsinki	1.07107	0.48579	2.205	0.027469	*
aitiedsp	1.20306	0.42868	2.806	0.005009	**
aika	0.13733	0.06555	2.095	0.036178	*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

log-scale coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
vaat_yh	2.5804	0.2032	12.7	<2e-16	***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:

	Estimate	Std. Error	z value
Äidin yks.huolt. Yhteishuolto, asum. äidillä.(Intercept)	-6.6706	1.5233	-4.379
Yhteishuolto, asum. äidillä Jako tai vuoroas..(Intercept)	2.2064	0.7598	2.904
Jako tai vuoroas. Yhteishuolto, asum. isällä.(Intercept)	2.5263	0.7690	3.285
Yhteishuolto, asum. isällä Isän yks.huolt..(Intercept)	11.7984	1.7312	6.815
Äidin yks.huolt. Yhteishuolto, asum. äidillä.vaat_vuoroas	2.0581	3.5776	0.575
Yhteishuolto, asum. äidillä Jako tai vuoroas..vaat_vuoroas	-1.7910	0.6928	-2.585
Jako tai vuoroas. Yhteishuolto, asum. isällä.vaat_vuoroas	2.0036	0.7849	2.553
Yhteishuolto, asum. isällä Isän yks.huolt..vaat_vuoroas	6.6963	12.4190	0.539

(113 observations deleted due to missingness)

```
> ennustedata <- data.frame(
  valittajaisa=valittajaisa,
  vanhlapsa7=vanhlapsa7,
  vanhinlapsipoika=vanhinlapsipoika,
  yhteishuolto=yhteishuolto,
  vakiintunut_aidille=vakiintunut_aidille,
```

```

vakiintunut_isalle=vakiintunut_isalle,
vakiintunut=vakiintunut,
suositusaidille=suositusaidille,
suositusisalle=suositusisalle,
suositus=suositus,
oiktodaitia1=oiktodaitia1,
oiktodisaa1=oiktodisaa1,
oiktodsyty1=oiktodsyty1,
syytisiaa1=syytisiaa1,
syytaitia2=syytaitia2,
syytaitia1=syytaitia1,
syytisiaa4=syytisiaa4,
ikaero=ikaero,
aitityoton=aitityoton,
helsinki=helsinki,
aitiedsp=aitiedsp,
oikeudessa=oikeudessa,
aika=aika,huoltoennen=huoltoennen,
vaat_vuoroas=vaat_vuoroas,
vaat_yh=vaat_yh)

> ennustetut_class <- predict(fit,newdata=ennustedata,type="class")$fit[!is.na(paa)]
> ennustetut_prob <- predict(fit,newdata=ennustedata,type="prob")$fit[!is.na(paa),]

# Yhteensopivuusindeksi
> library(DescTools)
> (SomersDelta(ennustetut_class,paa[!is.na(paa)]+1)/2
> (SomersDelta(ennustetut_class,paa[!is.na(paa)]+1)/2
[1] 0.8016604
# Yhteiskorrelaatiokerroin
> ridits <- (rank(paa[!is.na(paa)])-0.5)/sum(!is.na(paa))
> table(ridits)
> pred.ridit <- 0.1082995951417*ennustetut_prob[,1]+
0.364372469635628*ennustetut_prob[,2]+
0.536437246963563*ennustetut_prob[,3]+
0.69331983805668*ennustetut_prob[,4]+
0.912955465587045 *ennustetut_prob[,5]

> cor(ridits,pred.ridit,use="pairwise")
[1] 0.534463

```



## C.2 Ensimmäinen malli

```
> fit1 <- clm(paa~vanhlapsa7+vanhinlapsipoika+
              vakiintunut_aidille+vakiintunut_isalle+
              suositusaidille+suositusisalle+
              oiktodaitia1+oiktodisaa1+
              syytisaa4+
              ikaero+
              aitivityoton+
              helsinki+
              aitedsp+
              aika,
              scale=~vaat_yh,nominal=~vaat_vuoroas)
```

```
> summary(fit1)
```

```
formula:
```

```
paa ~ vanhlapsa7 + vanhinlapsipoika +
vakiintunut_aidille + vakiintunut_isalle +
suositusaidille + suositusisalle +
oiktodaitia1 + oiktodisaa1 +
syytisaa4 + ikaero + aitivityoton +
helsinki + aitedsp + aika
scale: ~vaat_yh
nominal: ~vaat_vuoroas
```

```
link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 384 -321.46 688.93 14(2) 5.37e-14 1.9e+05
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z )
vanhlapsa7	-0.70253	0.43314	-1.622	0.10481
vanhinlapsipoika	0.84083	0.43949	1.913	0.05572 .
vakiintunut_aidille	-2.90245	0.63878	-4.544	5.53e-06 ***
vakiintunut_isalle	2.65217	0.62082	4.272	1.94e-05 ***
suositusaidille	-2.54497	0.89632	-2.839	0.00452 **
suositusisalle	3.29240	0.64999	5.065	4.08e-07 ***
oiktodaitia1	2.38192	0.80464	2.960	0.00307 **
oiktodisaa1	-2.46075	0.99038	-2.485	0.01297 *
syytisaa4	4.29345	1.73868	2.469	0.01354 *
ikaero	-0.11869	0.03630	-3.270	0.00108 **

```

aitityoton      0.94830    0.53319    1.779    0.07531 .
helsinki        1.09460    0.50693    2.159    0.03083 *
aitiedsp        1.14743    0.44164    2.598    0.00937 **
aika            0.13100    0.06755    1.939    0.05248 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

log-scale coefficients:

```

      Estimate Std. Error z value Pr(>|z|)
vaat_yh  2.5879    0.2028   12.76 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Threshold coefficients:

	Estimate	Std. Error	z value
Äidin yks.huolt. Yhteishuolto, asum. äidillä.(Intercept)	-6.2028	1.7187	-3.609
Yhteishuolto, asum. äidillä Jako tai vuoroas..(Intercept)	2.9963	1.0877	2.755
Jako tai vuoroas. Yhteishuolto, asum. isällä.(Intercept)	3.3274	1.0943	3.041
Yhteishuolto, asum. isällä Isän yks.huolt..(Intercept)	12.7047	1.9379	6.556
Äidin yks.huolt. Yhteishuolto, asum. äidillä.vaat_vuoroas	2.3029	3.5870	0.642
Yhteishuolto, asum. äidillä Jako tai vuoroas..vaat_vuoroas	-1.7130	0.7232	-2.369
Jako tai vuoroas. Yhteishuolto, asum. isällä.vaat_vuoroas	2.5280	0.8729	2.896
Yhteishuolto, asum. isällä Isän yks.huolt..vaat_vuoroas	6.9967	12.4038	0.564

(116 observations deleted due to missingness)

```

> ennustetut_class1 <- predict(fit1,newdata=ennustedata,type="class")$fit[!is.na(paa)]
> ennustetut_prob1 <- predict(fit1,newdata=ennustedata,type="prob")$fit[!is.na(paa),]
> # Yhteensopivuusindeksi
> library(DescTools)
> (SomersDelta(ennustetut_class1,paa[!is.na(paa)])+1)/2
[1] 0.8024116
> # Yhteiskorrelaatiokerroin
> pred.ridit1 <- 0.1082995951417*ennustetut_prob1[,1]+
0.364372469635628*ennustetut_prob1[,2]+
0.536437246963563*ennustetut_prob1[,3]+
0.69331983805668*ennustetut_prob1[,4]+
0.912955465587045 *ennustetut_prob1[,5]
> cor(ridits,pred.ridit1,use="pairwise")
[1] 0.5423951

```

```
# Uskottavuusosamäärätestit vakiintuneen asuinpaikan, olosuhdeselvityksen suosituksen  
ja todeksi katsotun väkivalta-, päihteidenkäyttö- tai mielenterveyssyytöksen merkitykselle.
```

```
fit11 <- clm(paa~vanhlapsa7+vanhinlapsipoika+  
            I(vakiintunut_aidille-vakiintunut_isalle)+  
            suositusaidille+suositusisalle+  
            oiktodaitia1+oiktodisaa1+  
            syytisaa4+  
            ikaero+  
            aitivityoton+  
            helsinki+  
            aitiedsp+  
            aika,  
            scale=~vaat_yh,nominal=~vaat_vuoroas)
```

```
fit12 <- clm(paa~vanhlapsa7+vanhinlapsipoika+  
            vakiintunut_aidille+vakiintunut_isalle+  
            I(suositusaidille-suositusisalle)+  
            oiktodaitia1+oiktodisaa1+  
            syytisaa4+  
            ikaero+  
            aitivityoton+  
            helsinki+  
            aitiedsp+  
            aika,  
            scale=~vaat_yh,nominal=~vaat_vuoroas)
```

```
fit13 <- clm(paa~vanhlapsa7+vanhinlapsipoika+  
            vakiintunut_aidille+vakiintunut_isalle+  
            suositusaidille+suositusisalle+  
            I(oiktodaitia1-oiktodisaa1)+  
            syytisaa4+  
            ikaero+  
            aitivityoton+  
            helsinki+  
            aitiedsp+  
            aika,  
            scale=~vaat_yh,nominal=~vaat_vuoroas)
```

```
> anova(fit1, fit11)
```

```

      no.par   AIC  logLik LR.stat df Pr(>Chisq)
fit11      22 687.00 -321.50
fit1       23 688.93 -321.46  0.0696  1      0.792

```

```
> anova(fit1, fit12)
```

```

      no.par   AIC  logLik LR.stat df Pr(>Chisq)
fit12      22 687.37 -321.69
fit1       23 688.93 -321.46  0.4444  1      0.505

```

```
> anova(fit1, fit13)
```

```

      no.par   AIC  logLik LR.stat df Pr(>Chisq)
fit13      22 686.93 -321.47
fit1       23 688.93 -321.46  0.0042  1      0.9485

```

### C.3 Toinen malli

```

> fit2 <- clm(paa~huoltoennen+vakiintunut_isalle+ikaero+
              aitivityoton+syytaitia1+syytaitia2+
              syytisaa4+vakiintunut_aidille+vanhlapsa7
              ,
              scale=~vaat_yh,nominal=~vaat_vuoroas)
> summary(fit2)

```

```

formula: paa ~ huoltoennen + vakiintunut_isalle + ikaero + aitivityoton + syytaitia1 +
syytaitia2 + syytisaa4 + vakiintunut_aidille + vanhlapsa7
scale:    ~vaat_yh
nominal:  ~vaat_vuoroas

```

```

link threshold nobs logLik AIC niter max.grad cond.H
logit flexible 414 -389.51 815.01 13(2) 4.28e-08 1.9e+05

```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
huoltoennen	0.54784	0.30395	1.802	0.071478	.
vakiintunut_isalle	1.62836	0.53372	3.051	0.002281	**
ikaero	-0.11006	0.02929	-3.757	0.000172	***
aitivityoton	0.90547	0.40355	2.244	0.024847	*
syytaitia1	1.11433	0.37107	3.003	0.002673	**

```

syytaitia2          -1.29346    0.51781   -2.498 0.012492 *
syytisaa4           4.49410    1.68570    2.666 0.007676 **
vakiintunut_aidille -1.44896    0.53474   -2.710 0.006736 **
vanhlapsa7          -0.75215    0.35270   -2.133 0.032961 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

log-scale coefficients:

```

          Estimate Std. Error z value Pr(>|z|)
vaat_yh  2.3878     0.2037   11.72  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Threshold coefficients:

	Estimate	Std. Error	z value
Äidin yks.huolt. Yhteishuolto, asum. äidillä.(Intercept)	-5.4148	1.4541	-3.724
Yhteishuolto, asum. äidillä Jako tai vuoroas..(Intercept)	1.3460	0.9852	1.366
Jako tai vuoroas. Yhteishuolto, asum. isällä.(Intercept)	1.5570	0.9864	1.578
Yhteishuolto, asum. isällä Isän yks.huolt..(Intercept)	8.1420	1.4699	5.539
Äidin yks.huolt. Yhteishuolto, asum. äidillä.vaat_vuoroas	0.6845	2.1748	0.315
Yhteishuolto, asum. äidillä Jako tai vuoroas..vaat_vuoroas	-1.6436	0.6545	-2.511
Jako tai vuoroas. Yhteishuolto, asum. isällä.vaat_vuoroas	1.5720	0.6355	2.473
Yhteishuolto, asum. isällä Isän yks.huolt..vaat_vuoroas	6.4774	10.2095	0.634

(86 observations deleted due to missingness)

```

> ennustetut_class2 <- predict(fit2,newdata=ennustedata,type="class")$fit[!is.na(paa)]
> ennustetut_prob2 <- predict(fit2,newdata=ennustedata,type="prob")$fit[!is.na(paa),]
> # Yhteensopivuusindeksi
> library(DescTools)
> (SomersDelta(ennustetut_class2,paa[!is.na(paa)])+1)/2
[1] 0.7587166
> # Yhteiskorrelaatiokerroin
> pred.ridit2 <- 0.1082995951417*ennustetut_prob2[,1]+
0.364372469635628*ennustetut_prob2[,2]+
0.536437246963563*ennustetut_prob2[,3]+
0.69331983805668*ennustetut_prob2[,4]+
0.912955465587045 *ennustetut_prob2[,5]
> cor(ridits,pred.ridit2,use="pairwise")
[1] 0.4277471

```

## C.4 Kolmas malli

```
> fit3 <- clm(paa~valittajaisa+
              huoltoennen+
              syytisaal+
              ikaero,
              scale=~vaat_yh,nominal=~vaat_vuoroas)
> summary(fit3)

formula: paa ~ valittajaisa + huoltoennen + syytisaal + ikaero
scale:    ~vaat_yh
nominal:  ~vaat_vuoroas

link threshold nobs logLik AIC    niter max.grad cond.H
logit flexible 438 -369.74 765.48 12(3) 8.73e-08 5.8e+04

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
valittajaisa -4.87638    0.57700  -8.451 < 2e-16 ***
huoltoennen  0.94739    0.25165   3.765 0.000167 ***
syytisaal   -0.72044    0.42151  -1.709 0.087412 .
ikaero      -0.07126    0.03281  -2.172 0.029869 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

log-scale coefficients:
              Estimate Std. Error z value Pr(>|z|)
vaat_yh      2.2576     0.1802  12.53 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold coefficients:
              Estimate Std. Error z value
Äidin yks.huolt.|Yhteishuolto, asum. äidillä.(Intercept) -11.8961    1.4520 -8.193
Yhteishuolto, asum. äidillä|Jako tai vuoroas..(Intercept) -4.3282    0.9910 -4.367
Jako tai vuoroas.|Yhteishuolto, asum. isällä.(Intercept) -4.0075    0.9807 -4.086
Yhteishuolto, asum. isällä|Isän yks.huolt..(Intercept)    2.1053    1.2693  1.659
Äidin yks.huolt.|Yhteishuolto, asum. äidillä.vaat_vuoroas  0.6315    1.9469  0.324
Yhteishuolto, asum. äidillä|Jako tai vuoroas..vaat_vuoroas -2.9033    0.8737 -3.323
Jako tai vuoroas.|Yhteishuolto, asum. isällä.vaat_vuoroas  2.2312    0.6641  3.360
```

```
Yhteishuolto, asum. isällä|Isän yks.huolt..vaat_vuoroas      6.1330      6.3947      0.959
(62 observations deleted due to missingness)
```

```
> ennustetut_class3 <- predict(fit3,newdata=ennustedata,type="class")$fit[!is.na(paa)]
> ennustetut_prob3 <- predict(fit3,newdata=ennustedata,type="prob")$fit[!is.na(paa),]
> # Yhteensopivuusindeksi
> library(DescTools)
> (SomersDelta(ennustetut_class3,paa[!is.na(paa)])+1)/2
[1] 0.8518661
> # Yhteiskorrelaatiokerroin
> pred.redit3 <- 0.1082995951417*ennustetut_prob3[,1]+
0.364372469635628*ennustetut_prob3[,2]+
0.536437246963563*ennustetut_prob3[,3]+
0.69331983805668*ennustetut_prob3[,4]+
0.912955465587045 *ennustetut_prob3[,5]
> cor(ridits,pred.redit2,use="pairwise")
[1] 0.4277471
```