

PRO GRADU -TUTKIELMA

Ilkka Valjus

Lasten BMI- käyrien muutos vuosien 1974 ja 2001 välillä

TAMPEREEN YLIOPISTO

Informaatiotieteiden yksikkö

Tilastotiede

Maaliskuu 2012

Tampereen yliopisto

Informaatiotieteiden yksikkö

Valjus, Ilkka: Lasten BMI- käyrien muutos vuosien 1974 ja 2001 välillä

Pro gradu -tutkielma, 26 s.

Tilastotiede

Maaliskuu 2012

Tiivistelmä

Pituuden ja painon kehittyminen ovat lapsen terveydentilan herkkiä mittareita. Painon ja pituuden neliön osamäärää kutsutaan painoindeksiksi (Body Mass Index, BMI). Tässä tutkimuksessa tarkastellaan, millaisia muutoksia lasten BMI-käyrissä on tapahtunut vuoden 1974 jälkeen.

Tarkastelun kohteena on pääasiallisesti painoindeksin palauttava nousu eli painoindeksin käännepeiste (adiposity rebound). Lapsilla on syntymänjälkeisen painonnousun jälkeen aika, jolloin BMI keskimäärin laskee. Myöhemmin BMI kuitenkin alkaa kasvaa uudestaan, ja on huomattu, että ikä jolloin tämä tapahtuu, saattaa olla aikaistunut.

Lasten painoindeksikäyriä mallinnetaan tässä tutkimuksessa kuutiotasoitetuilla regressiosplineillä. Näiden estimoitujen käyrien ja niiden derivaattojen avulla voidaan löytää ennuste painoindeksin käännepeisteelle. Tarkastelemalla tämän kohdan siirtymän suuruutta eri tilanteissa voidaan päätellä, onko siinä ollut muutosta ajan kuluessa.

Lisäksi lasketaan lihaviiden 15-vuotiaiden lasten BMI-käyristä todennäköisyyksiä, että onko painoindeksin käännepesteen ajankohta varhainen, normaali vai onko sitä edes olemassa. Tutkitaan myös, mitkä ominaisuudet saattavat vaikuttaa painoindeksin käännepesteen ajankohtaan.

Asiasanat: kasvukäyrät, painoindeksi, BMI, painoindeksin käännepeiste, tasoitetut regressiosplinit.

Sisältö

1	Johdanto.....	1
2	Aineisto.....	1
3	Menetelmän teoria.....	4
3.1	Käytettävä malli.....	4
3.1.1	Estimaattien laskeminen.....	5
3.1.2	Keskineliövirheen estimointi.....	7
3.2	Bootstrap-estimointi.....	8
4	Tutkimustuloksia.....	8
4.1	Onko painoindeksin käänne­piste varhaistunut vuosien 1974 ja 2001 välillä?.....	8
4.2	Syntymäpainoindeksin vaikutus painoindeksin käänne­pisteen ajankohtaan.....	16
4.3	Mikä selittää varhaista painoindeksin käänne­pistettä.....	21
4.4	Millainen painoindeksin käänne­piste 15-vuotiaina lihavilla lapsilla on ollut.....	23
5	Päätelmät.....	26
6	Lähteet.....	27

1 Johdanto

Tässä tutkimuksessa tarkastellaan suomalaisten lasten BMI- käyrien muutosta ajan kuluessa. BMI eli painoindeksi määritetään

$$BMI = \frac{\text{paino}(kg)}{\text{pituus}(m)^2}. \quad (1.1)$$

Tutkimuksen kohteena on painoindeksikäyrissä esiintyvä painoindeksin käännepeiste. Yleensä lasten painoindeksissä on lasku ensimmäisen ja seitsemännen vuoden välillä, minkä jälkeen ne lähtevät nousuun. Tätä painoindeksin nousuun kääntymistä kutsutaan painoindeksin käännepeisteeksi. On mahdollista, että 1974 ja 2006 välillä tämä käännepeiste on varhaistunut eli siis lapset alkaisivat lihoa aikaisemmin kuin ennen.

On kuitenkin syytä tutkia, onko painoindeksin laskua kaikkein ylipainoisimmilla edes olemassa tai onko se varhaisempi ja onko täten painon nousu tapahtunut aikaisemmin. Kysyä tietenkin voi, mitkä asiat vaikuttavat varhaisen painoindeksin käännepeisteen ilmestymiseen. Olisiko esimerkiksi syntymäpainolla, -pituudella tai -painoindeksillä vaikutusta asiaan tai onko havaittavissa normaalia suurempi painonnousu ensimmäisen puolen vuoden aikana tai mihin painoluokkaan havainnot kuuluvat ensimmäisen ja toisen ikävuoden kohdalla.

Toisaalta voidaan myös tarkastella, millainen painoindeksin käännepeiste lihavilla nuorilla on ollut. Tämä voisi selittää, aiheuttavatko muutokset käännepeisteessä ongelmia paljon myöhemminkin.

2 Aineisto

Tutkimuksessa käytetään aineistoa pirkanmaalaisista lapsista vuosien 1974 ja 2006 väliltä lasten syntymästä maksimissaan viidenteentoista ikävuoteen asti. Muuttujina ovat muun muassa pituus, paino, ikä ja syntymäaika. Näistä saadaan laskettua BMI eri ikäluokissa ja voidaan tutkia painoindeksin käännepeistettä. Aineisto sisältää kokonaisuudessaan 4 305 lapsen kasvukäyrät ja sen on kerännyt Nina Vuorela väitöskirjaansa varten (Vuorela 2011).

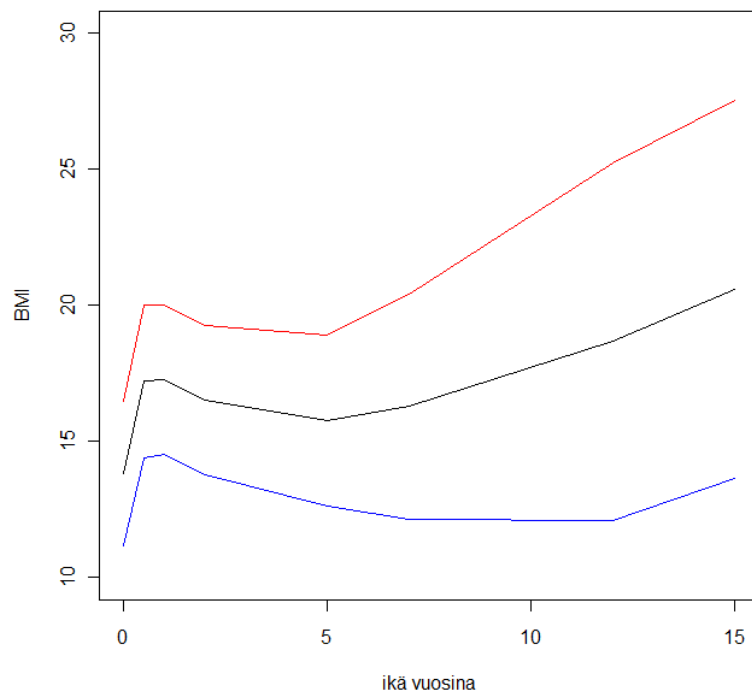
Lasten painoindeksikäyrät on jaettu yhdeksään osaan syntymävuosien perusteella (taulukko 1).

Taulukko 1. Havaintojen lukumäärät syntymävuosittain.

Vuosi	1974	1981	1990	1991	1992	1993	1994	1995	2001
Hav.lkm	1 109	987	28	553	5	5	209	643	766

Taulukon 1 perusteella on syytä yhdistää joitain vuosiluokkia, jotta saadaan mielekkäitä tutkittavia kokonaisuuksia sekä riittäviä havaintojen määriä. Vuosina 1974 ja 1981 syntyneet ovat varsin hyvin edustettuina aineistossa, mutta vuosina 1990 - 1993 syntyneet sekä vuosina 1994 ja 1995 syntyneet yhdistetään omiksi luokikseen. Lisäksi on huomattava, että aineisto sisältää dataa korkeintaan vuoteen 2006 asti, joten vuonna 2001 syntyneiden luokka sisältää kasvukäyriä vain viidenteen ikävuoteen asti. Tämän lisäksi vuosina 1994 ja 1995 syntyneistä lapsista havaintoja vain 12-vuotiaaksi asti.

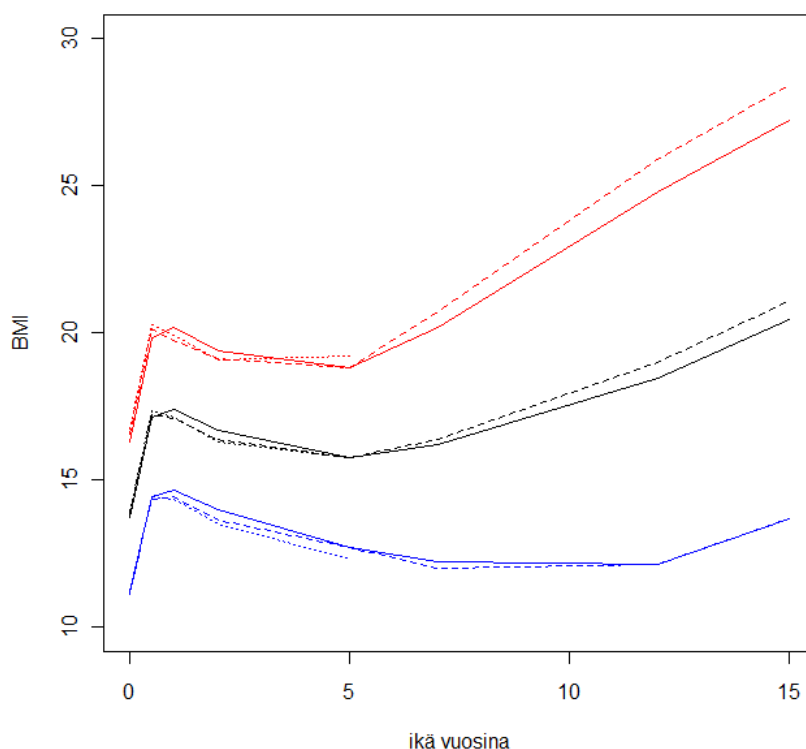
Kuvassa 1 esitellään, mille välille koko aineistoa käyttäen havainnot 95 %:n todennäköisyydellä sijoittuvat eri ikävuosipisteissä käyttämällä normaalijakauma-approksimaatiota luottamusvälien laskemiseen. Tästä voidaan tarkastella kuinka BMI-käyrät käyttäytyvät ja minne ne yleensä sijoittuvat. Hajonnat ja keskiarvot jokaisen ikäluokan kohdalla ovat siis laskettu erikseen.



Kuva 1: Koko datan lasten BMI-käyrän 95%:n luottamusväli

Kuvassa 1 näkyy varhainen syntymänjälkeinen nopea painonnousu, jota seuraa painon lasku. Tämän jälkeen on havaittavissa BMI-käyrän kääntyminen nousuun eli painoindeksin käänne piste, joka on yleensä olemassa, vaikka tässäkin aineistossa oli lapsia, joiden BMI on kasvanut aina tai jopa aina laskenut 1 ikävuoden jälkeen. Tarkastellaan, onko kyseessä olevaa laskua datan lihavimmilla lapsilla ollenkaan vai jatkuuko heidän painoindeksinsä kasvu koko ajan.

Kuvassa 2 näkyvät eri ryhmistä piirretyt BMI-käyrät. Yhtenäinen viiva on laskettu aineiston siitä ryhmästä, joka sisältää vuodet 1974-1981, katkoviiva on siitä ryhmästä, joka sisältää vuodet 1990-1995 ja viimeinen pisteviiva on vuoden 2001 perusteella.



Kuva 2: Lasten keskimääräisten BMI-käyrien 95% LV eri ryhmissä

Silmämääräisesti voidaan huomata käyrissä olevan jonkinlaista eroa, mutta onko kysymyksessä merkitsevää eroa? Aikaisemmissa tutkimuksissa on huomattu lihavuuden määrän kasvua, mutta tarkempaa tarkastelua vaatii erikseen painoindeksin käänne piste ja sen muutokset.

3 Menetelmän teoria

3.1 Käytettävä malli

Käytetään kuutiotasoitettuja regressiosplinejä mallintamaan BMI-käyriä, niin että saadaan mahdollisimman hyvä sovitus, sillä kyseessä olevat lasten BMI-käyrät eivät selvästikään ole lineaarisia.

Tämän alaluvun alkuosan esitys perustuu Simon N. Woodin kirjaan *Generalized additive models: An introduction with R*.

Tasoitusfunktiot voidaan esittää käyttämällä mallia, jossa on yksi tasoitusfunktio yhdelle kovariaatille. Tämä on muotoa

$$y_i = f(t_i) + \epsilon_i, \quad (3.1)$$

missä y_i on vastemuuttuja, t_i on kovariaatti, f tasoitusfunktio ja ϵ_i ovat samoin riippumattomia ja samoinjakautuneita $N(0, \sigma^2)$ satunnaismuuttujia.

Splinejä on useita tyyppejä, mutta periaate on samankaltainen eli määritellään paloittainen funktio solmukohtien väleille. Yksinkertaisin muoto on regressiosplini, jossa määritellään paloittain regressiosuoria. Tässä tutkimuksessa kuitenkin käytetään hieinan monimutkaisempaa menetelmää eli kuutiotasoitettuja regressiosplinejä.

Kuutiotasoitettut regressiosplinit ovat paloittain määriteltyjä kolmannen asteen polynomiregressiokäyriä solmukohtien väleillä. Solmukohtat ovat siis niitä paikkoja, joissa edellinen polynomi yhtyy seuraavan kanssa muodostaen yhtenäisen käyrän. Funktio f valitaan siis siten, että se on kolmannen asteen polynomi eli se on muotoa

$$f(t) = \beta_1 + t\beta_2 + t^2\beta_3 + t^3\beta_4, \quad (3.2)$$

jolloin kaava (3.1) on muotoa

$$y_i = \beta_1 + t_i\beta_2 + t_i^2\beta_3 + t_i^3\beta_4 + \epsilon_i. \quad (3.3)$$

Näiden polynomien kaarevuutta on mitattu niiden toisilla derivaatoilla. Näistä saadaan tasoituksen määrä laskemalla näiden neliöityjen toisen asteen derivaattojen integraaleja eli siis saadaan lopputuloksena kuutiollinen tasoitus. Mallit estimoidaan minimoimalla seuraavaa kaavaa

$$\sum_i [y_i - f(t_i)]^2 + \lambda \int [f''(t)]^2 dt, \quad (3.4)$$

jossa λ on tasoitusparametri ja $\int [f''(t)]^2 dt$ on käyrän toisen derivaatan integ-

roitu neliö, joka mittaa käyrän tasaisuutta.

Seuraavan osan kaavat ja teorian esitys perustuu J. S. Simonoffin kirjaan *Smoothing Methods in Statistics*.

Regressiossa pienimmän neliösumman kriteeri on hyvä estimoinnin tarkkuuden mita. Käyrän karkeus voidaan estimoida minimoimalla lauseke (3.4), josta saadaan

$$\sum_{i=1}^n [y_i - f(t_i)]^2 + \lambda \int f^{(m)}(t)^2 dt, \quad (3.5)$$

joka on kaavan (3.4) yleistetympi muoto.

Kaavassa (3.5) olevan funktion f m . derivaatta on $f^{(m)}$. Jotta lauseke (3.5) olisi määritelty, pitää olettaa, että f on m kertaa derivoituva ja että sen neliö on integroituva. Välillä $[a, b]$ nämä ehdot toteuttavaa funktioluokkaa merkitään $W_2^m[a, b]$.

Edellä olevassa kaavassa λ on tasoitusparametri ja sen suuruus määrää tasoituksen voimakkuuden. Jos tämä parametri on 0, on lopputulos interpolaatiosplini, joka kulkee suoraan jokaisen havaintopisteen kautta. Kun parametri lähestyy ääretöntä, alkaa estimoitu käyrä lähestyä regressiosuoraa pienimmän neliösumman mielessä. Tämä parametri kannattaakin valita siten, että se poistaa vain liian satunnaisvaihtelun.

Tässä tutkimuksessa ja yleensä m on 2 eli käytetään kaavaa (3.4).

Kyse on siis kuutiotasoitetuista splineistä, joita tutkimuksessa on käytetty painoindexikäyrien ennustamiseen. Kuutiollinen splini on funktio, joka on paloittain määritelty kuutiollinen polynomi vierekkäisten solmujen mukaan.

Tasoitusparametrin valinta voidaan suorittaa joko silmämääräisesti valitsemalla siten, että saatavat sovitteet näyttävät sopivilta. Tässä tutkimuksessa on tehty juuri näin. Toisaalta voidaan myös käyttää menetelmiä, joilla laskettaisiin hyvä tasoitusparametrin arvo, kuten ristiinvalidointi. Tämä menetelmä ei kuitenkaan juuri tässä tutkimuksessa ollut sopiva. Sovitteisiin jäi paljon liiallista satunnaisvaihtelua.

3.1.1 Estimaattien laskeminen

Tässä osiossa näytetään, kuinka kuutiotasoitettujen spliniestimaattien saadaan laskettua. Esitys perustuu Simon N. Woodin ja Nicole H. Augustin julkaisuun ”GAMs with integrated model selection using penalized regression splines and applications to environmental modelling”.

Oletetaan, että tasoitusparametri sekä solmukohdat tunnetaan. Merkitään solmukohtia $\{t_a^* : a=1, 2, \dots, j\}$. Määritellään kuutiollinen splinikanta siten, että paloittaiset

kuutiolliset polynomit yhtyvät solmukohdissa ja lisäksi niiden ensimmäinen ja toinen derivaatta ovat jatkuvia. Tämä saadaan määrittelemällä $b_a(t) = |t - t_a^*|^3$ ja lisäksi $b_{j+1}(t) = 1$ ja $b_{j+2}(t) = x$ sekä

$$f(t) = \sum_{a=1}^{j+2} \beta_a b_a(t). \quad (3.6)$$

Asetetaan vielä kertoimille rajoitteet $\sum_{a=1}^j \beta_a = 0$ ja $\sum_{a=1}^j \beta_a t_a^* = 0$. Kun $f(t)$ määritellään edellä mainitulla tavalla, on kyseessä luonnollinen splini. Määritellään nyt X seuraavasti:

$$X = \begin{bmatrix} |t_1 - t_1^*|^3 & |t_1 - t_2^*|^3 & \dots & |t_1 - t_j^*|^3 & 1 & t_1 \\ |t_2 - t_1^*|^3 & |t_2 - t_2^*|^3 & \dots & |t_2 - t_j^*|^3 & 1 & t_2 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ |t_n - t_1^*|^3 & |t_n - t_2^*|^3 & \dots & |t_n - t_j^*|^3 & 1 & t_n \end{bmatrix}. \quad (3.7)$$

Kun merkitään $y = (y_1, \dots, y_n)$ ja $\beta = (\beta_1, \dots, \beta_{j+2})$ voidaan kaavassa (3.4) lauseke $\sum_{a=1}^n [y_a - f(t_a)]^2$ kirjoittaa muodossa $\|y - X\beta\|^2$. Nyt voitaisiin jo etsiä rajoitteet mukaan ottaen pienimmän neliösumman menetelmällä parametrien β_a estimaatit, jotka minimoivat kaavan (3.4) alkuosan, mutta tässä ei olisi vielä tasoitusta mukana ollenkaan. Seuraavaksi esitetään, miten tasoituksen vaikutus saadaan lisättyä.

Kaavan (3.4) loppuosa $\lambda \int [f''(t)]^2 dt$ voidaan myös kirjoittaa matriisimuodossa.

Tarkastellaan splinifunktion f derivaattoja ja integraaleja.

$$f'(t) = \sum_{a=1}^j \beta_a b'_a(t), \quad f''(t) = \sum_{a=1}^j \beta_a b''_a(t), \quad \int f(t) dt = \sum \beta_a \int b_a(t) dt. \quad (3.8)$$

Kirjoitetaan siis

$$f''(t) = \sum_{a=1}^j \beta_a b''_a(t) = \mathbf{b}''(t)^T \beta, \quad (3.9)$$

missä $\mathbf{b}''(t)$ on kantafunktioiden pisteessä t laskettujen toisten derivaattojen vektori. Tästä edelleen saadaan seuraava neliömuoto

$$[f''(t)]^2 = \beta^T \mathbf{b}''(t) \mathbf{b}''(t)^T \beta = \beta^T K(t) \beta, \quad (3.10)$$

missä K on matriisi

$$K(t) = \begin{bmatrix} b_1''(t)^2 & b_1''(t)b_2''(t) & \dots & b_1''(t)b_j''(t) \\ b_2''(t)b_1''(t) & b_2''(t)^2 & \dots & b_2''(t)b_j''(t) \\ \vdots & \vdots & \ddots & \vdots \\ b_j''(t)b_1''(t) & b_j''(t)b_2''(t) & \dots & b_j''(t)^2 \end{bmatrix}. \quad (3.11)$$

Käyttämällä hyväksi integraalia (3.8), voidaan kirjoittaa

$$\int [f''(t)]^2 dt = \beta^T \int K(t) dt \beta = \beta^T H \beta, \quad (3.12)$$

missä H on kerroinmatriisi. Nyt kaava (3.4) voidaan kirjoittaa muodossa

$$\|y - X\beta\|^2 + \lambda \beta^T H \beta, \quad (3.13)$$

joka minimoidaan $\beta: n$ suhteen. Lauseke (3.13) voidaan kirjoittaa muotoon

$$(y - X\beta)^T (y - X\beta) + \lambda \beta^T H \beta = \beta^T (X^T X + \lambda H) \beta + 2\beta^T X^T y + y^T y, \quad (3.14)$$

josta saadaan ratkaistua derivointikeinolla $\beta: n$ estimaatti

$$\hat{\beta} = (X^T X + \lambda H)^{-1} X^T y, \quad (3.15)$$

kun λ on annettu. Niin sanottu hattumatriisi on muotoa

$$S = X (X^T X + \lambda H)^{-1} X^T \quad (3.16)$$

ja käyrän sovitteet

$$\hat{u} = S y. \quad (3.17)$$

3.1.2 Keskineliövirheen estimointi

Sovitteen virhettä voidaan estimoida matriisin S avulla. Tämä on siis niin sanottu hattumatriisi, josta voidaan estimoida keskineliövirheen suuruutta sovitetussa mallissa. Vaikuttavien parametrien määrä mallissa voidaan ajatella olevan hattumatriisin jälki eli $\text{trace}(S)$. Tästä seuraa, että keskineliövirhe on

$$MSE = v^{-1} \sum_{i=1}^n [y_i - f(t_i)]^2, \quad (3.18)$$

missä v , residuaalien vapausasteen määrä, on muotoa

$$v = \text{trace}[I - S] = n - \text{trace}(S). \quad (3.19)$$

Keskineliövirheestä voidaan laskea neliöjuuri ja saadaan $RMSE$ eli keskineliövirheen neliöjuuri. Näitä on käytetty tämän tutkimuksen mallien tulosteissa.

3.2 Bootstrap-estimointi

Tutkimuksessa on käytetty bootstrap-estimointia painoindeksin käännepesteen ajan kohdan hajonnan estimointiin. Koska käytössä on vain yksi aineisto, on muutoin vaikea laskea kyseistä estimaattia, jota kuitenkin tarvitaan käytettyjen luottamusvälien selvittämiseksi.

Merkitään alkuperäistä otosta X ja $\hat{\theta} = \hat{\theta}(X)$ otoksesta laskettavaa tuntemattoman parametrin estimaattia. Valitaan alkuperäisestä otoksesta uusia otoksia poimimalla havainnot satunnaisesti palauttaen $b = 1, 2, \dots, B$ kappaletta. Merkitään näitä otoksia X_b eli $[x_{1b}, x_{2b}, \dots, x_{nb}]$. Lasketaan tutkittava estimaatti $\hat{\theta}_b = \hat{\theta}_b(X_b)$ jokaisen bootstrap-otoksen kohdalla. Nyt $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_B$ ovat alkuperäisen otoksen estimaatin empiirinen jakauma. Esimerkiksi tässä tutkimuksessa $\hat{\theta}_b$ ovat painoindeksin käännepesteen kohtia ja näiden hajonnasta saadaan estimoitua alkuperäisen aineiston käännepesteen hajontaa.

Bootstrap-estimointi on siis yksinkertaisesti sitä, että jo olemassa olevan aineiston oletetaan tulevan jostakin jakaumasta. Kun valitaan havaitusta aineistosta satunnaisesti ”uusia” dataa ja lasketaan näistä haluttu estimaatti useita kertoja, saadaan sellaisia estimaatteja, joita on muutoin vaikea laskea.

4 Tutkimustuloksia

4.1 Onko painoindeksin käännepeste varhaistunut vuosien 1974 ja 2001 välillä?

Sovitetaan eri syntymävuosien perusteella muodostetuissa luokissa lasten painoindeksille (BMI) kuutiotasoitettuja regressiosplini -malleja. Lasketaan näistä estimaatit ja tutkitaan, minne painoindeksin käännepeste sijoittuu. Tämä käännepeste saadaan helposti tarkastelemalla estimoitujen käyrien derivaattoja ja etsimällä niistä nollakohdat.

Eri vuosiluokille muodostetut mallit estimoidaan ja estimoitujen splinikäyrien kuvaajat piirretään käyttämällä, R-paketin ”fields” sisältämää ”sreg” funktiota. Tämä funktio laskee kuutiotasoitettuja splinejä, joista voidaan laskea halutuissa pisteissä estimaatteja ja näiden derivaattoja käyttämällä ”predict” funktiota.

Mallit ovat muotoa

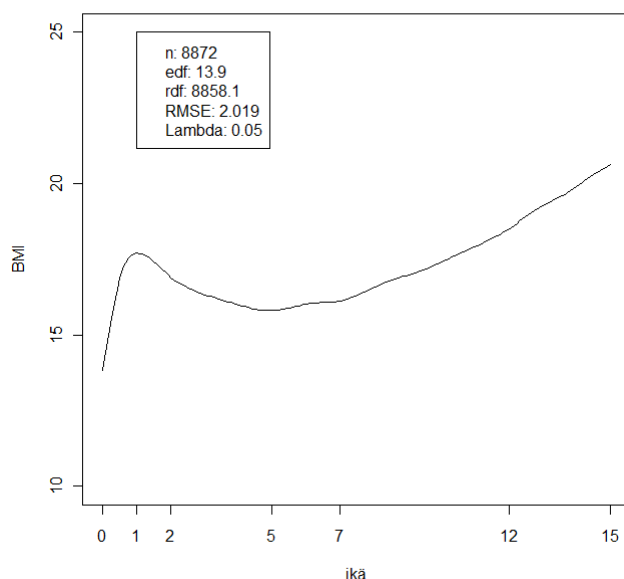
$$y_i = \mu(t_i) + \varepsilon_i, \quad i = 1, \dots, n. \quad (4.1)$$

Funktio laskee estimaatit minimoimalla kaavaa, joka on samaa muotoa kuin kaava (3.4). Funktiossa kaavaan sisältyy mahdollisuus painottaa havaintoja, mutta tässä tutkimuksessa sitä ominaisuutta ei tarvittu. Predict-funktiolle voidaan antaa parametri-
na, että halutaanko funktion estimaatit vai funktion derivaatat.

Käytettävä tasoitusparametri voidaan joko antaa tai sitten funktio laskee ristiinvali-
dointimenetelmällä sille arvon. Tässä tutkimuksessa tasoitusparametri valittiin sil-
mämääräisesti, sillä ristiinvalidointimenetelmä antoi arvon, joka oli liian pieni.

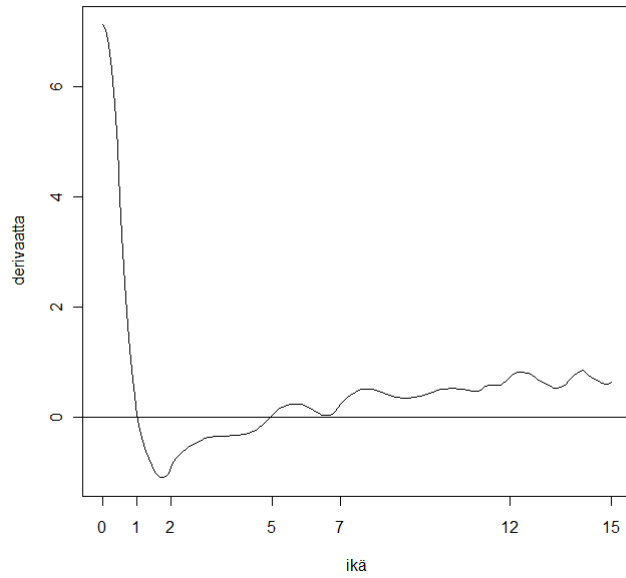
Edellä mainitut funktiot ja niiden määreet ovat ne, joita tässä tutkimuksessa on käy-
tetty, vaikka ne sisältävätkin mahdollisuuden paljon laajempaan toiminnallisuuteen.
Lisätietoja voidaan tarkastella esimerkkien kanssa paketin sivuilta [http://cran.r-
project.org/web/packages/fields/fields.pdf](http://cran.r-project.org/web/packages/fields/fields.pdf).

Kuvissa 3, 5, 7, 9 ja 11 on y-akselilla estimoidut painoindeksin arvot ja x-akselilla
lapsen ikä vuosina. Tasoitusparametriksi kaikissa kuvissa on valittu 0.05, jotta liialli-
nen satunnaisvaihtelu saadaan tasoitettua pois. Mallin antamien sovitteiden arvot on
laskettu siten, että jokainen vuosi on jaettu 12 osaan, eli siis mallin tarkkuuden voisi
ajatella olevan kuukauden tarkkuus. Lapsien iät oli ilmoitettu yhden desimaalin
tarkkuudella aineistossa, joten tarkkuus on sopiva. Kuvassa 3 on vuonna 1974 synty-
neiden lasten painoindeksikäyrä.



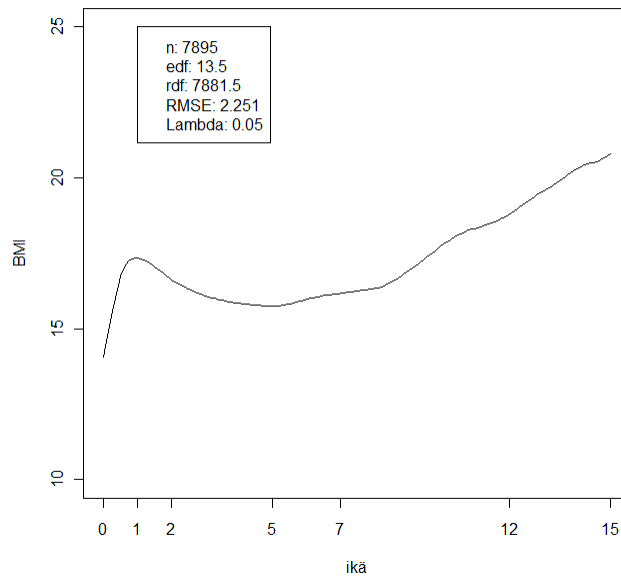
Kuva 3: Vuonna 1974 syntyneiden lasten estimoitu BMI-käyrä

Kuvassa 4 on vuonna 1974 syntyneiden lasten painoindeksikäyrän derivaatan
kuvaaja. Derivaatan nollakohta eli painoindeksin käännepiste on 4.92 vuoden iässä.

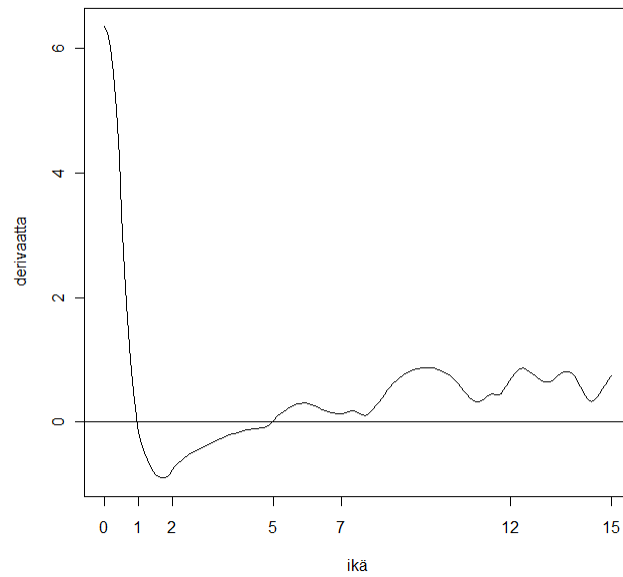


Kuva 4: Vuonna 1974 syntyneiden lasten BMI-käyrän derivaatan kuvaaja

Kuvassa 5 on vuonna 1981 syntyneiden BMI-käyrä ja sen derivaatan käyrä on kuvassa 6. Derivaatasta nähdään painoindeksin käännepisteen olevan 5 vuoden iässä. Havaitaan, että käyrät kuvissa 3 ja 5 sekä näiden derivaatat kuvissa 4 ja 6 ovat varsin samankaltaiset.

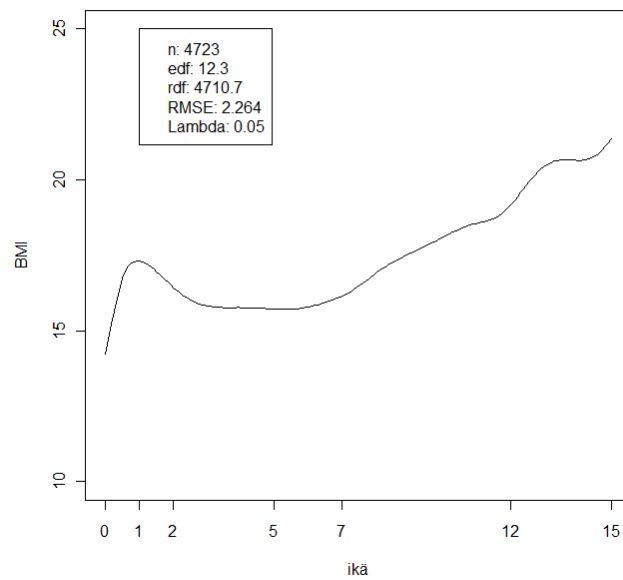


Kuva 5: Vuonna 1981 syntyneiden lasten estimoitu BMI-käyrä

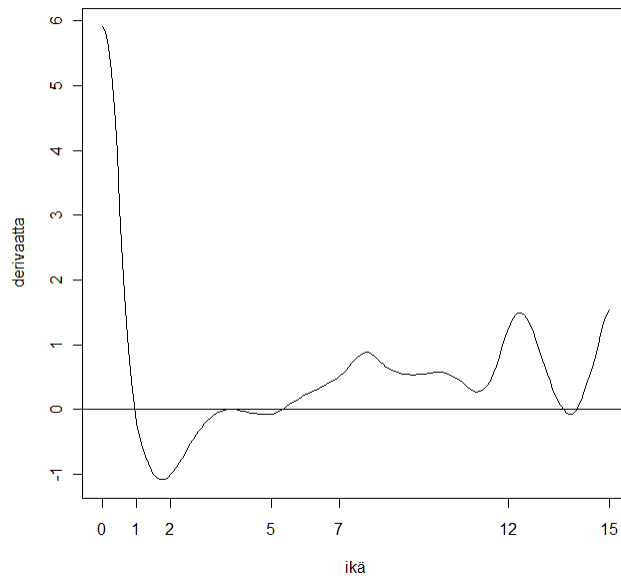


Kuva 6: Vuonna 1981 syntyneiden lasten BMI-käyrän derivaatan kuvaaja

Kuvassa 7 on vuosien 1990-1993 BMI-käyrä ja sen derivaatan käyrä on kuvassa 8. Painoindeksin käännepointe on 5.33 vuoden iässä. Derivaatan kuvaajasta huomataan, että ikävälillä 12 - 15 on hieman suurempaa vaihtelua kuin kuvissa 4 ja 6.

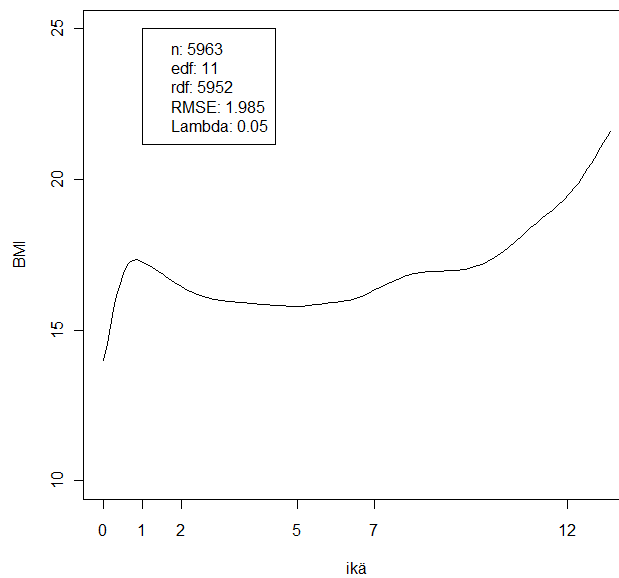


Kuva 7: Vuosina 1990-1993 syntyneiden lasten estimoitu BMI-käyrä

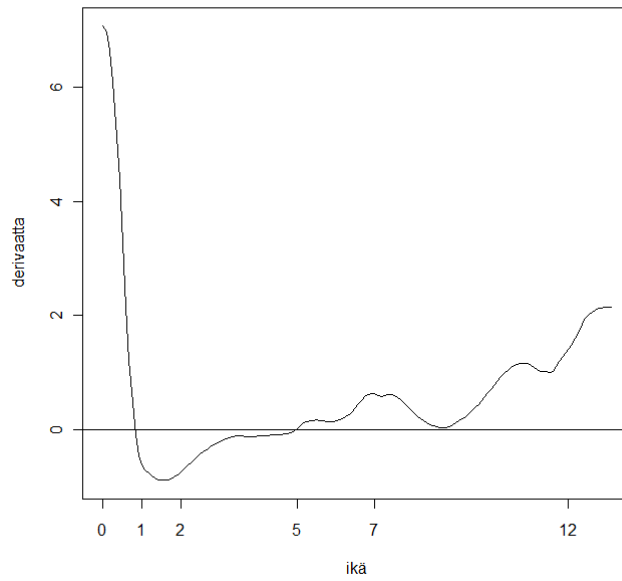


Kuva 8: Vuosina 1990-1993 syntyneiden lasten BMI-käyrän derivaatan kuvaaja

Kuvissa 9 ja 10 on vuosien 1994-1995 BMI-käyrä ja sen derivaatan käyrä. Painoindeksin käännepointe on 5 vuoden iässä. Kuten edellisessäkin kuvaajassa, BMI-käyrän lopussa näkyy voimakkaampaa muutosta. Tästä on kuitenkin vaikea sanoa, olisiko sama voimakas muutos olemassa vuoden 2001 käyrässä, sillä sen osan aineisto puuttuu.

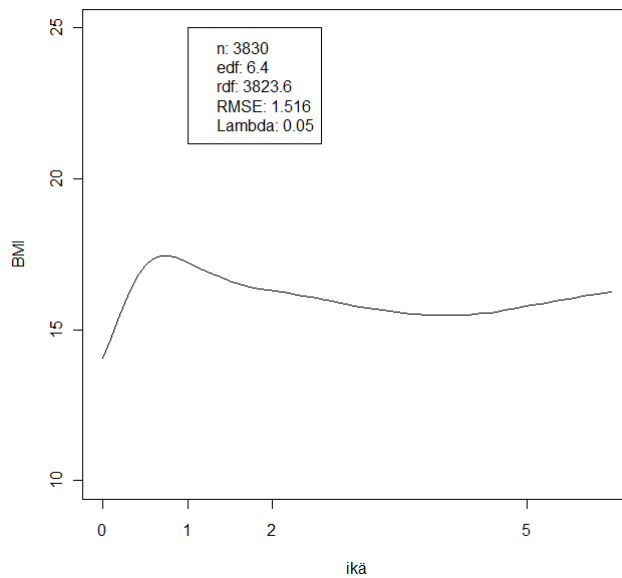


Kuva 9: Vuosina 1994-1995 syntyneiden lasten estimoitu BMI-käyrä

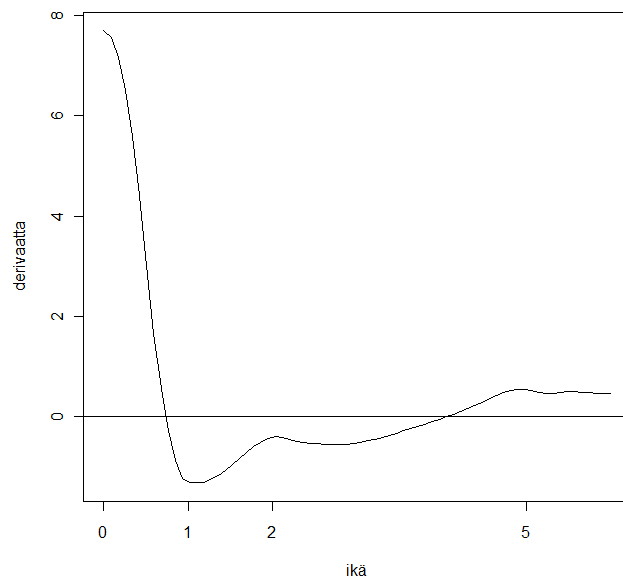


Kuva 10: Vuosina 1994-1995 syntyneiden lasten BMI-käyrän derivaatan kuvaaja

Kuvassa 11 on vuoden 2001 BMI-käyrä ja kuvassa 12 sen derivaatan käyrä. Painoindeksin käännepiste on 4.08 vuoden iässä.

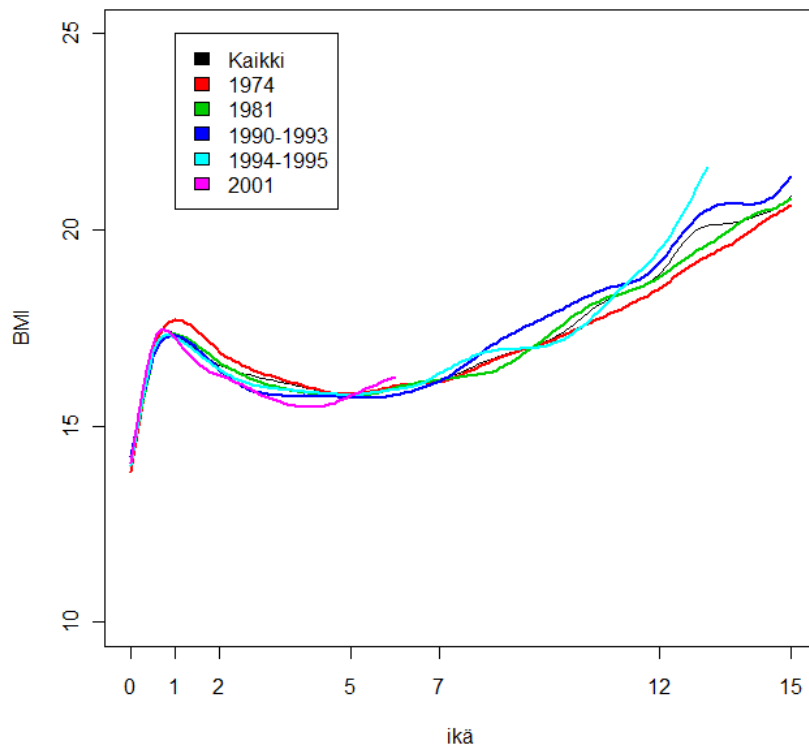


Kuva 11: Vuonna 2001 syntyneiden lasten estimoitu BMI-käyrä

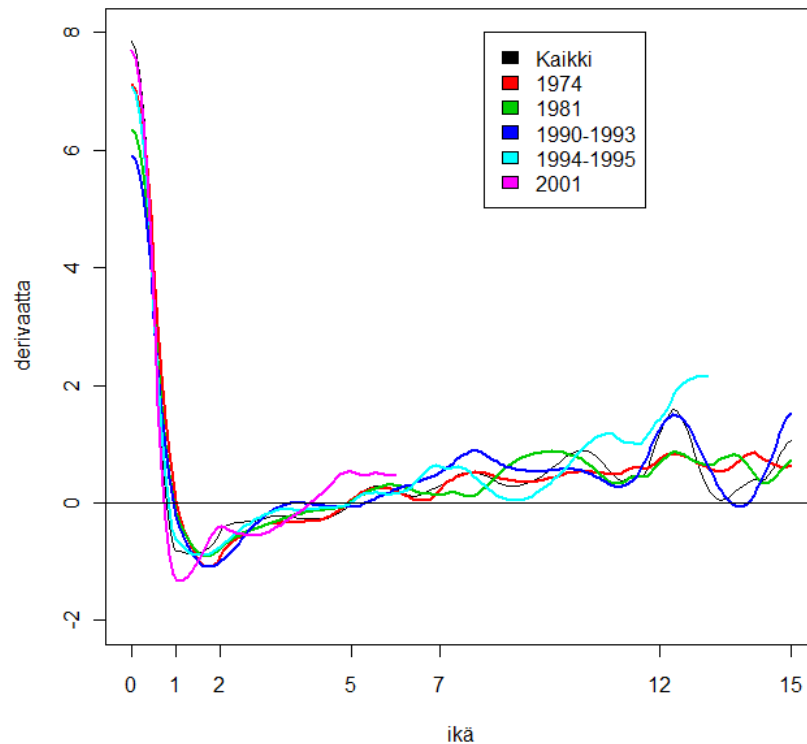


Kuva 12: Vuonna 2001 syntyneiden lasten BMI-käyrän derivaatan kuvaaja

Kuvassa 13 on esitetty kaikki estimoidut painoindexikäyrät yhdessä ja kuvassa 14 on näiden käyrien derivaatat.



Kuva 13: Estimoidut BMI-käyrät eri syntymävuosien mukaan



Kuva 14: Estimoitujen BMI-käyrien derivaattojen kuvaajat eri syntymävuosien mukaan

Nähdään, että kaikkien derivaattakäyrien nollakohdat sijoittuvat lähelle viidettä ikävuotta. Jos koko aineistosta tehdään vastaavanlainen malli, sijoittuu painoindeksin käänne piste keskimäärin tasan 5 ikävuoden kohdalle. Eri luokkien painoindeksien käänne pisteet ajoittuvat taulukon 2 mukaisesti eri vuosina syntyneiden luokissa.

Taulukko 2. Painoindeksien käänne pisteet luokittain.

Syntymävuosi	1974	1981	1990–1993	1994–1995	2001
Painoindeksin käänne piste (ikävuotta)	4.92	5	5.33	5	4.08

Se, onko muutosta tapahtunut määritellään siten, että bootstrap-estimoidaan vuonna 1974 syntyneiden painoindeksin käänne pisteelle hajonta ja keskiarvo ja tämän jälkeen lasketaan sille 95 %:n luottamusväli käyttämällä normaalijakauma-approksimaatiota. Tämän jälkeen voidaan katsoa, mitkä muiden vuosien käänne pisteet sijoittuvat tälle välille.

Tämä 95 %:n luottamusväli vuoden 1974 keskimääräiselle käänne pisteelle olisi vuosien [4.55, 5.42] välillä. Tälle välille mahtuvat kaikkien mallien estimaatit paitsi vuonna 2001 syntyneiden lasten tapaus. Tämä aineisto oli kuitenkin puutteellinen juuri halutusta kohdasta, eli 5 ikävuoden kohdalta, joten saattaa olla, että sekin todell-

lisuudessa mahtuisi tälle välille. Toisaalta, jos tarkastellaan myös edellistä luokkaa, vuosia 1994 ja 1995, saattaa olla mahdollista, että painoindeksin käänneaste aikalaisuusi siirryttäessä ajassa eteenpäin. Tässä tutkimuksessa ei kuitenkaan voida sanoa, kumpi vaihtoehto on tapahtunut todellisuudessa eli siis onko 1994-1995 luokan lopussa näkyvä voimakas nousu ja painoindeksin käänneaste aikaistuneet vuoteen 2001 mennessä vai ei.

Koska voimakas kasvu on näkyvissä kahden viimeisen luokan käyrien lopussa, on mahdollista, että kyseessä olisi estimoinnissa ilmentynyt epätarkkuus, sillä näiden molempien käyrien aineisto oli juuri tämän loppuosan kohdalla puutteellinen.

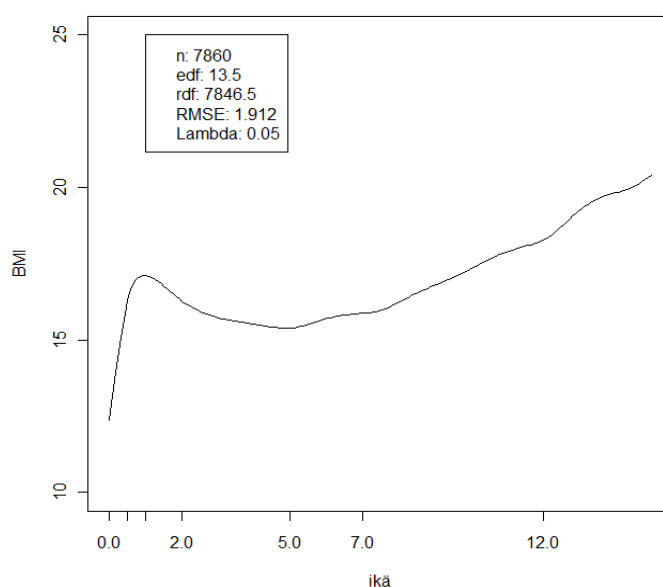
Jos kuitenkin luotetaan estimoinnin tarkkuuteen, päädytään seuraavanlaiseen tulokseen: On epätodennäköistä, että on tapahtunut merkittävä muutosta painoindeksin käänneasteen ajankohdassa vuoden 1974 jälkeen, mahdollisena poikkeuksena 2001 syntyneet.

4.2 Syntymäpainoindeksin vaikutus painoindeksin käänneasteen ajankohtaan

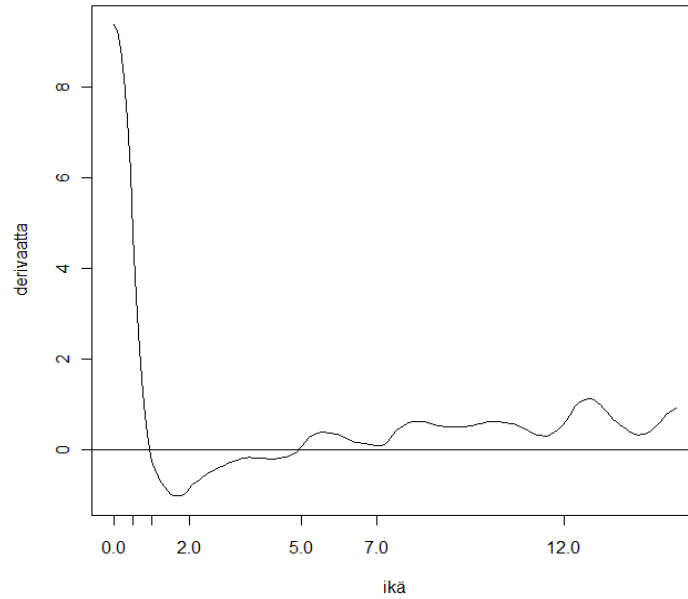
Tätä tutkimusongelmaa on lähestytty samaan tapaan kuin tutkittaessa BMI-käyrän käänneasteen varhaistumista. Nyt aineisto on jaettu neljään luokkaan syntymäpainoindeksin kvartiilien mukaan. Tämän jälkeen niiden perusteella on laskettu käyrät ja näiden derivaatat, joista lopulta saadaan ennustettu painoindeksin käänneasteen ajankohta.

Malleista saadaan seuraavanlaisia tuloksia painoindeksin käänneasteen ajankohdalle:

Alakvartiilin mukaan muodostettu luokka (kuva 15) sisältää ne lapset, joiden syntymäpainoindeksi oli 13.02 tai alle. Painoindeksin käänneaste on mallin mukaan noin 4.9 vuoden kohdalla (kuva 16).

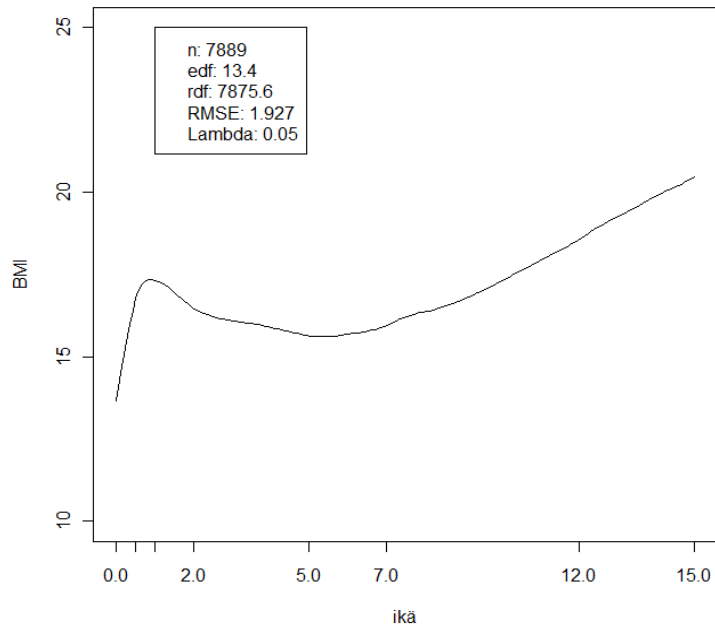


Kuva 15: Alakvartiilin estimoitu BMI-käyrä

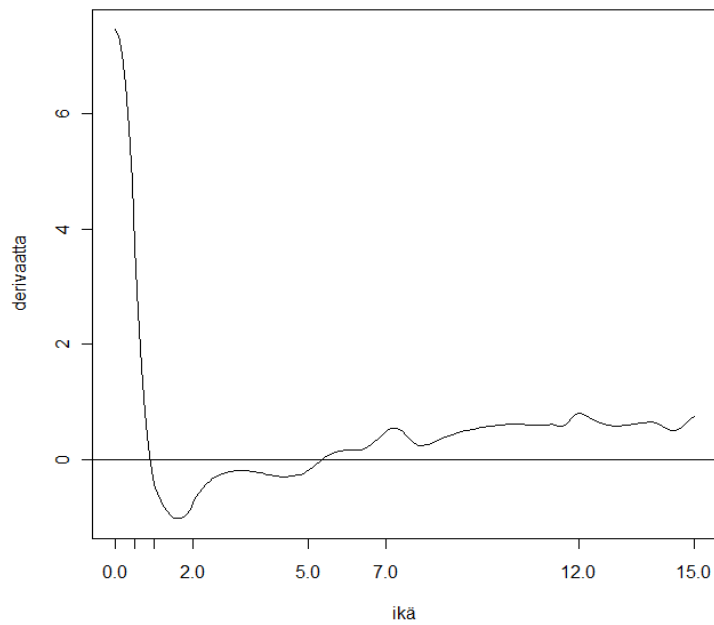


Kuva 16: Alakvartiilin estimoidun BMI-käyrän derivaatan kuvaaja

Toinen kvartiili (kuva 17), 25–50 %, sisältää ne lapset, joiden syntymäpainoindeksi oli suurempi kuin 13.02 ja pienempi tai yhtä suuri kuin 13.87. Painoindeksin käännepiste on mallin mukaan noin 5.3 ikävuoden kohdalla. (kuva 18).

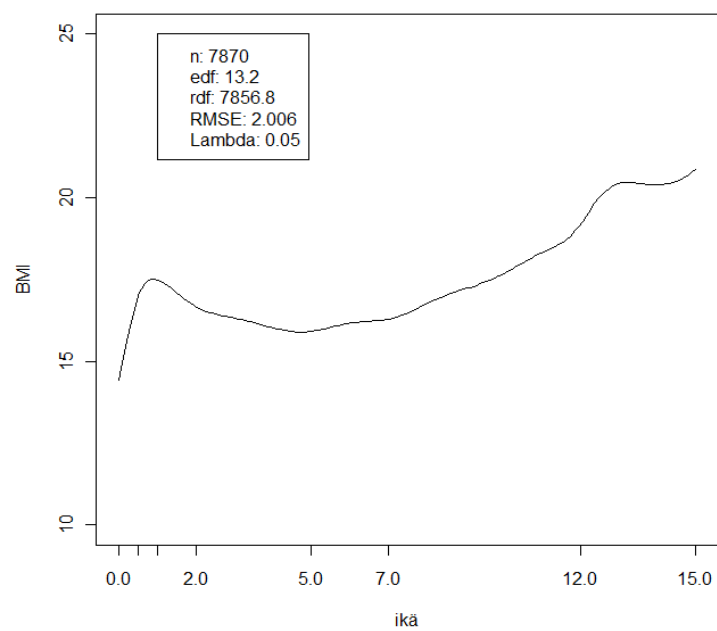


Kuva 17: 25-50 % kvartiilin estimoitu BMI-käyrä

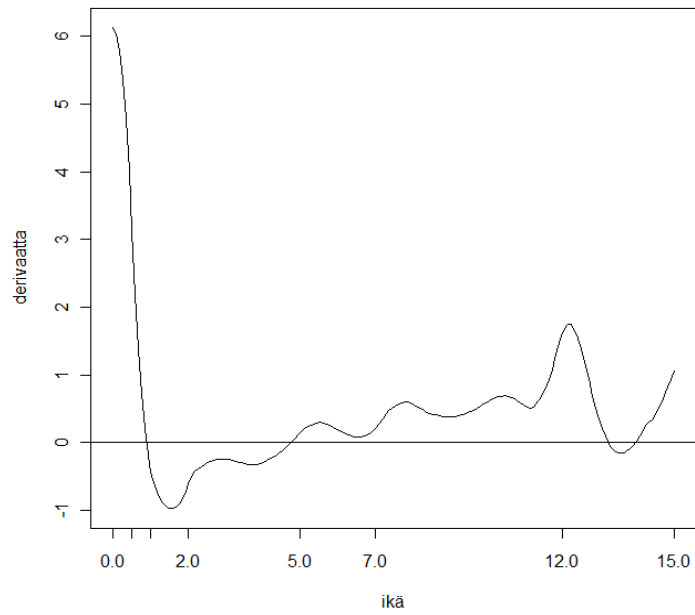


Kuva 18: 25-50 % kvartiilin estimoidun BMI-käyrän derivaatan kuvaaja

Kolmas kvartiili (kuva 19), 50–75 %, sisältää ne lapset, joiden syntymäpainoindeksi oli suurempi kuin 13.87 ja pienempi tai yhtä suuri kuin 14.7. Painoindeksin käänne-
piste on mallin mukaan noin 4.75 ikävuoden kohdalla (kuva 20).

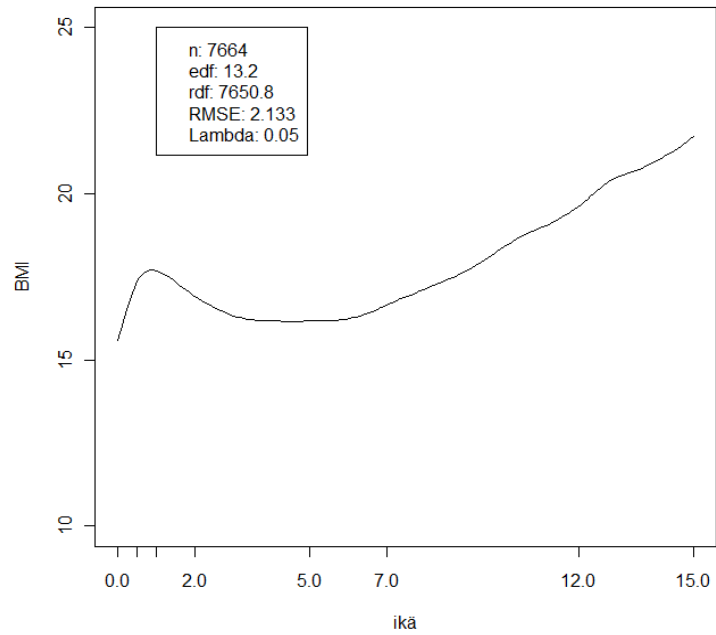


Kuva 19: 50-75 % kvartiilin estimoitu BMI-käyrä

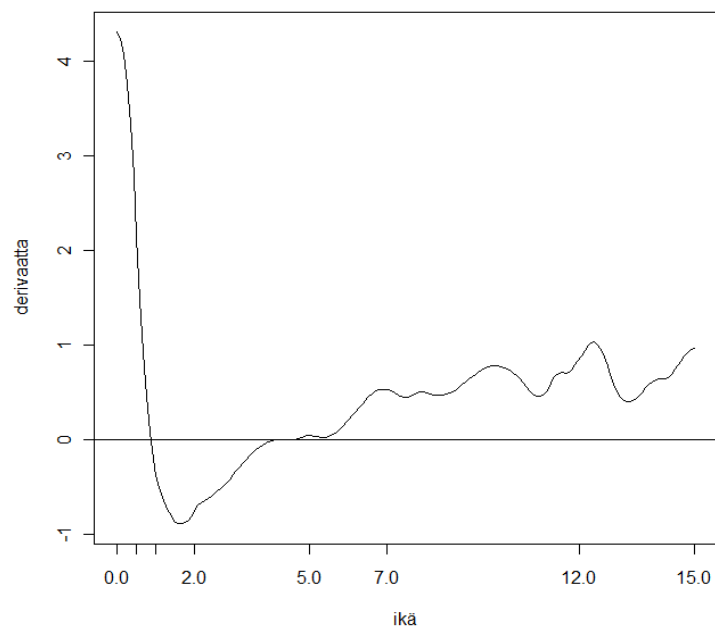


Kuva 20: 50-75 % kvartiilin estimoidun BMI-käyrän derivaatan kuvaaja

Yläkvartiilin mukaan muodostettu luokka (kuva 21) sisältää ne lapset, joiden syntymäpainoindeksi oli suurempi kuin 14.7. Painoindeksin käännepiste on mallin mukaan noin 4.66 ikävuoden kohdalla (kuva 22).

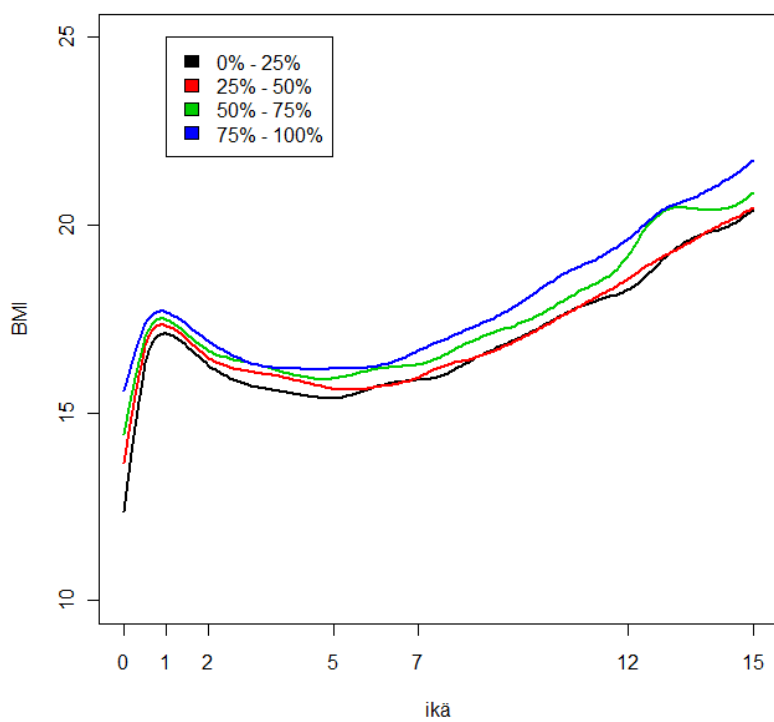


Kuva 21: Yläkvartiilin estimoitu BMI-käyrä

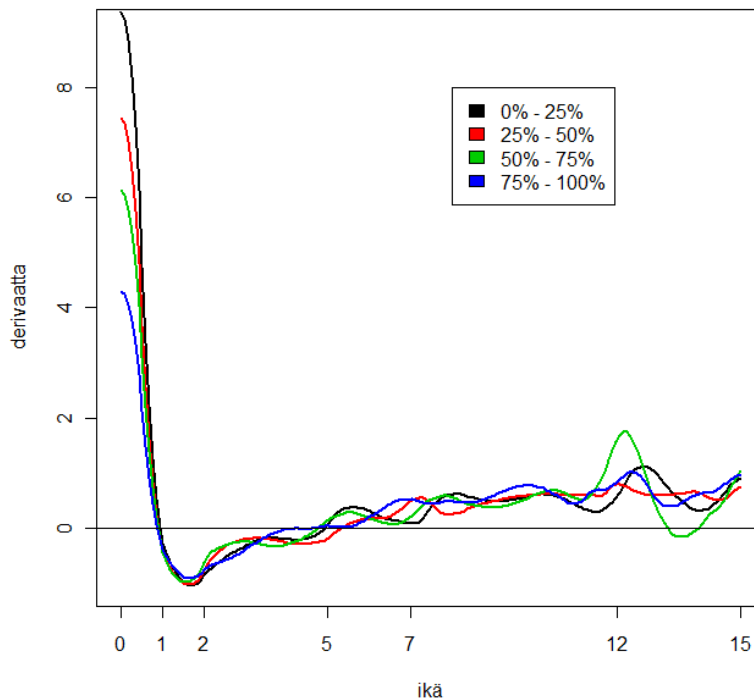


Kuva 22: Yläkvartiilin estimoidun BMI-käyrän derivaatan kuvaaja

Kuvassa 23 on esitetty kaikki estimoidut BMI-käyrät ja vastaavasti kuvassa 24 niiden derivaattojen käyrät yhdessä.



Kuva 23: Ennustetut BMI-käyrät neljän syntymäpainoindeksi kvartiilin mukaan



Kuva 24: Ennustettujen BMI-käyrien derivaatat neljän syntymäpainoindeksi kvartiilin mukaan

Kuvaajista näkyy, että kaksi alinta syntymäpainoindeksin kvartiilia ovat suunnilleen samanlaiset varsinkin loppua kohden. Kun taas tarkastellaan kahta korkeinta syntymäpainoindeksin kvartiilia huomataan, että ennusteet ovat siirtyneet korkeammalle. Lisäksi kaikkein korkein kvartiili nousut voimakkaammin kuin sitä edellinen. Tämä kertoisi siitä, että ne, joilla on jo syntyessä korkea painoindeksi, ovat myöhemminkin korkeammalla ja vieläpä lihovat voimakkaammin kuin muut.

Painoindeksi kuitenkin käyttäytyy varsin samalla tavalla jokaisessa eri kvartiilissa. Derivaattojen kuvaajat ovat alussa lähestulkoon samanlaiset. Riippumatta siitä, kuinka korkea painoindeksi on syntyessä, lasten painoindeksien käyrät käyttäytyvät aluksi samalla tavalla.

Painoindeksin käännepisteen 95 %:n luottamusväli kaikkein laihimmilla antaa välin [4.55,5.40], kun hajonta on estimoitu bootstrap-estimoinnilla. Jos tutkitaan, minne kaikkein lihavimpien malli sijoittaa arvioitun painoindeksin käännepisteen, saadaan se noin 4.66 ikävuoden kohdalle. Kaikkein lihavimpien arvioitu painoindeksin käännepisteen ajankohta siis sijoittuu kaikkein laihimpien luottamusvälin reunalle. Estimoinin epätarkkuuden takia saattaa olla mahdollista, että kaikkein lihavimpien painoindeksin käännepiste saattaa olla aikaisempi verrattuna laihempiin, mutta tässä se vielä ei olisi tilastollisesti merkitsevän poikkeava.

4.3 Mikä selittää varhaista painoindeksin käännepistettä

Tarkastellaan painoindeksin käännepistettä yksilötasolla. Määritellään painoindeksin

käännepisteen ajankohta hieman eri tavalla kuin aikaisemmin. Katsotaan missä painoindeksi on matalimmillaan ja mistä sen nousu alkaa. Tämän jälkeen yksinkertaisesti mallinnetaan tätä ajankohtaa käyttäen aivan yksinkertaista lineaarista mallia, voidaan nähdä eri muuttujien vaikutus. Tällaiset mallit kykenevät selittämään painoindeksin käännepistettä huomattavan hyvin, joten parempaa menetelmää ei tarvittu.

Tutkimuksessa erityisesti haluttiin tarkastella ainakin tiettyjä muuttujia. Näitä ehdotettuja muuttujia olivat syntymäpaino, -BMI, -pituus, varhainen nopea painonnousu 0–6 kuukauden iässä ja painoluokat ensimmäisenä ja toisena vuotena. Painoluokat on määritelty niin, että BMI < 18.5 on laiha, 18.5 ≤ BMI ≤ 25 on normaali ja BMI > 25 on lihava.

Positiiviset muuttujien kertoimet tarkoittavat, että painoindeksin käännepisteen ajankohta siirtyy myöhemmäksi ja negatiiviset taas tarkoittavat aikaistumista.

Mikään näistä ehdotetuista muuttujista ei tuntunut vaikuttavan juuri ollenkaan painoindeksin käännepisteen ajankohtaan. Mallit, joissa käytettiin vain näitä muuttujia, eivät saaneet kuin korkeintaan noin 1 %:n selitysasteita. Jos malliin lisättiin syntymävuosi, oli tällä pieni negatiivinen vaikutus. Olisi siis mahdollista, että painoindeksin käännepiste olisi aikaistunut hieman ajan kuluessa.

Taulukko 3. Ehdotettujen muuttujien malli.

	Estimaatti	Keskivirhe	T-arvo	P-arvo	
Vakio	101.27602	6.19705	16.343	< 2e-16	***
Sukupuoli	-0.214687	0.063041	-3.406	6.66E-004	***
Syntymäpaino	-0.456157	0.24219	-1.883	0.05970	.
Syntymäpituus	0.047656	0.034503	1.381	0.16728	
SyntymäBMI	0.106541	0.063665	1.673	0.09431	.
Varhainen painonnousu	0.048877	0.039603	1.234	0.21720	
Painoluokka 1v normaali	0.262552	0.093802	2.799	0.00515	**
Painoluokka 2v lihava	-1.789575	1.126769	-1.588	0.11231	
Painoluokka 2v normaali	-0.254515	0.119853	-2.124	0.03376	*
Vuosi	-0.049492	0.003016	-16.412	< 2e-16	***
Lihava 15v	-2.152082	0.125278	-17.178	< 2e-16	***
Residuaalien keskivirhe	1.943 , jonka vapausaste on 4320				
Selitysaste R^2	0.1149				
Korjattu selitysaste R^2	0.1129				
F-arvo	56.09 , jonka vapausasteet ovat 10 ja 4320, p-arvo: < 2.2e-16				

Parhaiten painoindeksin käännepisteen ajankohtaa tuntuivat selittävän käännepisteen pituus ja paino. Jos jompikumpi näistä oli mallissa, selitysasteet nousivat jo varsin huomattaviksi. Voisi siis ajatella, että painoindeksin käännepiste selittyy varsin hyvin sillä, kuinka "suuri" lapsi on. Toisaalta, jos on myöhäinen käännepiste, on lapsi ehtinyt tietenkin jo kasvaa ja näiden kahden muuttujan selitysvoima saattaa olla itsestään

selvä. Jos nämä kaksi muuttujaa olivat edellä mainittujen lisäksi mallissa, selitysaste oli jo ~92 %. Molemmilla oli positiivinen kerroin.

Taulukko 4. Ehdotettujen muuttujien, painon ja pituuden malli.

	Estimaatti	Keskivirhe	T-arvo	P-arvo	
Vakio	-0.34969	0.80669	-0.433	0.665	
Sukupuoli	-0.01064	0.01906	-0.558	0.577	
Syntymäpaino	-0.33166	0.07291	-4.549	5.53E-006	***
Syntymäpituus	-0.05172	0.01039	-4.976	6.74E-007	***
SyntymäBMI	0.07772	0.01922	4.043	5.37E-005	***
BMI	-0.19895	0.02306	-8.628	< 2e-16	***
Varhainen painonnousu	-0.26092	0.01285	-20.308	< 2e-16	***
Painoluokka 1v normaali	0.15444	0.02889	5.346	9.43E-008	***
Painoluokka 2v lihava	0.38324	0.34328	1.116	0.264	
Painoluokka 2v normaali	0.06072	0.0382	1.589	0.112	
Paino	0.13522	0.01705	7.933	2.70E-015	***
Pituus	0.08928	0.00571	15.637	< 2e-16	***

Residuaalien keskivirhe 0.5849 , jonka vapausaste on 4319.

Selitysaste R² **0.9199**

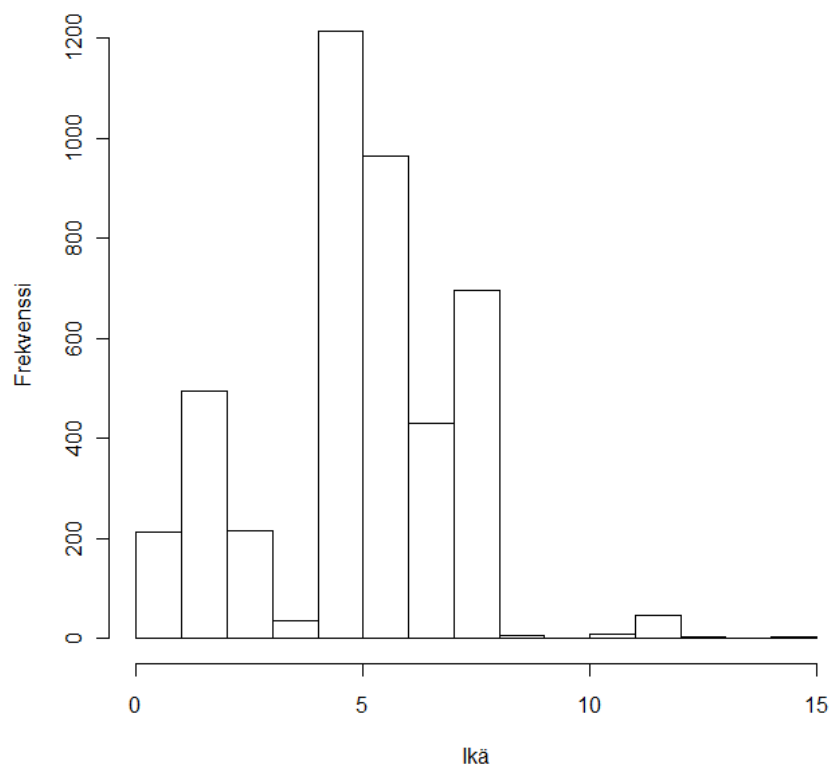
Korjattu selitysaste R² 0.9196

F-arvo 4506 , jonka vapausasteet ovat 11 ja 4319, p-arvo: < 2.2e-16

Edelliset tutkimuksen kohdat kuitenkin viittasivat, että tämä ei olisi vielä tilastollisesti merkitsevää, vaikka yllä olevassa taulukossa se näyttäisi siltä.

4.4 Millainen painoindeksin käänneaste 15-vuotiaana lihavilla lapsilla on ollut

Määritellään tavanomainen painoindeksin käänneaste siten, että lasketaan koko aineistoon perustuva painoindeksien keskiarvo ja hajonta. Katsotaan kuinka suuri osa 15-vuotiaana lihavien käänneasteista jää yhden hajonnan sisään, joka on noin [2.9,7.1]. Aikaisia ovat tällöin ne, jotka ovat aikaisempia kuin 2.9 vuotta. Kaikkien painoindeksien käänneasteiden jakauma on kuvassa 25.



Kuva 25: Painoindeksin käänne pisteiden jakauma

Huomioidaan kuvaa 25 tarkasteltaessa, että käytetyssä aineistossa oli vain tiettyinä ikävuosilta kerätyt tiedot ja siitä seuraa kuvassa oleva kuoppa. Tarkemmissa mittausajankohdissa on kuitenkin hajontaa, sillä mittauksia ei ole välttämättä suoritettu juuri valitun ikävuoden kohdalla.

Määritellään, että ei ole ollut painoindeksin käänne pistettä, jos BMI on aina noussut. Tällöin painoindeksi on ollut alimmillaan 1. ikävuoden kohdalla.

Lasketaan yksinkertaisesti tietynlaisten havaintojen määrä verrattuna kaikkien havaintojen määrään, jotta saadaan taulukoissa 5 ja 6 olevat todennäköisyydet.

Taulukko 5. Painoindeksin käänne pisteiden tyyppien yleisyys, jos lapsi 15-vuotiaana lihava.

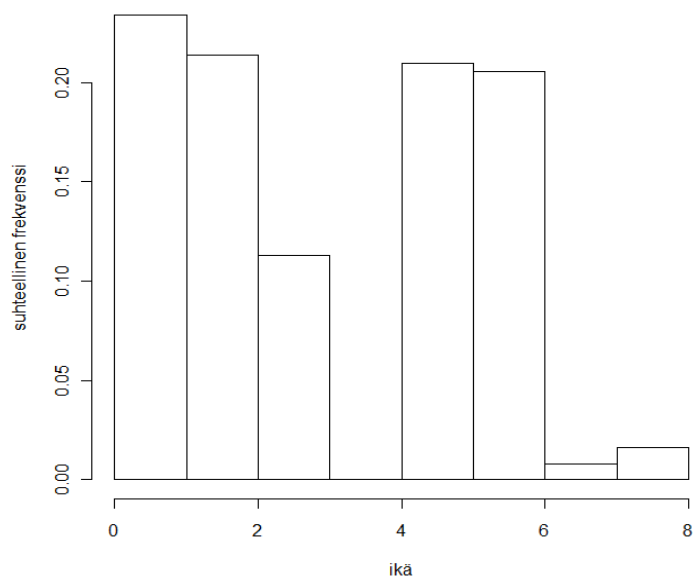
	ei painoindeksin käänne pistettä	varhainen	tavanomainen
Kaikki	0.23	0.29	0.39
Poika	0.21	0.29	0.43
Tyttö	0.25	0.30	0.35

Verrataan taulukossa 5 olevia todennäköisyyksiä koko aineistosta saatuihin todennäköisyyksiin.

Taulukko 6. Painoindeksin käänneasteiden tyyppien yleisyys koko aineistossa.

	ei painoindeksin käänneasteettä	varhainen	tavanomainen
Kaikki	0.07	0.15	0.57
Poika	0.05	0.13	0.6
Tyttö	0.1	0.17	0.54

Todennäköisyyksissä on selvä ero verrattuna 15-vuotiaana lihavien vastaaviin lukemiin. Aineistosta 15-vuotiaana lihavia lapsia on 266, joka on vain noin 6 % kaikista lapsista.



Kuva 26: Painoindeksin käänneasteen ikävuosien jakauma 15-vuotiaana lihavilla

Huomioidaan kuvassa 26 sama kuin kuvassa 25, että mittauksia on vain tietyiltä ikävuosilta ja näistä seuraa kuvassa näkyvä kuoppa. Todellisuudessa tälläkin välillä olisi havaintoja.

Kuvista 25 ja 26 näkyy samankaltaisia jakauman eroja kuin taulukot 5 ja 6 jo ehdotavatkin. 15-vuotiaana lihavilla näyttäisi olevan yleisempää aikainen painoindeksin käänneaste. Verrattuna koko aineiston painoindeksin käänneasteiden jakaumaan, jossa näkyy paljon voimakkaampi yleisyys viidennen ikävuoden kohdalla, on 15-vuotiaana lihavilla ollut useammin aikainen tai ei ollenkaan käänneasteettä. Aikainen käänneaste on näkyvillä myös, jos tarkastellaan taulukossa 3 olevaa kerrointa, joka

on negatiivinen.

5 Päätelmät

Vaikka ajassa eteenpäin mentäessä onkin havaittavissa mahdollista painoindeksin käännepisteen ajankohdan aikaistumista, kuitenkin tässä tutkimuksessa näillä menetelmillä tullaan siihen tulokseen, että se ei ole vielä tilastollisesti merkitsevää. Ainoastaan vuonna 2001 syntyneillä olikin poikkeavuutta. Näiden 2001 syntyneiden aineistossa kuitenkin havainnot loppuivat juuri sille kohdalle, jossa ilmiö oletettavasti tapahtuisi, joten tätä kohtaa tuloksista sietää ajatella varauksella.

Myöskään eri syntymäpainoindeksin mukaan luokitelluissa havainnoissa ei ollut tulosten mukaan tilastollisesti poikkeavan aikaista painoindeksin käännepistettä. Kun laihimpia verrattiin lihaviimpiin, oli lihaviimpien painoindeksin käännepiste kuitenkin aivan rajalla, että se olisikin merkitsevästi poikkeava.

Näitä molempia asioita tutkittiin käyttämällä bootstrap-estimointia tarvittavien parametrien estimointiin. Lopputulokset olivat rajoja hipovia ja päätelmissä täytyy ottaa huomioon tilastollinen inferenssi eli ovatko havainnot riittäviä.

Tutkimuksen viimeisessä osassa verrattiin 15-vuotiaina lihaviiden painoindeksin käännepisteitä kaikkien havaintojen käännepisteisiin. Havaitaan, että 15-vuotiaina lihavilla on ollut yleisempää aikainen käännepiste tai sitä ei ole edes ollut.

6 Lähteet

Fields Development Team (2006). *fields: Tools for Spatial Data*. National Center for Atmospheric Research, Boulder, CO. URL <http://www.cgd.ucar.edu/Software/Fields>

Simonoff, Jeffrey S. (1996), "Smoothing Methods in Statistics", *New York (N.Y.) : Springer*.

Vuorela Nina (2011), "Body Mass Index, Overweight and Obesity Among Children in Finland - A Retrospective Epidemiological Study in Pirkanmaa District Spanning Over Four Decades", *Tampere University Press*.

Wood, Simon N. (2006), "Generalized additive models : An introduction with R", *Boca Raton : Chapman & Hall/CRC*.

Wood, Simon N. & Augustin, Nicole H. (2002), "GAMs with integrated model selection using penalized regression splines and applications to environmental modelling", *Amsterdam: Elsevier*.