

KiEL-kuolevuuden arviointi

Kuolevuustason arviointi "pienelle" populaatiolle

Aktuaariyhdistyksen kuolevuusseminaari 26.5.2015

Risto Louhi

Kuolevuusselvityksen tausta

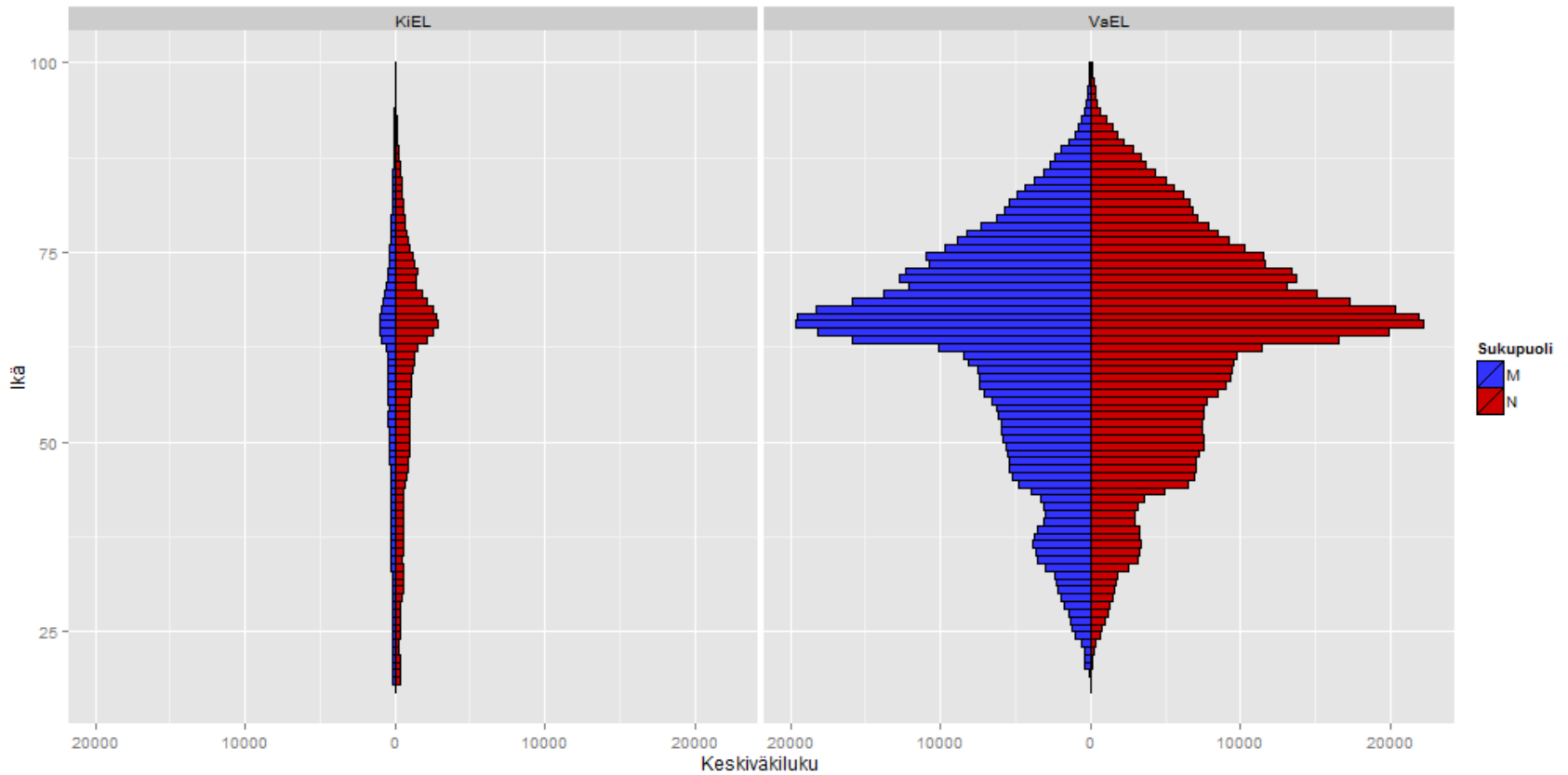
- KiEL-eläkemenoa ennustettaessa pitkän tähtäimen deterministisessä ennustemallinnuksessa käytetään joiltain osin suoraan VaEL:n siirtymätodennäköisyyksiä, näin oli aikaisemmin myös kuolevuuden osalta.
- Vuonna 2013 asiakkaan tilintarkastajan asiantuntija esitti auditoinnin yhteydessä kysymyksen siitä, tulisiko eläkemenon KiEL-mallinnuksessa käyttää mahdollisesti erillistä kuolevuutta.
- Asiakas pyysi meiltä myöhemmin selvityksen KiEL-kuolevuudesta.
- Selvitys tehtiin syksyllä 2013 ja päivitettiin lisädatalla syksyllä 2014.

Selvityksen tavoitteet, käytössä oleva data ja odotettavissa olevat ongelmat

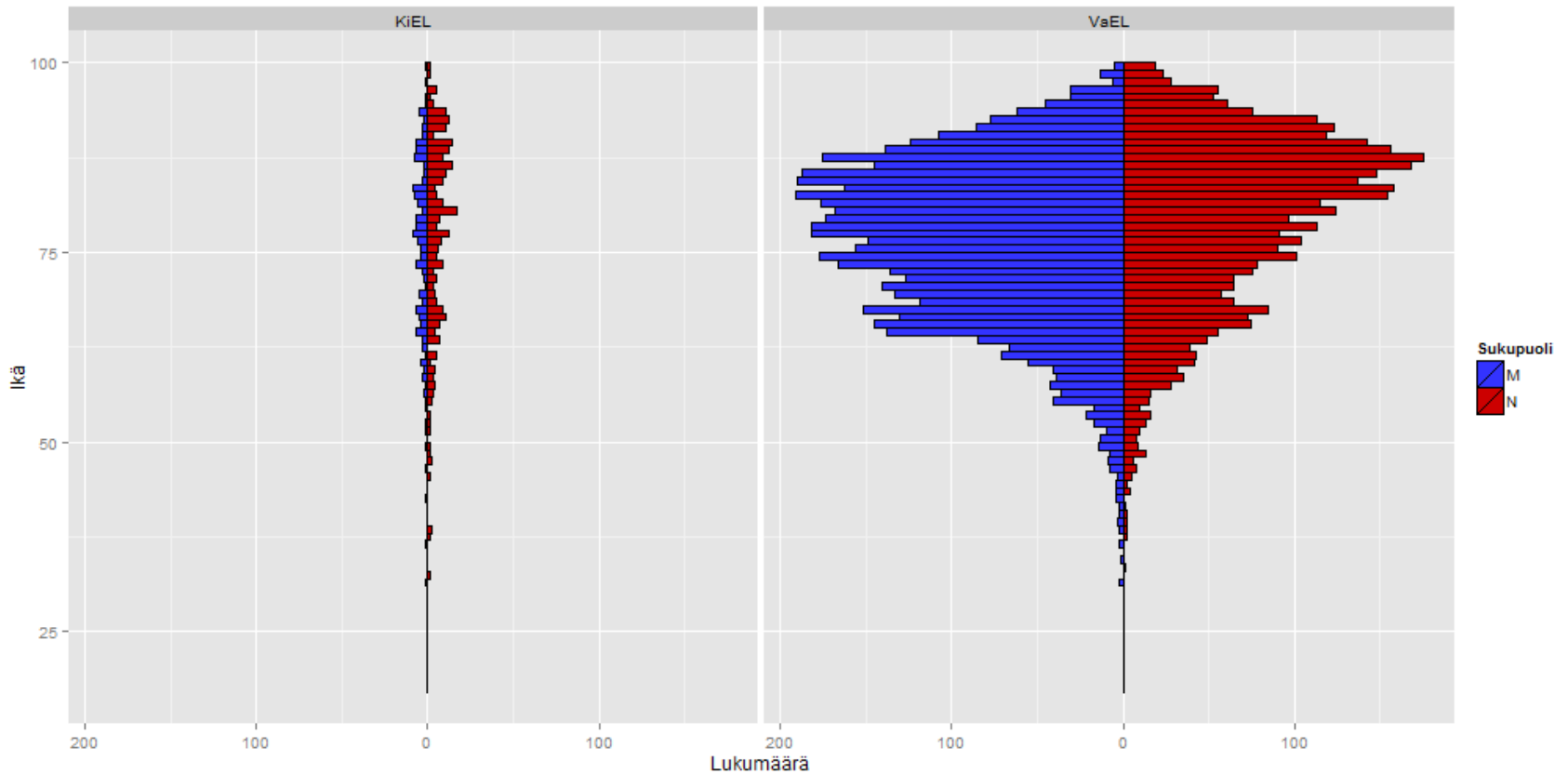
- Tässä selvityksessä tavoitteena on:
 1. Tutkia onko KiEL-eläkemenoa mallinnettaessa syytä siirtyä VaEL-kuolevuustodennäköisyyksistä erillisiin KiEL-kuolevuustodennäköisyyksiin.
 2. Mikäli tällainen siirtyminen näyttää perustellulta, niin lisätavoitteena on pyrkiä arviomaan KiEL-lukumääräkuolevuuden taso.
- KiEL:n suhteellisen pienestä koosta johtuvasta satunnaisuhojunnasta ja havaintojen puutteesta on odotettavissa ongelmia.
- Lähtötietoina on ETK:n toimittamat lakikohtaiset ikäluokittaiset kuolleet ja keskiväkiluvut

Havainnot KiEL-kuolevuudesta

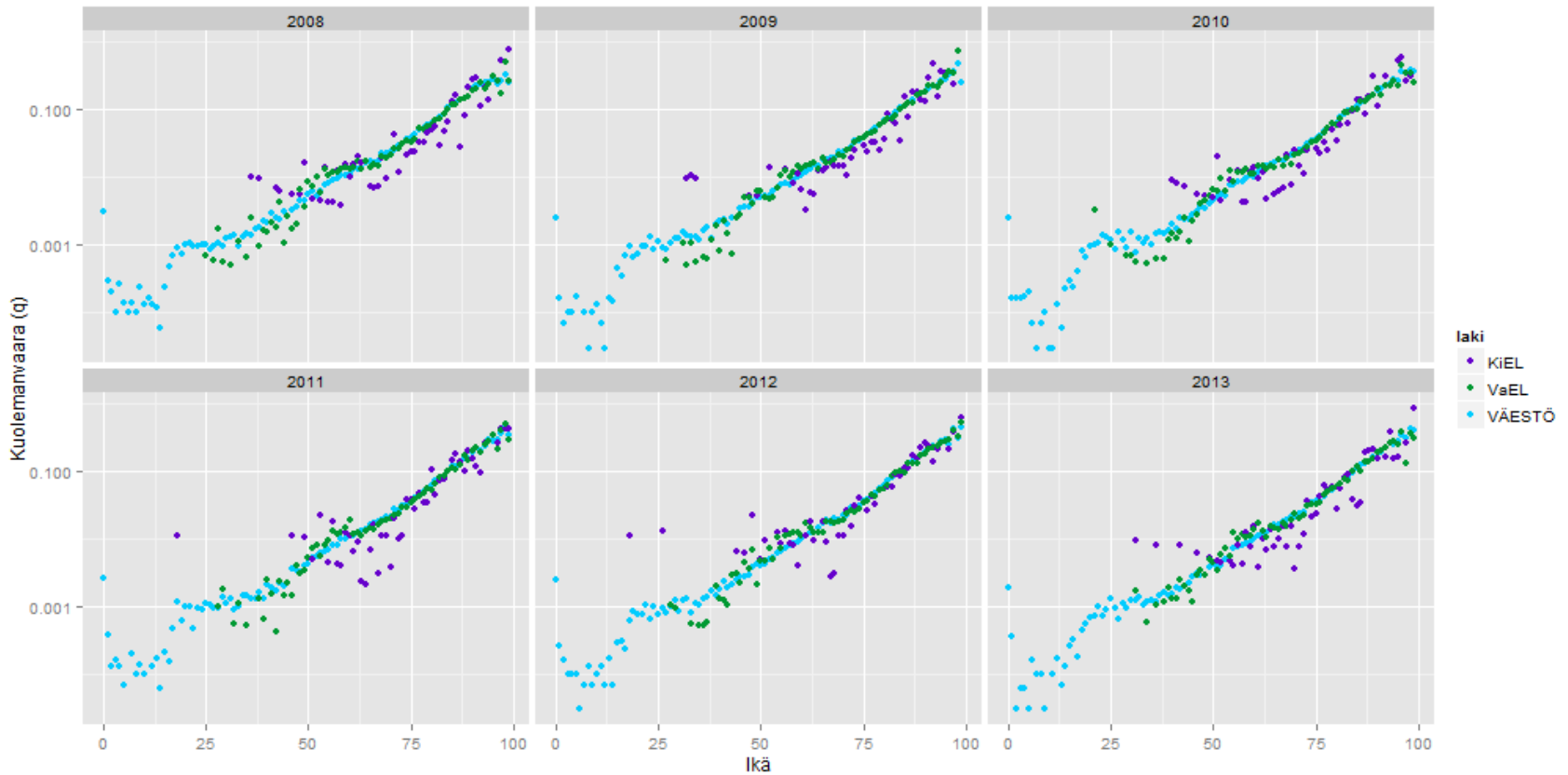
Vuoden 2013 keskiväkiluku



Vuonna 2013 kuolleet



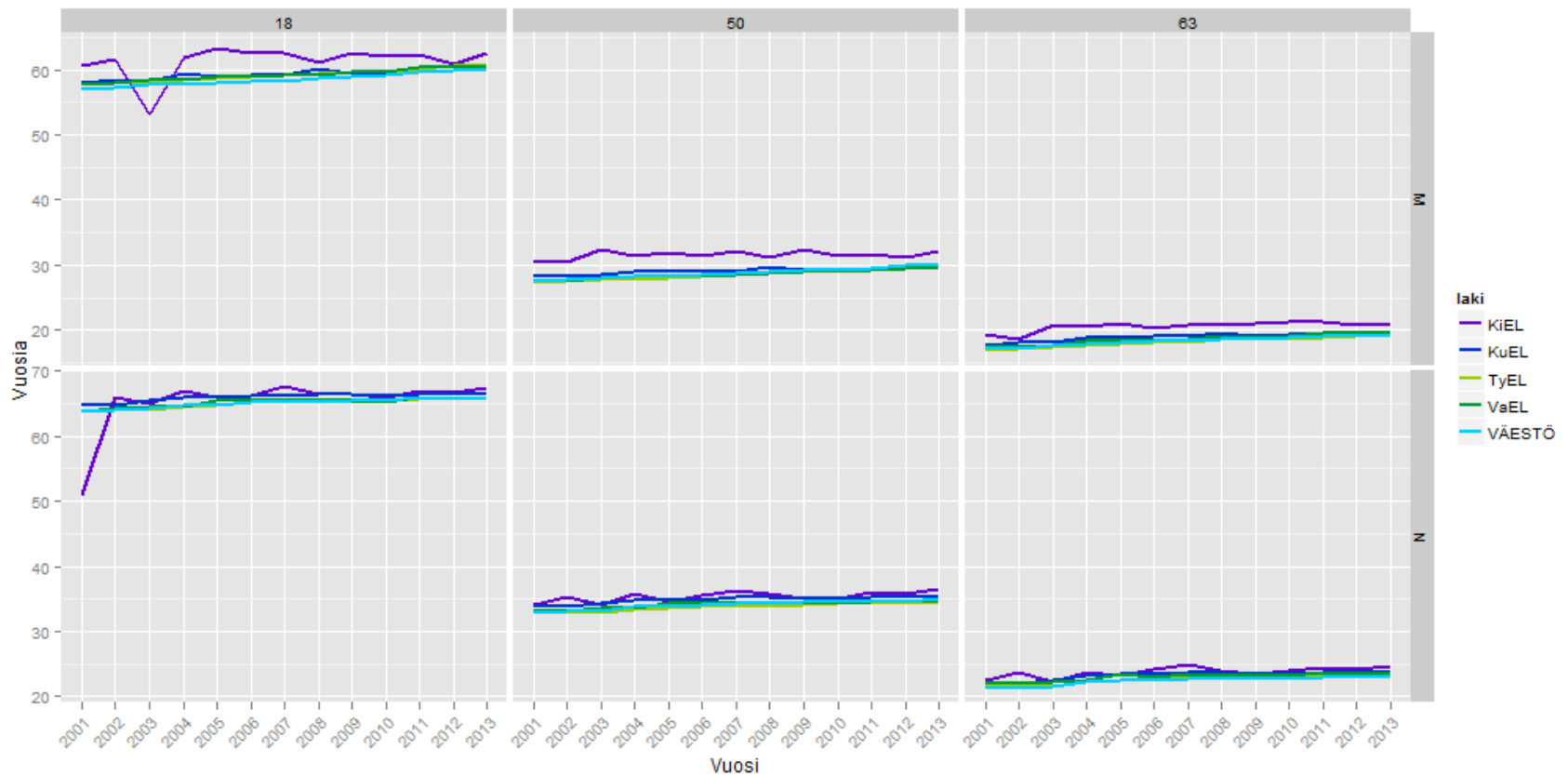
Havaitut kuolemanvaaraluvut 2008-2013, miehet



Havaitut kuolemanvaaraluvut 2008-2013, naiset



Havaittu elinajanodote 18, 50 ja 63-vuotiaina



Epävarmuus havaintojen merkityksestä

Epävarmuus epävarmuudesta

- Analyysia varten pitää kvantifioida myös havaintoihin liittyvä epävarmuus.
- Hyvin nuorten ja erittäin vanhojen ikien havainnoissa on erittäin paljon pienestä otoskoosta johtuvaa satunnaisvaihtelua. Tätä halutaan hiukan karsia, mutta kuitenkin niin, että niiden ikäluokkien osalta, joissa otoskoko on suurempi painavat varsinaiset havainnot riittävästi.
- Välineeksi Baysiläinen data-analyysi.

Epävarmuuden mallintaminen

- Melko luonnollinen mallinnus on, että sukupuolittain s , ikäluokittain x ja vuosittain v on olemassa jokin "todellinen" keskikuolevuus $\theta_{s,x,v}$ ja saadut havainnot $d_{s,x,v}$ kuolleiden määrästä seuraavat binomijakaumaa:

$$d_{s,x,v} \sim \text{Bin}(n_{s,x,v}, \theta_{s,x,v}),$$

missä $n_{s,x,v}$ on altistuneiden lukumäärä (\approx keskiikäkiluku).

- Sovelletaan Bayesin kaavaa

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

Etukäteistiedon huomiointi: priorijakauma

- Erityisesti

$$P(A|B) \sim P(B|A)P(A)$$

- Sovellettuna tilanteeseemme:

$$P(\theta_{s,x,v} = z | d_{s,x,v}) \sim P(d_{s,x,v} | \theta_{s,x,v} = z) P(\theta_{s,x,v} = z)$$

- Tällä kaavalla voimme laskea (suhteelliset) todennäköisyydet "todelliselle" keskikuolevuudelle $\theta_{s,x,v}$. Tarvitaan vain jokin järkevä priorijakauma $P(\theta_{s,x,v} = z)$.
- Luonnollinen valinta eläkelakikohtaiseksi keskikuolevuuden prioriksi (=paras arvaus ilman mitään tutkittavaan lakiin liittyviä havaintoja) ovat väestön keskikuolevuusluvut.

Priorista posterioriin

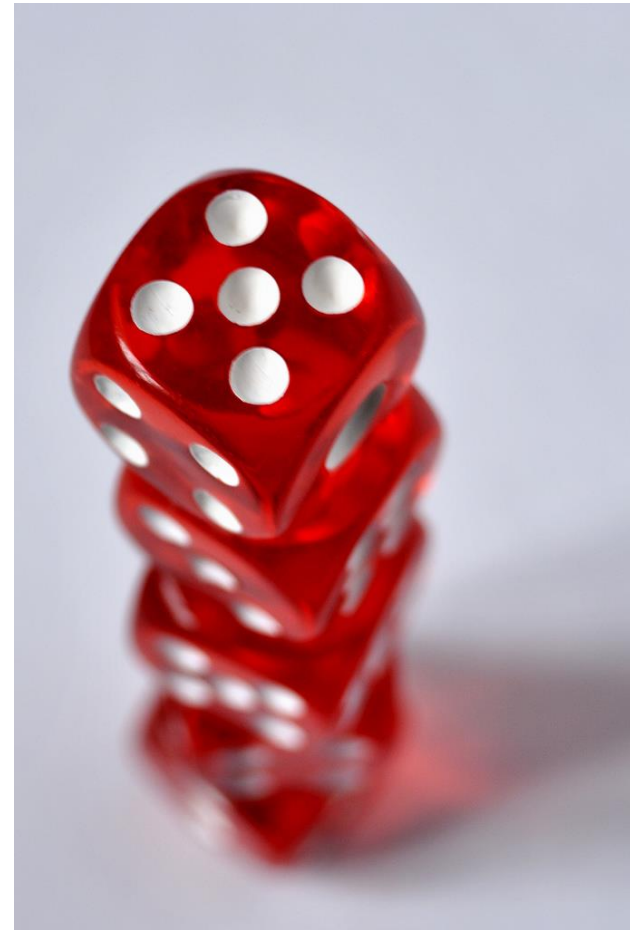
- Valitaan siis prioriksi kunkin vuoden väestödataa myötäilevä oletus, mutta tarpeeksi pienellä informaatiopainolla (piori otoskoko 25) jotta väestötasoa kohti tasoittava vaikutus on merkittävää vain erittäin pienillä ikäluokilla
- Näillä oletuksilla osoittautuu, että taustalla oleva "todellinen" keskikuolevuus on beta-jakautunut:

$$\theta_{s,x,v} \sim \text{Beta}(d_{s,x,v} + \alpha, (n_{s,x,v} - d_{s,x,v}) + \beta),$$

jossa $\alpha + \beta = 25$ ja $\frac{\alpha}{\alpha + \beta} = m_{\text{väestö}}$.

Alea jacta est!

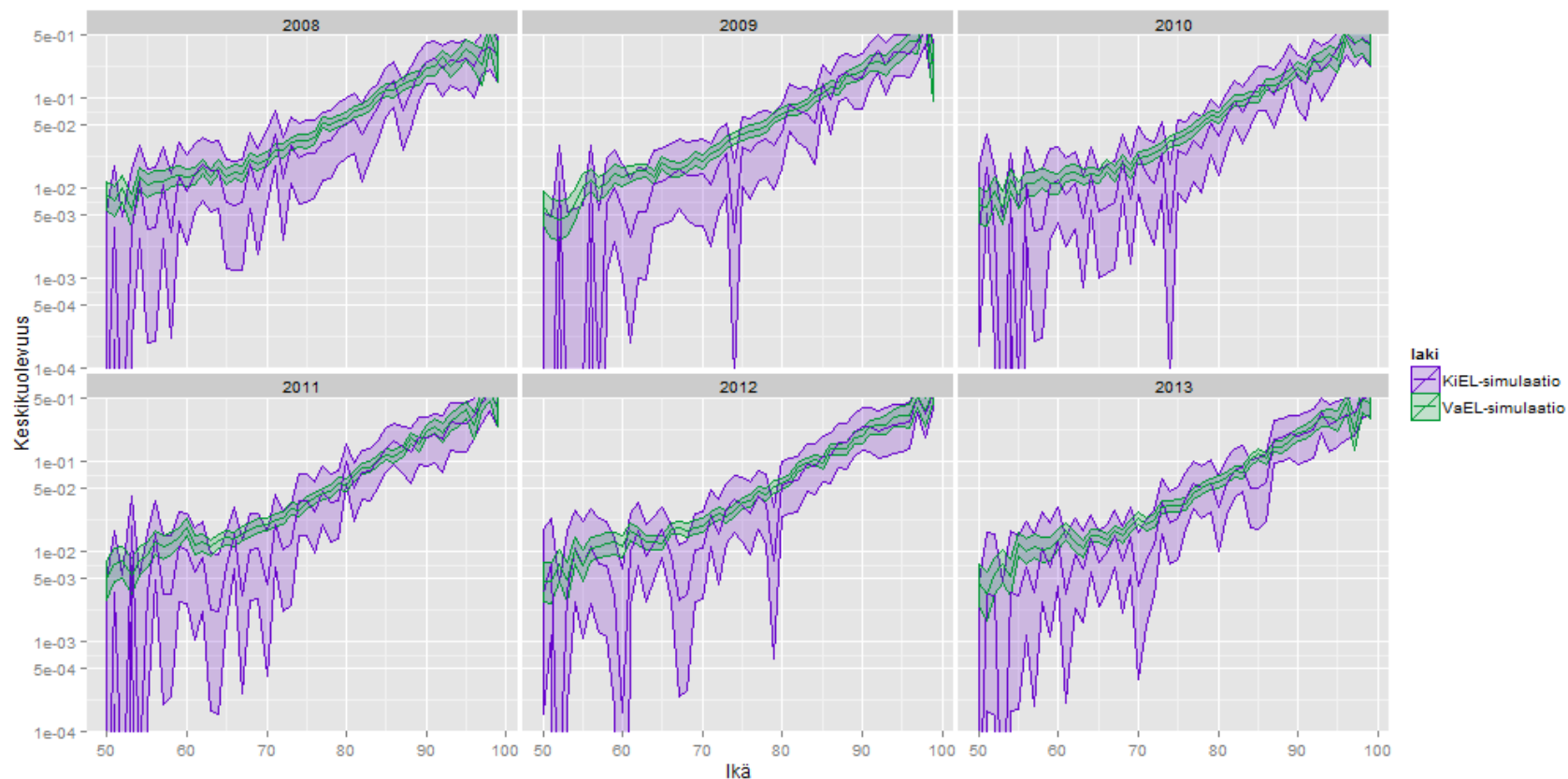
- Simuloidaan otokset edellä johdetusta kuolemanvaaralukujen *posteriori-jakaumasta* ikäluokittain Monte Carlo-menetelmällä, eli annetaan koneen heittää arpaa!
- Valitaan $n=10000$.
- Joudumme siis generoimaan (lakien lkm x sukupuoli x ikäluokat x vuodet x n) ≈ 40 miljoonaa satunnaislukua.
- Näiden avulla voimme edelleen laskea jakaumat elinajanodotteille jne.



Tietoa epävarmuudesta

Bayesiläinen analyysi KiEL-kuolevuudesta ja sen eroista
VaEL-kuolevuuteen

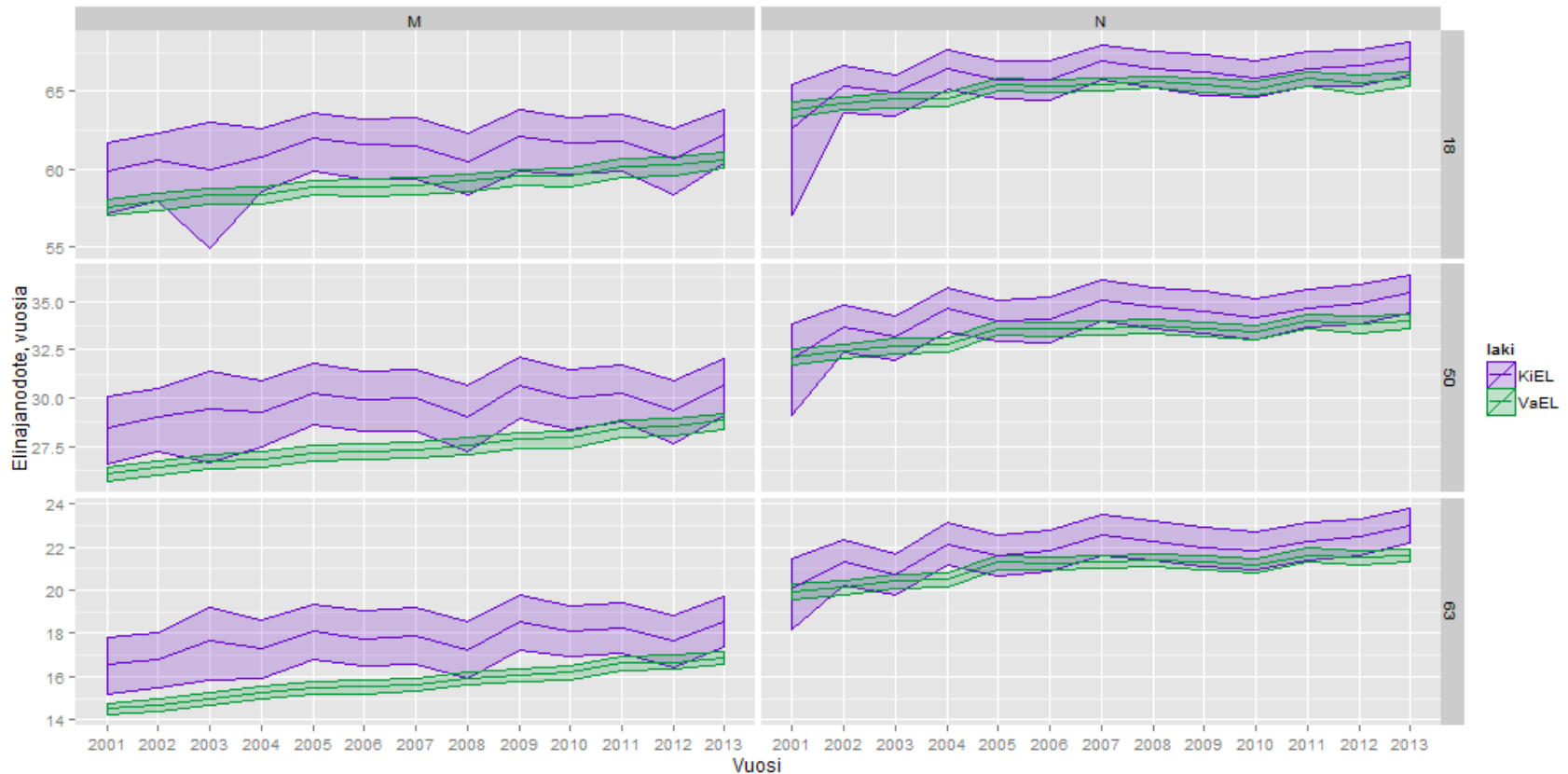
Simuloidut keskikuolevuudet 2008–2013, mediaani ja 95%-luottamusväli, 50-99 vuotiaat miehet



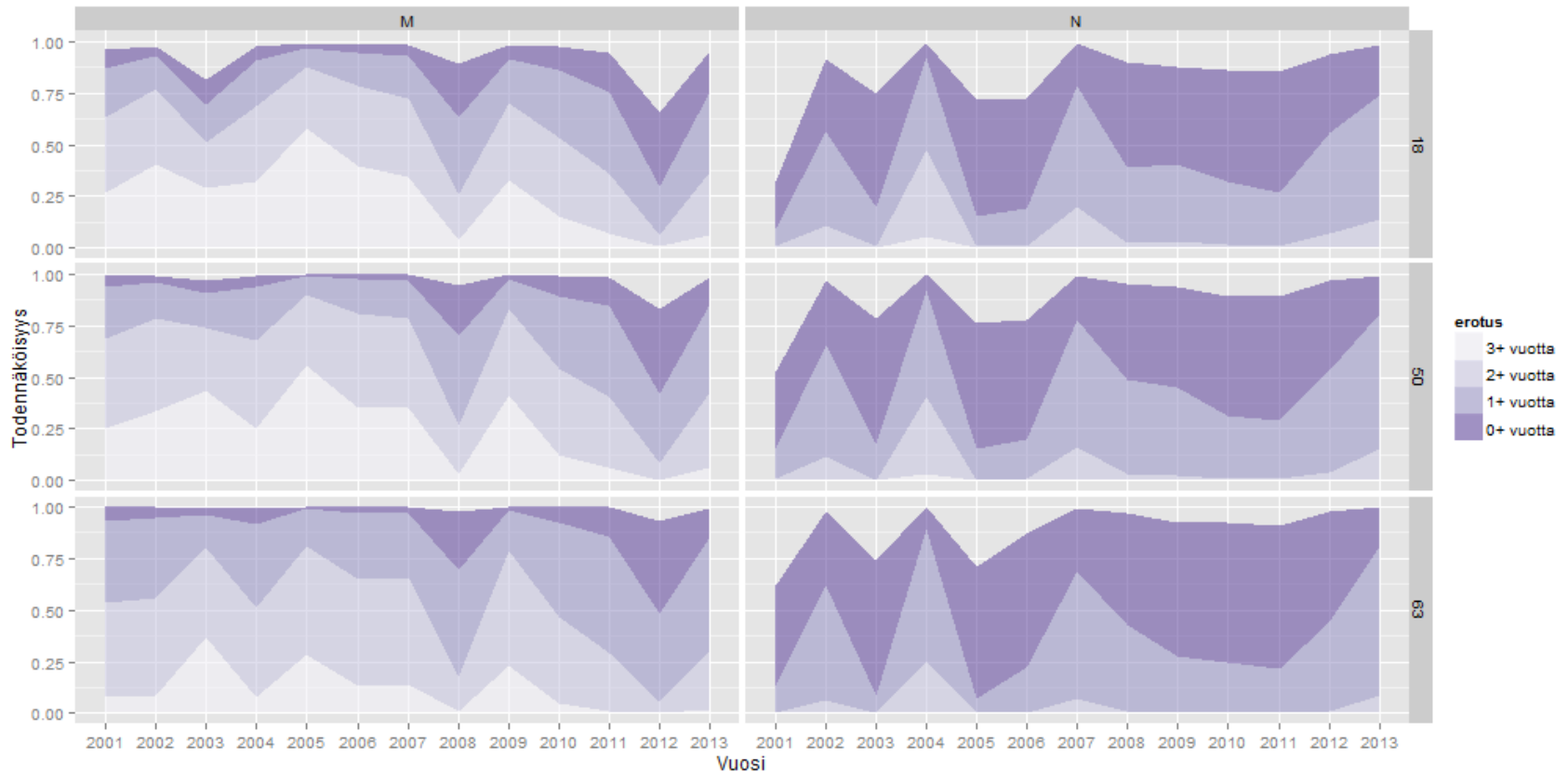
Simuloidut keskikuolevuudet 2008–2013, mediaani ja 95%-luottamusväli, 50-99 vuotiaat naiset



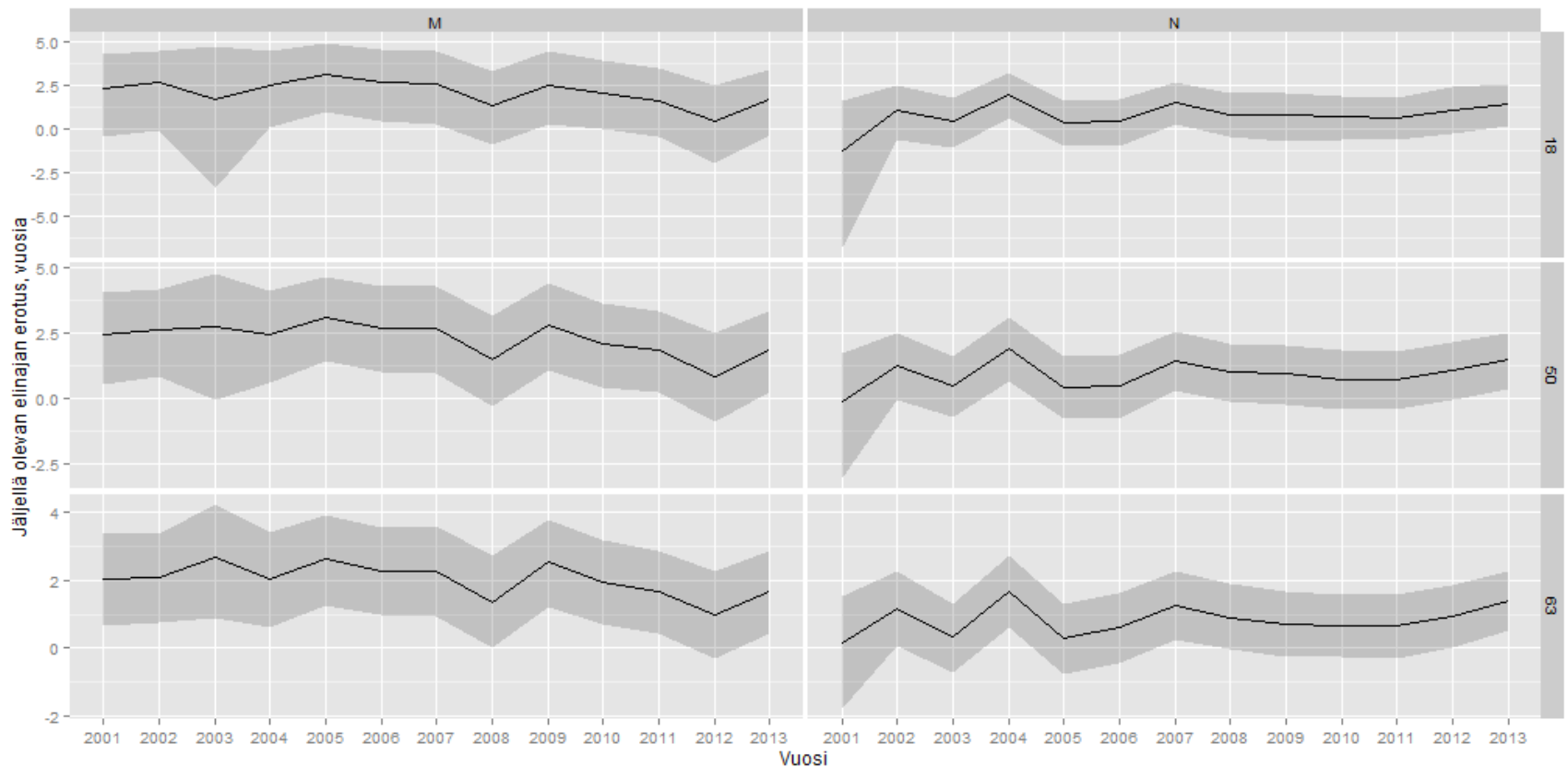
Simuloituihin kuolemanvaaralukuihin perustuva teorettinen elinajanodote (18-, 50- ja 63-vuotiaana): odotusarvo ja 95%-luottamusväli



Todennäköisyys että jäljellä oleva teoreettinen elinaika on KiEL-piirissä suurempi kuin VaEL-piirissä 18-, 50-, ja 63-vuotiaana



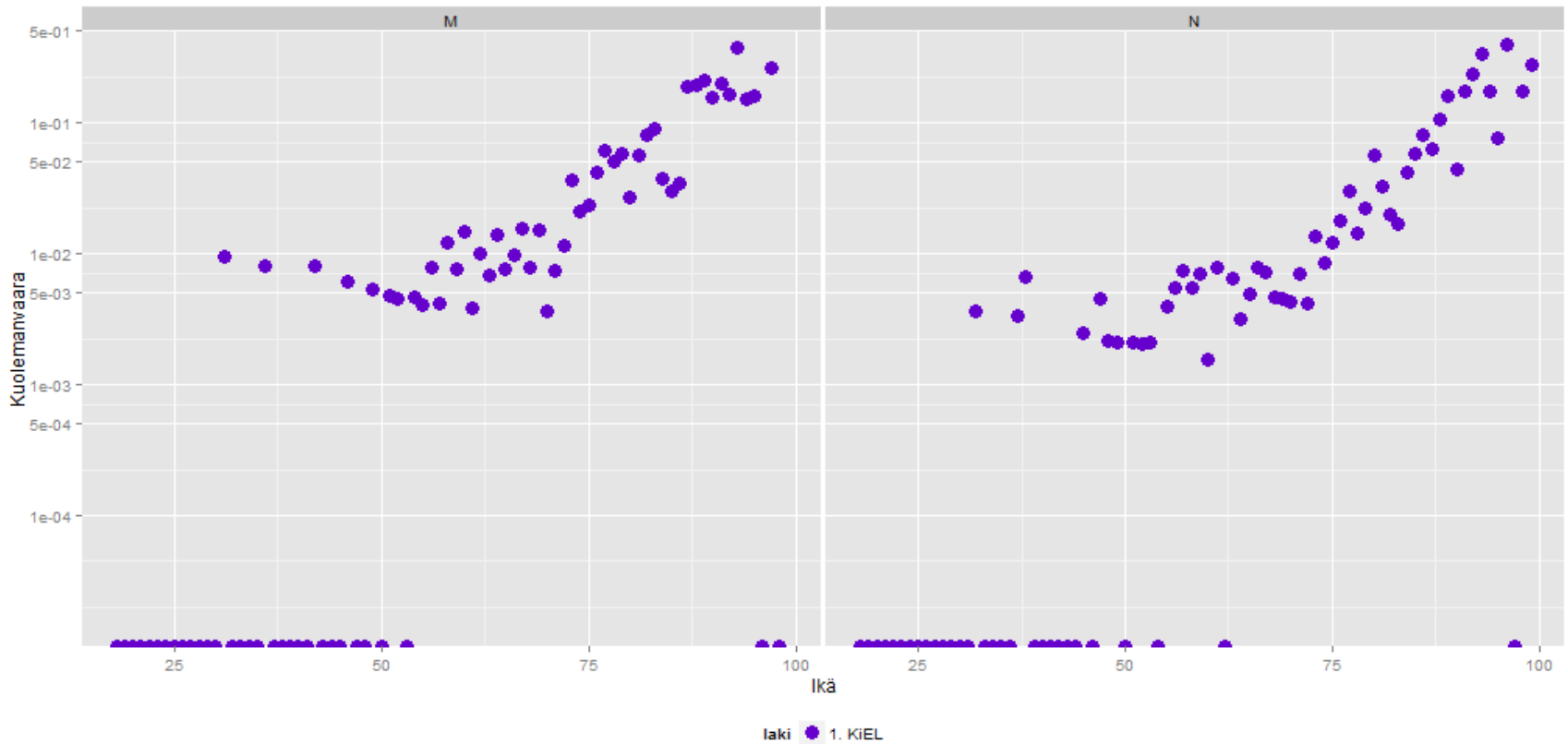
KiEL vs. VaEL, jäljellä oleva elinaika 18-, 50-, ja 63- vuotiaana, teoreettisten odotesimulaatioiden mediaani ja 95%-luottamusväli



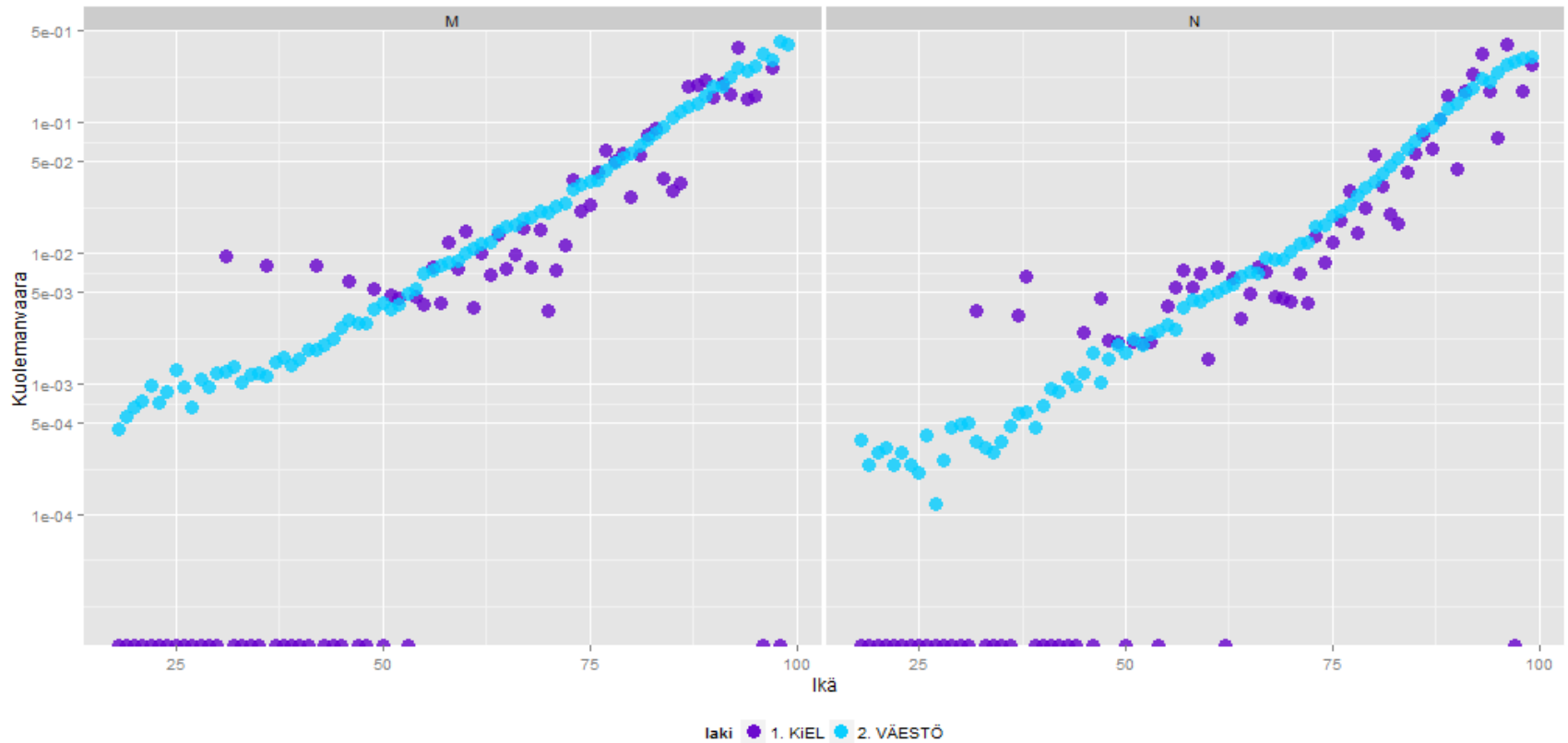
Lukumääräkuolevuuden arviointi

Konstruoidaan vuoden 2013 KiEL-kuolevuuden piste-estimaatti

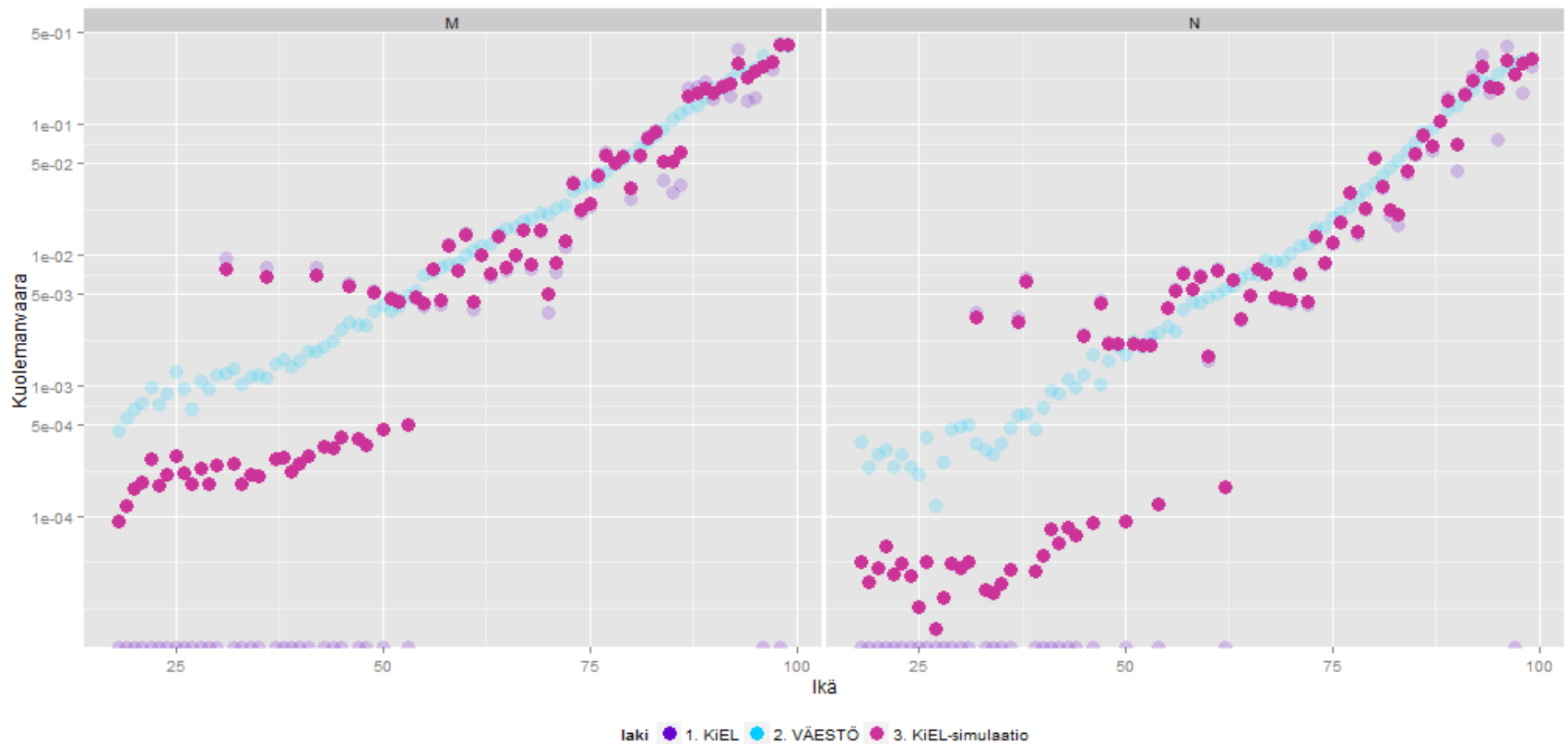
Havaitut KiEL-kuolemanvaarat vuonna 2013



Väestön kuolemanvaaraluvut avuksi



Simuloitujen kuolemanvaarojen odotusarvot vuonna 2013, KiEL

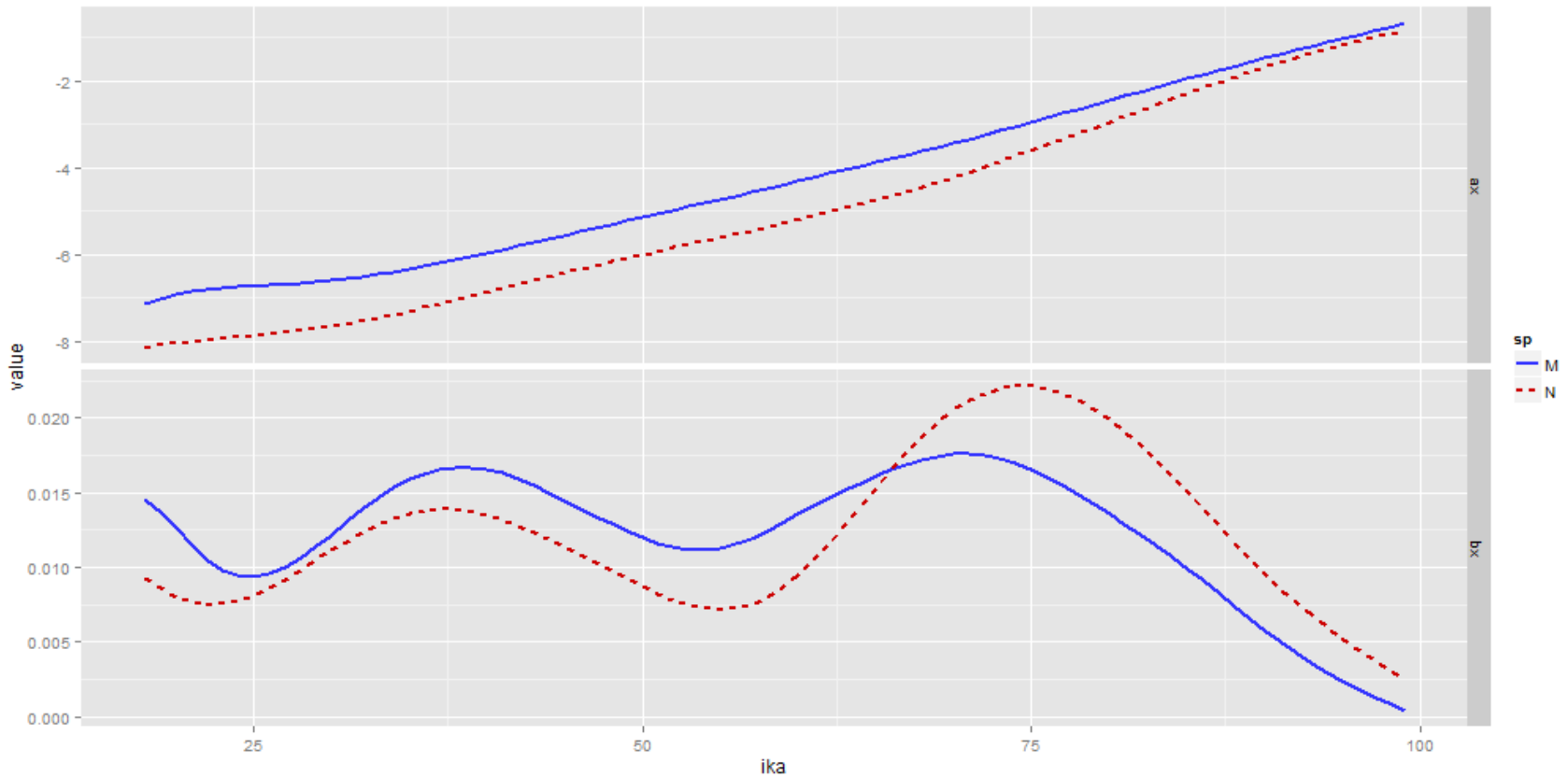


Lee-Carter apuun!

- Kuolevuuden tason hahmottamiseksi täytyy löytää jokin keino hyödyntää aikaisempina vuosina saatua informaatiota.
- Tämä vaatii tietysti ylimääräisiä oletuksia vuosien välisistä yhteyksistä.
- Tavoite on, että tehdyt oletukset olisivat kuitenkin luonnollisia ja perusteltavissa.
- Otetaan apuun väestöennusteiden nykystandardi: **Lee-Carter**.
- Eli sovitetaan väestön keskikuolevuuksiin 1986-2013 seuraava malli:

$$\ln(m_{x,t}) = a_x + b_x k_t + e_{x,t}$$

Väestödatan Lee-Carter ikäriippuvuus, parametrit a_x ja b_x



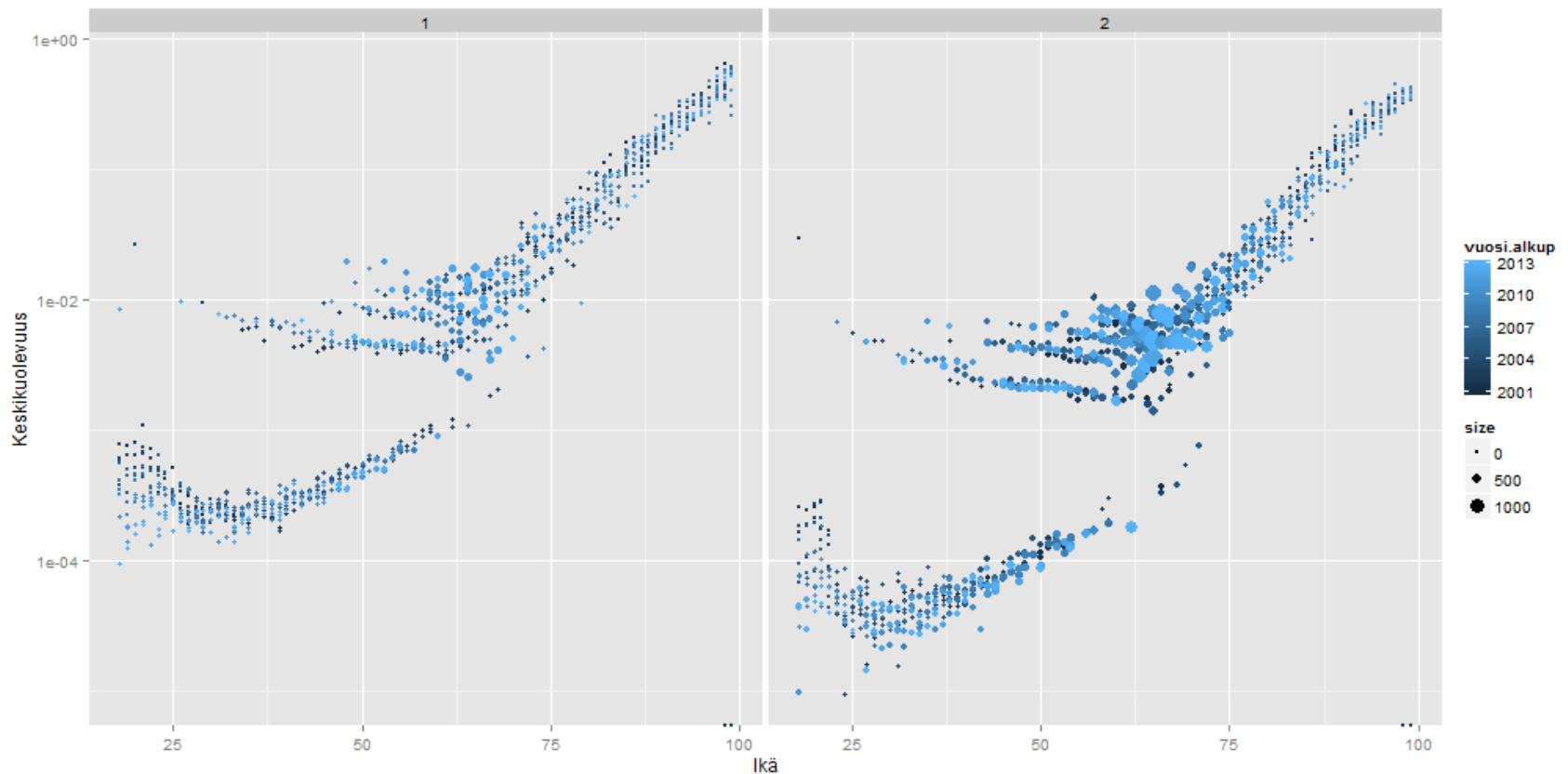
Lee-Carter sovittien aikariippuvuus k_t



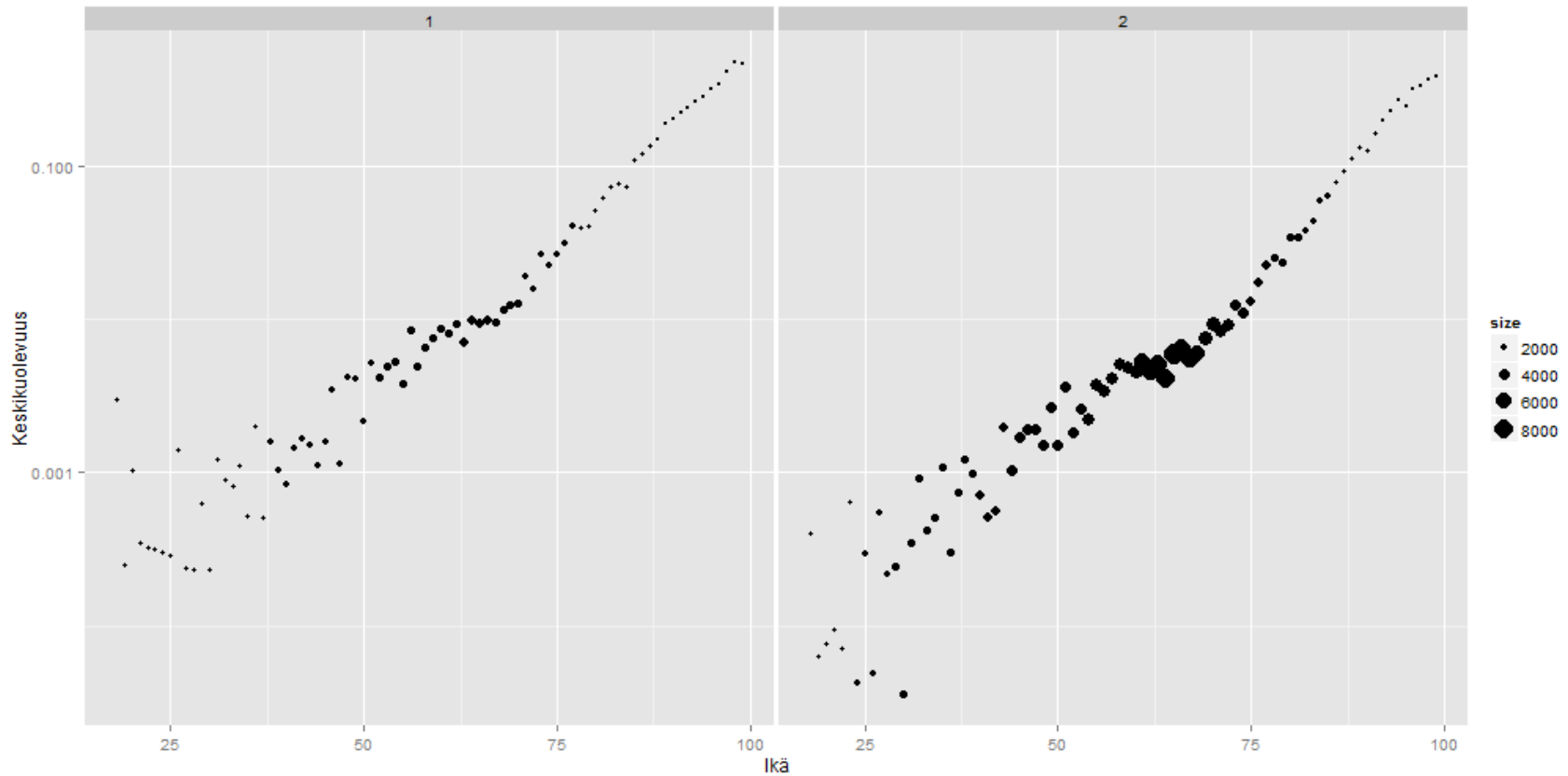
Aikahyppelyä

- Käyttämällä väestön aikariippuvuuskerrointa k_t ja ikäluokittaista herkkyysskerrointa b_x voimme nyt "siirtää" havaintojamme eri vuosien välillä.
- Oletamme siis implisiittisesti, että väestössä havaitut trendit kuolevuuden kehityksestä vaikuttavat myös lakikohtaisten kuolevuuksien taustalla.
- Kompensoimme kuitenkin siirtämisen tuomaa epätarkkuutta pienentämällä menneiden vuosien havaintojen posteriorijakaumien otoskokoa siirrettyjen vuosien määrän funktiona.

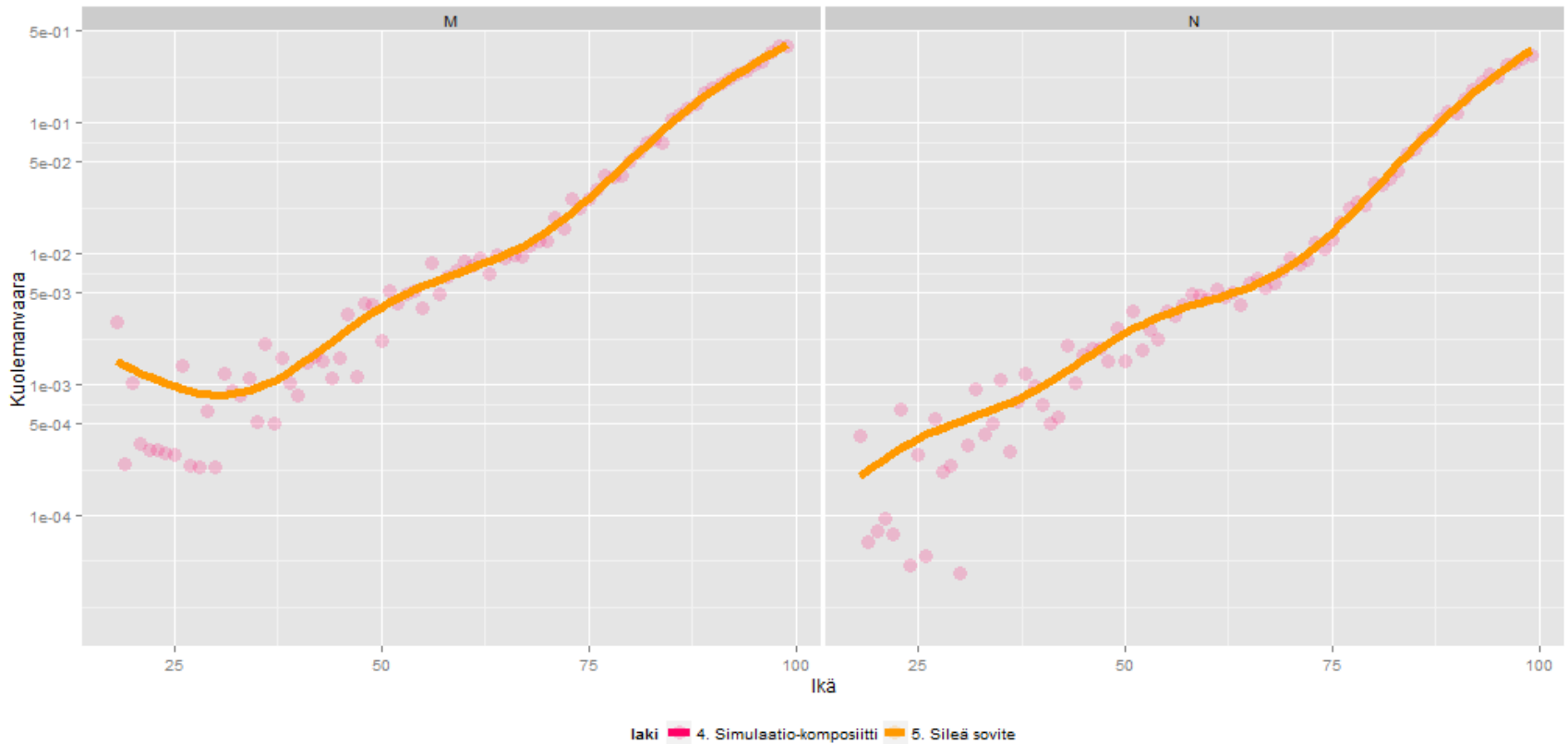
Lee–Carter menetelmän mukaisella aikariippuvaisuuskertoimella (k_t) vuoteen 2013 siirretyt ikäluokittaiset KiEL-posteriorijakaumat: alkuperäinen vuosi, odotusarvo ja informaatiokadolla modifioitu otoskoko. Miehet (1) ja naiset (2). Vuotuiseksi "informaatiokadoksi" on oletettu 7 %



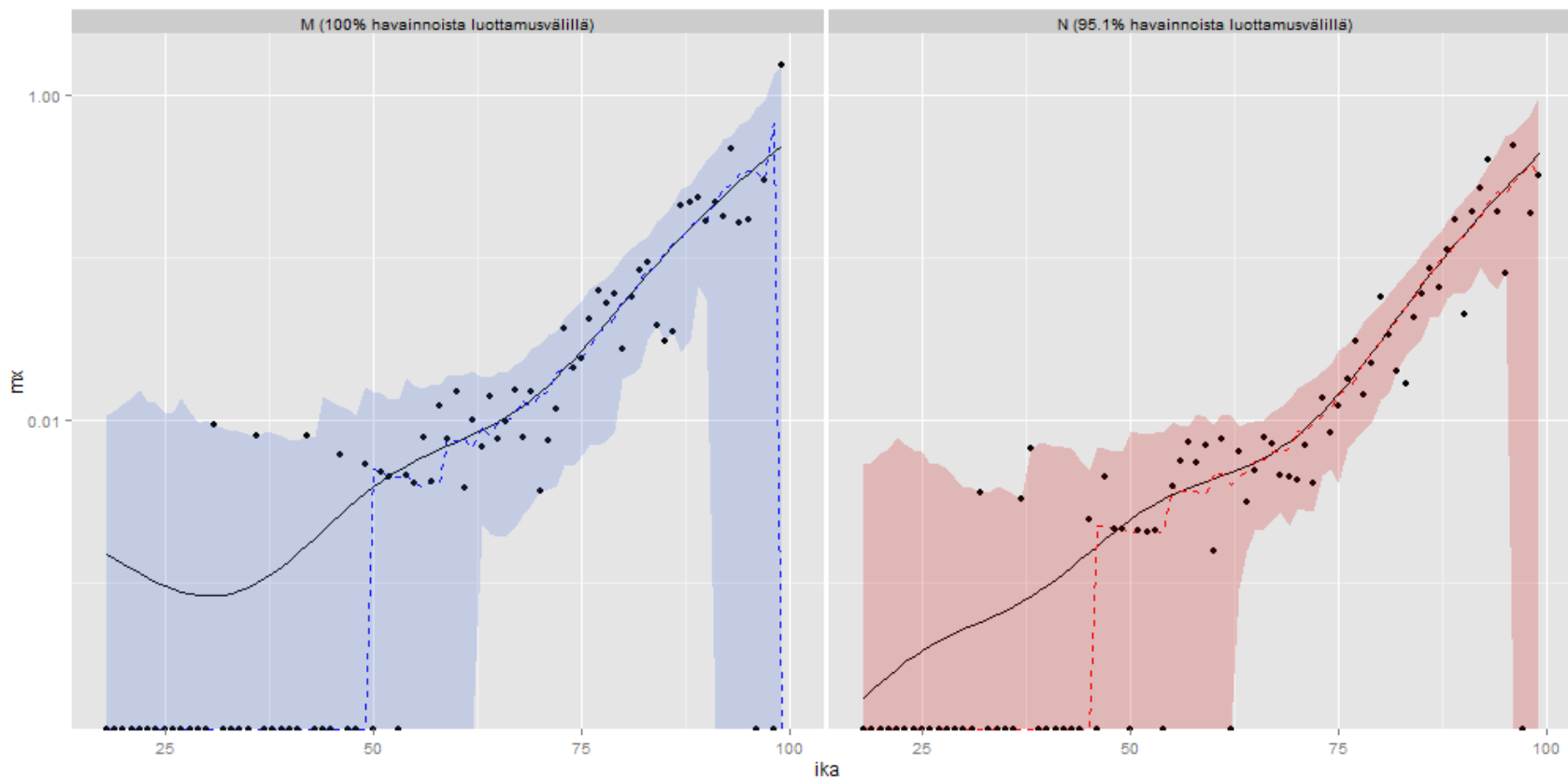
Yhdistetään eri vuosilta saadut siirretyt keskikuolevuusjakaumat ikäluokittain ja lasketaan yhdistettyjen jakaumien odotusarvot



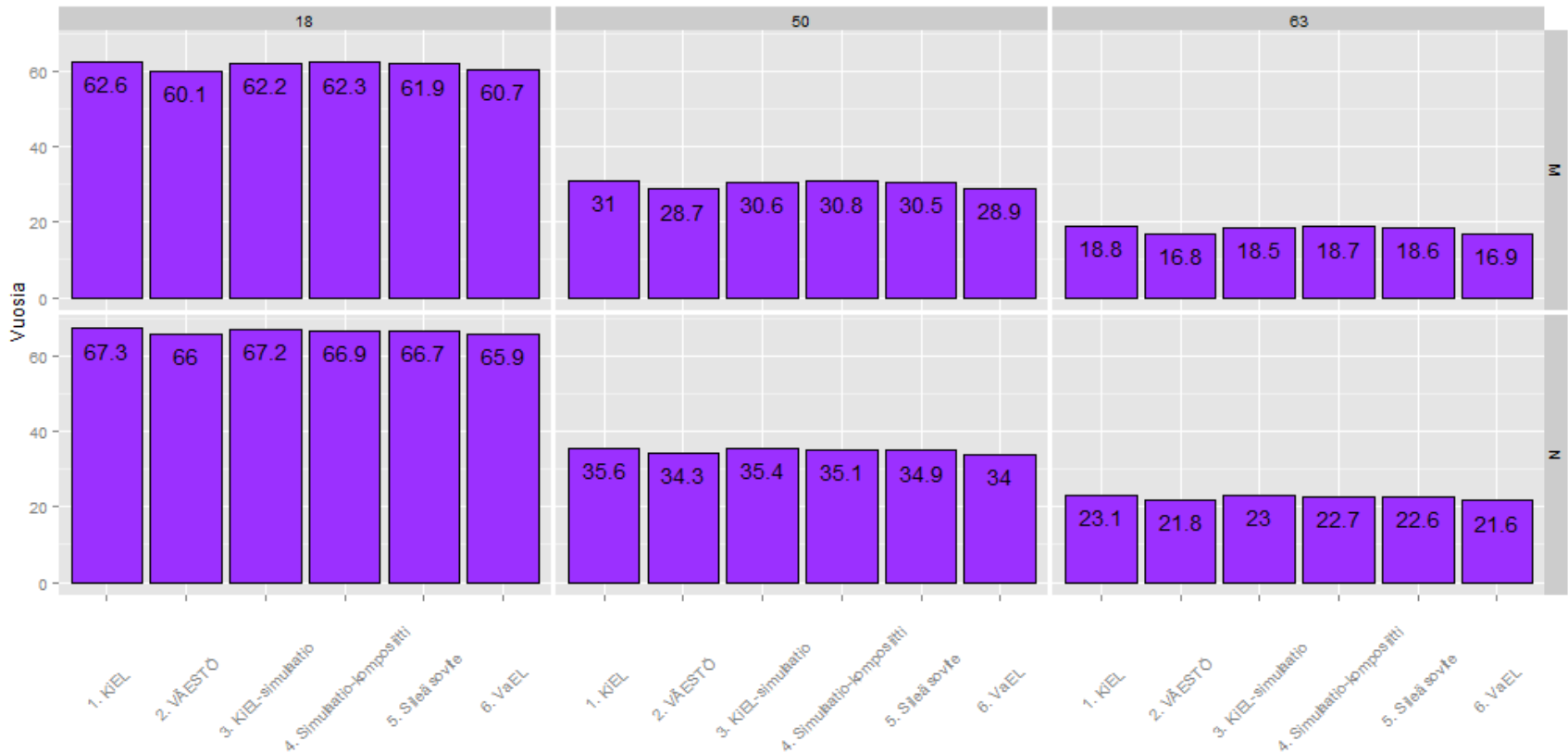
Siirrettyjen ja yhdistettyjen kuolemanvaara- jakaumien odotusarvoihin sovitettu splini- approksimaatio



Validointi: KiEL-kuolevuussovite 2013 ja tätä vastaava havaittavan KiEL-keskikuolevuuden mediaani ja 95%-luottamusväli vuoden 2013 KiEL-altistumin, sekä todella havaittu KiEL-keskikuolevuus 2013



Elinajanodotteet (18-, 50- ja 63-vuotiaana) vuonna 2013: havainnot, simulaatiot ja sovite



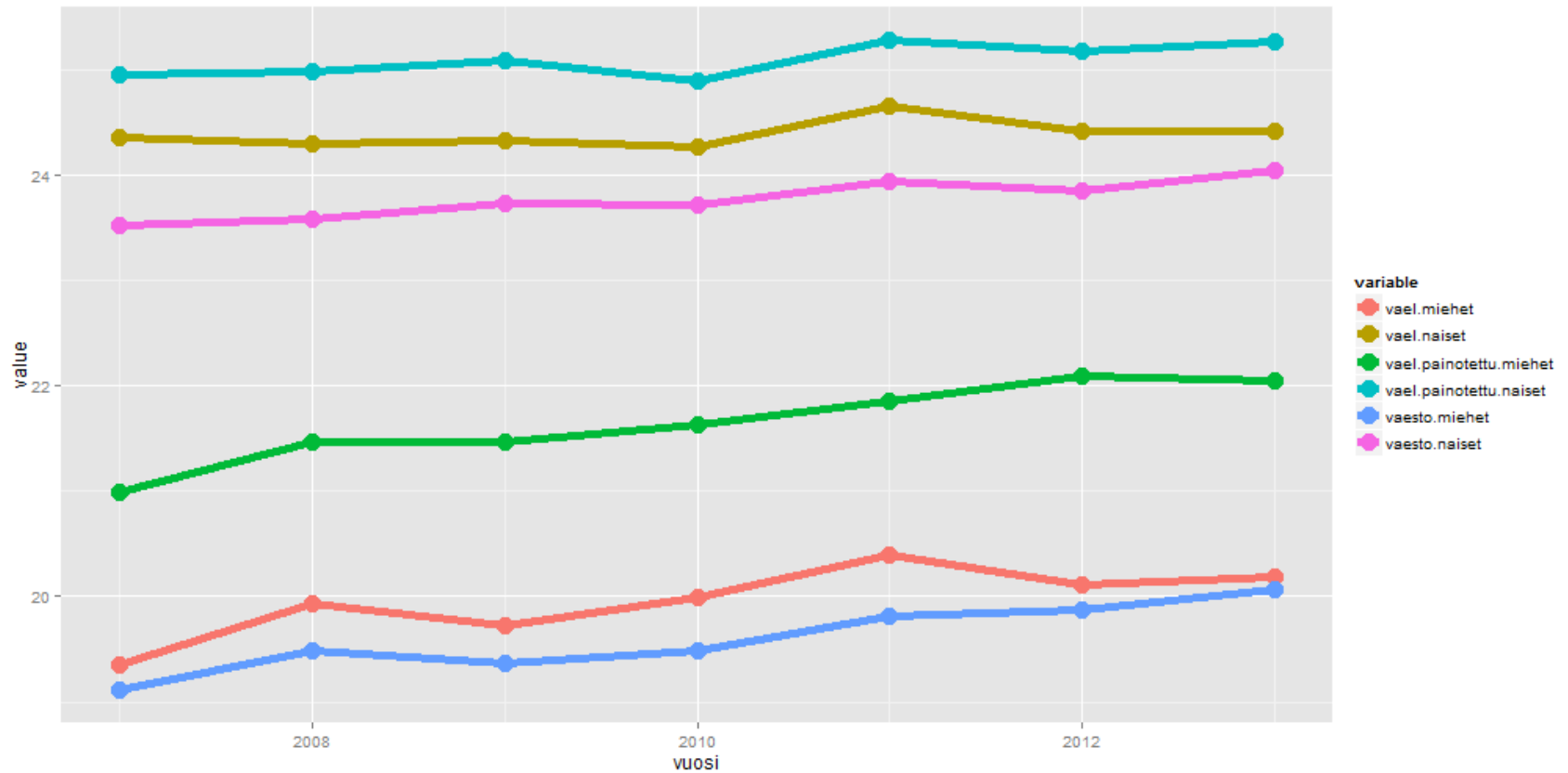
Moneta longa, vita brevis?

Rahakuolevuuden ja lukumääräkuolevuuden eroista

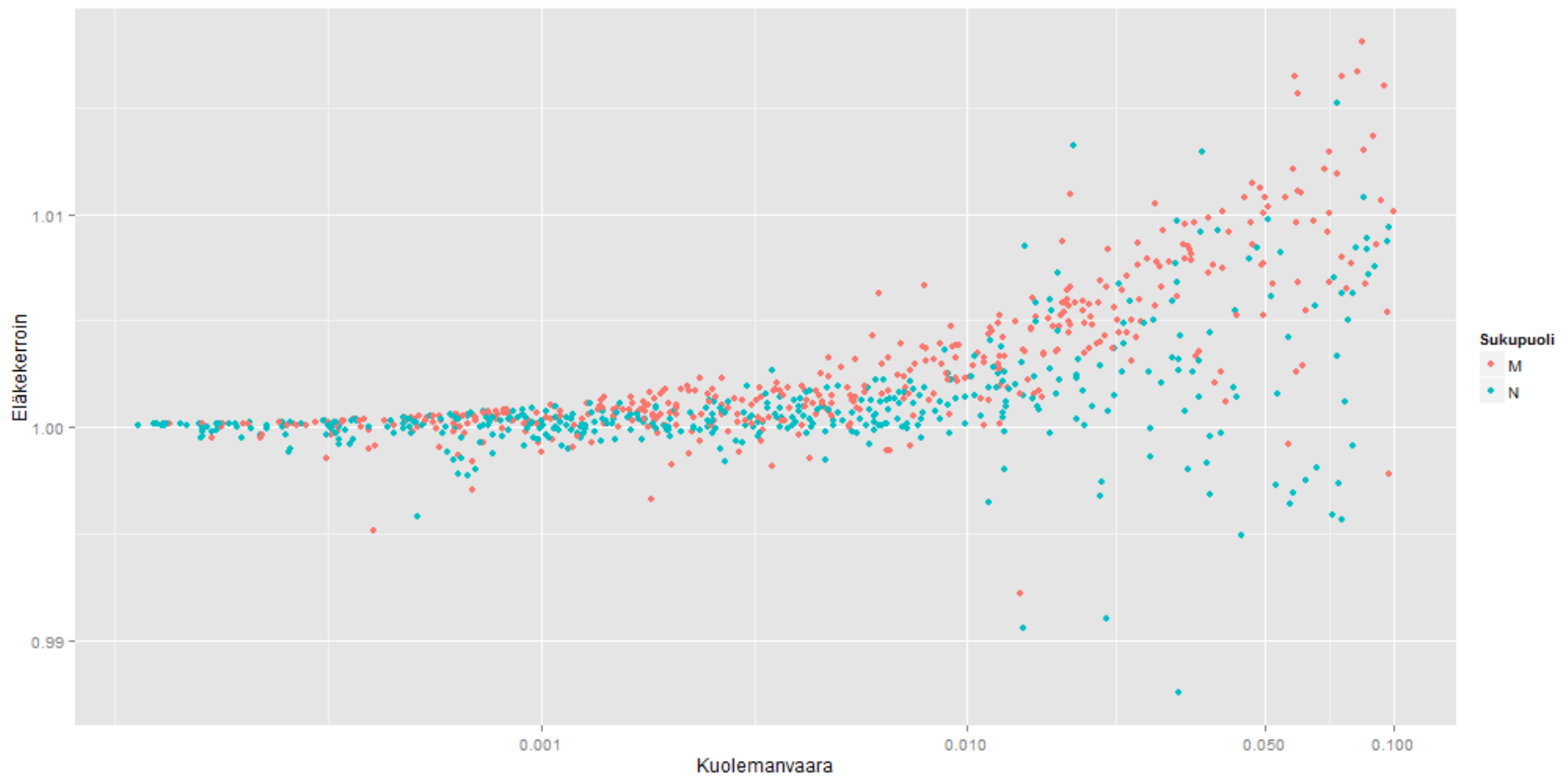
Kappalekuolevuus vs rahakuolevuus

- Tulotaso ja kuolevuus korreloivat tunnetusti käänteisesti (rahaefekti).
- Monissa sovelluksissa käytetäänkin kappalekuolevuuden sijasta raha(sto)kuolevuutta.
- Käytettäessä kappalekuolevuutta eläkemenoennusteissa tulee rahaefekti ottaa erikseen huomioon.

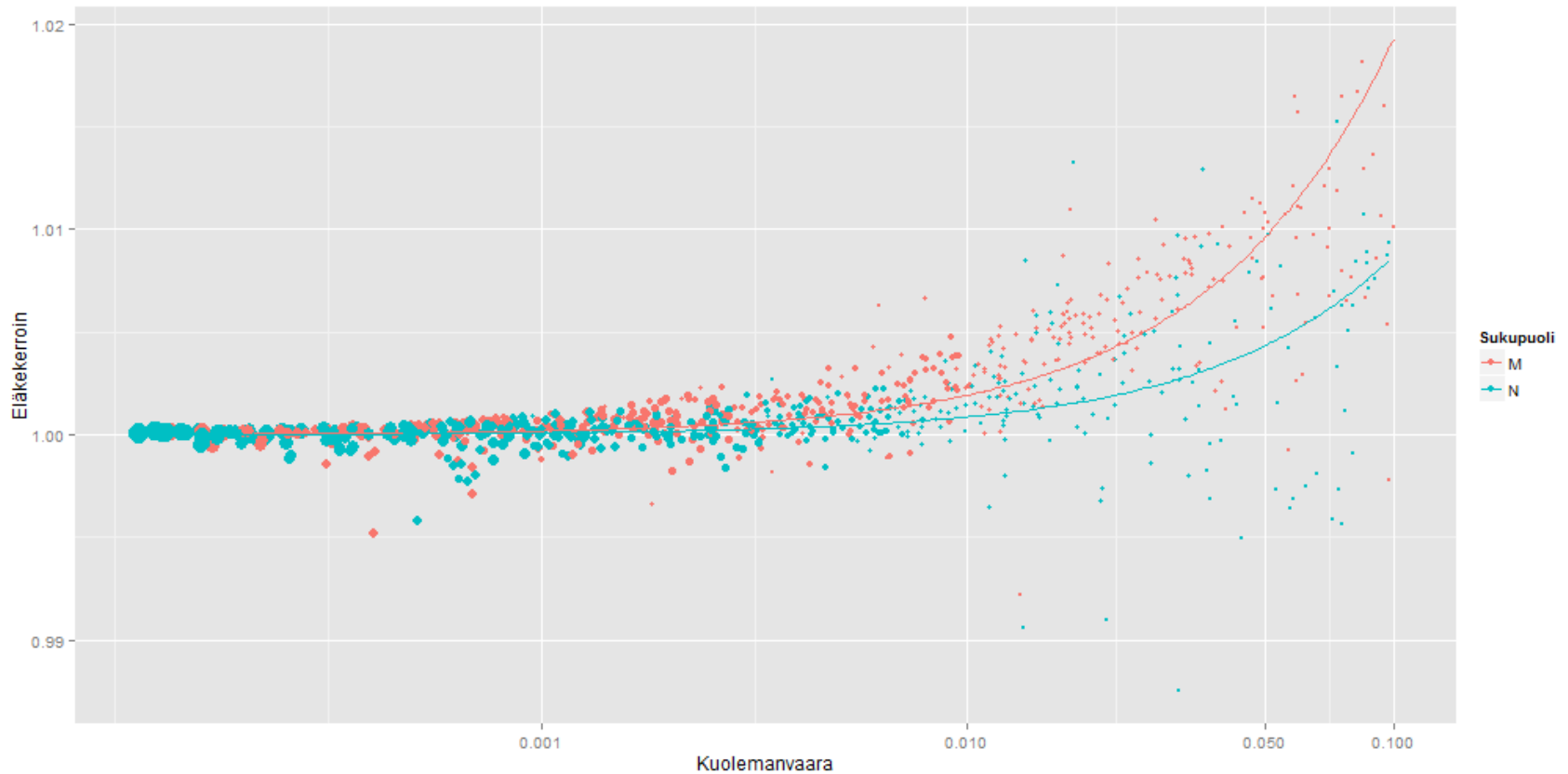
VaEL-piirissä havaittu elinajanodote 62-vuotiaana



Havaittu kuolemanvaara vs kohortin eloonjääneiden eläkkeiden/karttumien keskisuuruus suhteessa koko kohortin eläkkeisiin



Sovitetaan lineaarinen malli rahaefektidataan



Kiitos 2001-2013!

- **Lakikohtaiset kuolevuustiedot** 2001-2013: ETK/Sergei Lahti.
- **Väestön kuolevuustiedot** 1986-2013: Tilastokeskus.
- **VaEL-rahakuolevuustietojen** raakadata on peräisin Kevan rekistereistä ja ETK:n eläkkeiden perintätiedostosta.
- **R-ohjelmisto:**
Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Kuvat laadittin R:n **ggplot2**-paketilla:
H. Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009.
- Lee-Carter sovite ja sileät splini aproksi-maatiot laadittiin R:n **demography** paketilla:
Rob J Hyndman with contributions from Heather Booth, Leonie Tickle, John Maindonald, Simon Wood and R Core Team. (2014). demography: Forecasting mortality, fertility, migration and population data. R package version 1.17.
<http://CRAN.R-project.org/package=demography>
- **Valokuvat:** noppakuva (CC2.0 lisenssi): Mattia Merlo.