

# Eksponenttisen perheen tila-avaruusmallien sovellus alkoholikuolleisuusaineistoon

Elina Ahola

Tilastotieteen pro gradu -tutkielma

Jyväskylän yliopisto  
Matematiikan ja tilastotieteen laitos  
31. elokuuta 2009

## Tiivistelmä

Elina Ahola: *Eksponenttisen perheen tila-avaruusmallien sovellus alkoholikuolleisuusaineistoon.*

Tilastotieteen pro gradu -tutkielma, Jyväskylän yliopisto, 31. elokuuta 2009.  
Sivuja 37, 1 liite.

Tila-avaruusmalleilla analysoidaan aikasarjoja. Tila-avaruusmallien perustulokset, kuten Kalmanin suodin ja tasoint, pätevät vain lineaarisille ja gaussisille tila-avaruusmalleille. Näiden perustulosten avulla saadaan kuitenkin lineaarisia ja ei-gaussisia sekä erityisesti eksponenttisen perheen tila-avaruusmalleja koskevia tuloksia, kun käytetään Monte Carlo -integrointia, tärkeyspainotusta ja ei-gaussisen mallin linearisaatiota.

Suomalaisten alkoholikuolleisuusaineistoon vuosilta 1969–2007 voidaan sovittaa eksponenttisen perheen tila-avaruusmallien ryhmään kuuluva Poisson-jakaumaan perustuva malli. Tuloksina saadaan, että koko väestöä kerralla tarkasteltaessa alkoholikuolleisuus on tarkasteluvuosina kasvanut. Suurin kasvu on ajoittunut vuoden 2004 jälkeiseen aikaan. Kyseisenä vuonna Suomessa suoritettiin alkoholin veronalennus ja matkustajatuonti Virosta vapautui. Koko väestön kanssa samantyyppinen kehityskulku nähdään myös naisten ja miesten ryhmiä erikseen tarkasteltaessa. Ikäryhmittäisistä tuloksista nähdään, että alkoholikuolleisuus on melko suoraviivaisesti kasvanut 40–79-vuotiaiden joukossa, kun taas 20–39-vuotiaiden joukossa kuolleisuus on ollut vaihtelevampaa.

*Avainsanoja:* aikasarja-analyysi, alkoholikuolleisuus, eksponenttinen perhe, linearisaatio, Monte Carlo -integrointi, Poisson-jakauma, tila-avaruusmalli, tärkeyspainotus.

# Kiitosmaininnat

Haluan kiittää Tilastokeskuksen yliaktuaari Helena Korpea, joka ystävällisesti kokosi tarvittavan aineiston käyttööni. Lisäksi haluan kiittää työni ohjaajaa professori Jukka Nyblomia neuvoista ja kannustuksesta tutkielman teossa.

Jyväskylässä 31.8.2009  
Elina Ahola

# Sisältö

<b>1 Johdanto</b>	<b>1</b>
1.1 Merkinnöistä . . . . .	1
<b>2 Tila-avaruusmallien teoriaa</b>	<b>3</b>
2.1 Lineaarisista ja gaussisista tila-avaruusmalleista . . . . .	3
2.1.1 Suodatus . . . . .	4
2.1.2 Tasoitus tiloille . . . . .	4
2.1.3 Tasoitus häiriötermeille . . . . .	5
2.1.4 Havainnoilla ehdollistettu simulointi . . . . .	5
2.1.5 Parametrien estimointi . . . . .	8
2.1.6 Alustus . . . . .	8
2.2 Lineaarisista ja ei-gaussisista tila-avaruusmalleista . . . . .	8
2.2.1 Monte Carlo -integrointi ja tärkeyspainotus . . . . .	9
2.2.2 Linearisaatio . . . . .	11
2.2.3 Tasoitus tiloille . . . . .	15
2.2.4 Luottamusväli signaaleille . . . . .	15
2.2.5 Parametrien estimointi . . . . .	16
<b>3 Sovellus alkoholikuolleisuusaineistoon</b>	<b>21</b>
3.1 Yhteiskunnallinen tausta . . . . .	21
3.1.1 Alkoholilojen historiasta Suomessa . . . . .	21
3.1.2 Suomalaisten alkoholiasenteista . . . . .	22
3.2 Aineistosta . . . . .	23
3.2.1 Alkoholikuolleisuuden määrittely . . . . .	23
3.2.2 Aineiston kuvailu . . . . .	26
3.3 Mallin määrittely . . . . .	28
3.4 Mallin sovitus . . . . .	30
3.5 Johtopäätökset . . . . .	36
<b>Viitteet</b>	<b>37</b>
<b>Liitteet</b>	
<b>A R-ohjelmiston koodit</b>	<b>38</b>

# Luku 1

## Johdanto

Tila-avaruusmalleilla voidaan analysoida aikasarjoja. Tila-avaruusmallinnuksessa oletetaan, että tutkittavassa prosessissa ajassa tapahtuvaa muutosta voidaan kuvata tiloilla  $\alpha_1, \dots, \alpha_n$ . Näitä tiloja ei havaita, mutta havainnot  $y_1, \dots, y_n$  ovat tiedossa.

Tila-avaruusmalli koostuu kahdesta osasta, havaintoyhtälöstä ja tilayhtälöstä. Havaintoyhtälö kuvaa yhteyden havainnon  $y_t$  ja tilan  $\alpha_t$  välillä. Tilayhtälö puolestaan kuvaa yhteyden kahden peräkkäisen tilan  $\alpha_t$  ja  $\alpha_{t+1}$  välillä.

Tila-avaruusmallit voidaan jaotella pääryhmiin sen mukaan, ovatko havainto- ja tilayhtälöt lineaarisia vai eivät ja vastaavasti, ovatko havainto- ja tilayhtälöt gaussisia vai eivät. Tässä pro gradu -tutkielmassa pääpaino on lineaaristen ja ei-gaussisten mallien teorialla. Erityisesti keskitytään sellaisiin malleihin, joissa havaintoyhtälö on eksponenttisen perheen muotoa ja tilayhtälö on gaussinen.

Eksponenttisen perheen tila-avaruusmallien teoriaa sovelletaan suomalaisten alkoholikuolleisuusaineistoon vuosilta 1969-2007. Aineistoon sovitetaan tila-avaruusmalli, jossa havaintoyhtälö perustuu Poisson-jakaumaan ja tilayhtälö on gaussinen.

### 1.1 Merkinnöistä

Tässä tutkielmassa käytetään seuraavanlaisia merkintöjä:

- Ajanhetken  $t$  havainto on  $y_t$ .
- Ajanhetken  $t$  tila on  $\alpha_t$ .
- Yksinkertaisuuden vuoksi jakaumille käytetään geneerisiä merkintöjä. Merkinnöillä  $f(\cdot)$ ,  $f(\cdot, \cdot)$  ja  $f(\cdot|\cdot)$  tarkoitetaan jatkuvan satunnaismuuttujan tapauksessa jakauman tiheysfunktioita, yhteisjakauman tiheysfunktioita ja ehdollisen jakauman tiheysfunktioita. Diskreetin satunnaismuuttujan tapauksessa merkinnöillä tarkoitetaan vastaavasti jakauman pistetodennäköisyysfunktioita, yhteisjakauman pistetodennäköisyysfunktioita ja ehdollisen jakauman pistetodennäköisyysfunktioita. Lyhyesti näitä kaikkia kutsutaan jakaumiksi. Lineaarisen ja gaussisen mallin tapauksessa jakauman tunnus on  $g$ ; lineaarisen ja ei-gaussisen mallin tapauksessa jakauman tunnus on  $p$ .

- Mitkä tahansa yksittäiset vektorit  $x_1, \dots, x_n$  voidaan koota edelleen kokoelma-vektoriksi  $x = (x'_1, \dots, x'_n)'$ .
- Symboli  $0$  tarkoittaa tilanteesta riippuen joko arvoa nolla, nollavektoria tai nollamatriisia.

# Luku 2

## Tila-avaruusmallien teoriaa

Tässä mallin teoriaa käsittelevässä luvussa on käytetty pääasiallisena lähteenä Durbinin ja Koopmanin teosta *Time Series Analysis by State Space Methods* (2001) [4]. Aluksi esitellään lyhyesti lineaaristen ja gaussisten tila-avaruusmallien teoriaa. Sen jälkeen siirrytään tarkastelemaan tarkemmin lineaaristen ja ei-gaussisten tila-avaruusmallien teoriaa painottaen eksponenttisen perheen malleja.

### 2.1 Lineaarista ja gaussista tila-avaruusmalleista

Teoreettisesti yksinkertaisin tila-avaruusmallien perhe ovat lineaariset ja gaussiset tila-avaruusmallit. Näille malleille saatuja tuloksia hyödynnetään vahvasti myös muiden tila-avaruusmalliperheiden teoriaa rakennettaessa, joten siksi lineaaristen ja gaussisten mallien teoria esitellään ensimmäisenä.

Yleinen lineaarinen ja gaussinen tila-avaruusmalli määritellään seuraavasti:

**Määritelmä 2.1** (Lineaarinen ja gaussinen tila-avaruusmalli).

$$\begin{aligned}y_t &= Z_t \alpha_t + \epsilon_t, & \epsilon_t &\sim N(0, H_t), \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, & \eta_t &\sim N(0, Q_t),\end{aligned}$$

kaikilla  $t = 1, \dots, n$ , missä muuttujat  $\epsilon_t$  ja  $\eta_t$  ovat kaikki keskenään riippumattomia ja riippumattomia myös muuttujasta  $\alpha_1$ . Lisäksi oletetaan, että  $\alpha_1 \sim N(a_1, P_1)$ . Matriisin  $R_t$  oletetaan koostuvan  $m \times m$  -kokoisen yksikkömatriisin  $I_m$  sarakkeista. Vektoreiden ja matriisien dimensiot on esitelty taulukossa 2.1.

Määritelmän 2.1 ensimmäinen yhtälö on havaintoyhtälö ja toinen yhtälö tilayhtälö. Termejä  $\epsilon_t$  ja  $\eta_t$  kutsutaan häiriötermeiksi. Matriisia  $R_t$  kutsutaan valintamatriisiksi ja sen rakenteesta johtuen pätee, että  $\eta_t = R_t'(\alpha_{t+1} - T_t \alpha_t)$ .

Taulukko 2.1: Vektoreiden ja matriisien dimensiot.

Vektori		Matriisi	
$y_t$	$p \times 1$	$Z_t$	$p \times m$
$\alpha_t$	$m \times 1$	$T_t$	$m \times m$
$\epsilon_t$	$p \times 1$	$H_t$	$p \times p$
$\eta_t$	$r \times 1$	$R_t$	$m \times r$
		$Q_t$	$r \times r$
$a_1$	$m \times 1$	$P_1$	$m \times m$

### 2.1.1 Suodatus

Suodatus tarkoittaa tilan  $\alpha_t$  jakauman selvittämistä ehdolla havainnot  $y_1, \dots, y_{t-1}$ . Jakauma selviää Kalmanin suotimen avulla.

**Lause 2.2** (Kalmanin suodin). Olkoon määritelmän 2.1 mukainen lineaarinen ja gaussinen tila-avaruusmalli. Tällöin  $\alpha_t | y_1, \dots, y_{t-1} \sim N(a_t, P_t)$  kaikilla  $t = 1, \dots, n+1$ , missä  $a_t$  ja  $P_t$  saadaan seuraavien rekursiivisten kaavojen avulla:

$$\begin{aligned}
 v_t &= y_t - Z_t a_t, \\
 F_t &= Z_t P_t Z_t' + H_t, \\
 K_t &= T_t P_t Z_t' F_t^{-1}, \\
 L_t &= T_t - K_t Z_t, \\
 a_{t+1} &= T_t a_t + K_t v_t, \\
 P_{t+1} &= T_t P_t L_t' + R_t Q_t R_t'.
 \end{aligned}$$

**Todistus:** Durbin & Koopman (2001) [4], s. 65–68.  $\square$

Lauseessa 2.2 vektori  $v_t = y_t - E(y_t | y_1, \dots, y_{t-1})$  on yhden askeleen ennustevirhe muuttujalle  $y_t$ , kun muuttujat  $y_1, \dots, y_{t-1}$  tunnetaan.  $F_t = \text{var}(v_t)$  on yhden askeleen ennustevirheen varianssi. Matriisia  $K_t$  kutsutaan nimellä Kalman gain. Se ilmaisee, mikä on uuden havainnon vaikutus suodatettuun odotusarvoon  $a_t$ . Muuttujaa  $L_t$  käytetään vain laskuissa apuna eikä sille ole tilastollista tulkintaa.

### 2.1.2 Tasoitus tiloille

Tasoituksen tarkoituksena on selvittää tilan  $\alpha_t$  odotusarvo ja varianssi ehdolla kaikki havainnot  $y$ .



**Lause 2.3** (Tilojen tasoitusrekursio). Olkoon määritelmän 2.1 mukainen lineaarinen ja gaussinen tila-avaruusmalli. Merkitään  $\hat{\alpha}_t = E_g(\alpha_t|y)$  ja  $V_t = \text{var}_g(\alpha_t|y)$ .  $\hat{\alpha}_t$  ja  $V_t$  saadaan seuraavien rekursiivisten kaavojen avulla:

$$\begin{aligned} r_{t-1} &= Z_t' F_t^{-1} v_t + L_t' r_t, \\ N_{t-1} &= Z_t' F_t^{-1} Z_t + L_t' N_t L_t, \\ \hat{\alpha}_t &= a_t + P_t r_{t-1}, \\ V_t &= P_t - P_t N_{t-1} P_t, \end{aligned}$$

missä  $r_n = 0$  ja  $N_n = 0$ .  $F_t$ ,  $v_t$ ,  $L_t$ ,  $a_t$  ja  $P_t$  saadaan Kalmanin suotimesta (lause 2.2).

**Todistus:** Durbin & Koopman (2001) [4], s. 70–73.  $\square$

Lauseessa 2.3 muuttujille  $r_{t-1}$  ja  $N_{t-1}$  ei ole tilastollista tulkintaa. Ne ovat vain laskujen apumuuttujia.

### 2.1.3 Tasoitus häiriötermeille

Halutaan selvittää häiriötermien  $\epsilon_t$  ja  $\eta_t$  odotusarvo ja varianssi ehdolla kaikki havainnot  $y$ .

**Lause 2.4** (Häiriötermien tasoitusrekursio). Olkoon määritelmän 2.1 mukainen lineaarinen ja gaussinen tila-avaruusmalli.  $E_g(\epsilon_t|y)$ ,  $E_g(\eta_t|y)$ ,  $\text{var}_g(\epsilon_t|y)$  ja  $\text{var}_g(\eta_t|y)$  saadaan seuraavien rekursiivisten kaavojen avulla:

$$\begin{aligned} E_g(\epsilon_t|y) &= H_t(F_t^{-1} v_t - K_t' r_t), \\ E_g(\eta_t|y) &= Q_t R_t' r_t, \\ \text{var}_g(\epsilon_t|y) &= H_t - H_t(F_t^{-1} + K_t' N_t K_t) H_t, \\ \text{var}_g(\eta_t|y) &= Q_t - Q_t R_t' N_t R_t Q_t. \end{aligned}$$

Termit  $F_t$ ,  $v_t$  ja  $K_t$  saadaan Kalmanin suotimesta (lause 2.2); termit  $r_t$  ja  $N_t$  puolestaan tilojen tasoitusrekursiosta (lause 2.3).

**Todistus:** Durbin & Koopman (2001) [4], s. 73–77.  $\square$

### 2.1.4 Havainnoilla ehdollistettu simulointi

Havainnoilla ehdollistetulla simuloinnilla tarkoitetaan mallin simulointia ehdolla kaikki havainnot  $y$ . Havainnoilla ehdollistettua simulointia voidaan soveltaa sekä häiriötermeille  $\epsilon$  ja  $\eta$  että tiloille  $\alpha$ . Seuraavassa esiteltävän menetelmän ovat kehittäneet Durbin & Koopman (2001) [3].

## Havainnoilla ehdollistettu häiriötermien simulointi

Johdetaan aluksi menetelmä häiriötermien  $\epsilon$  ja  $\eta$  ehdolliseksi simuloimiseksi ehdolla havainnot  $y$ . Otetaan käyttöön merkinnät  $\omega = (\epsilon', \eta')'$ ,  $\hat{\omega} = E_g(\omega|y)$  ja  $W = \text{var}_g(\omega|y)$ . Tarkoituksena on siis simuloida jakaumaa  $g(\omega|y)$ . Koska malli on lineaarinen ja gaussinen, niin muuttujan  $\omega|y$  jakauma  $g(\omega|y)$  on normaalijakauma  $N(\hat{\omega}, W)$ . Menetelmän ideana on simuloida aluksi vektori jakaumasta  $N(0, W)$  ja lisätä tähän vektori  $\hat{\omega}$ .

Menetelmän toimivuus perustuu tulokseen, että  $W$  ei riipu muuttujasta  $y$  eli  $\text{var}_g(\omega|y) = W$  kaikilla  $y$ . Tulos seuraa yleisestä moniulotteista normaalijakaumaa koskevasta tuloksesta, jonka mukaan ehdollinen varianssimatriisi joidenkin muuttujien ollessa kiinnitettynä ei riipu kiinnitettyjen muuttujien arvoista (Anderson (1984) [1] s. 37).

Tiedetään, että  $\alpha_1 \sim N(a_1, P_1)$ . Olkoon  $\alpha_1^+$  yksi simulointi jakaumasta  $N(a_1, P_1)$ . Muuttujan  $\omega$  jakauma  $g(\omega)$  puolestaan on  $N(0, \Omega)$ , missä

$$\Omega = \begin{pmatrix} H_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & H_n & 0 & \cdots & 0 \\ 0 & \cdots & 0 & Q_1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & Q_n \end{pmatrix}.$$

Olkoon  $\omega^+ = ((\epsilon^+)', (\eta^+))'$  simulointi jakaumasta  $g(\omega)$ . Näiden normaalijakaumien  $N(a_1, P_1)$  ja  $g(\omega)$  simulointiin on olemassa monia menetelmiä. Tässä tutkielmassa on käytetty Choleskyn hajotelmaa. Generoidaan seuraavaksi vektori  $y^+$  lineaarisesta ja gaussisesta mallista (määritelmä 2.1), jossa muuttujille  $\alpha_1$ ,  $\epsilon$  ja  $\eta$  on asetettu simuloidut arvot  $\alpha_1^+$ ,  $\epsilon^+$  ja  $\eta^+$ . Saadaan siis satunnaismuuttuja  $y^+$ . Lasketaan  $\hat{\omega}^+ = E_g(\omega^+|y^+)$  häiriötermien tasoitusrekursiolla (lause 2.4) eli  $\hat{\omega}^+$  on muuttujasta  $y^+$  riippuva satunnaismuuttuja. Koska  $W$  on riippumaton muuttujan  $y^+$  arvosta, niin  $\text{var}_g(\omega^+|y^+) = W$ . Ehdollisen odotusarvon ja ehdollisen varianssin laskusääntöjen avulla saadaan

$$\begin{aligned} E_g(\omega^+ - \hat{\omega}^+) &= E_g[E_g(\omega^+ - \hat{\omega}^+|y^+)] = E_g[E_g(\omega^+|y^+) - E_g(\hat{\omega}^+|y^+)] \\ &= E_g[\hat{\omega}^+ - \hat{\omega}^+] = 0 \end{aligned}$$

ja

$$\begin{aligned} \text{var}_g(\omega^+ - \hat{\omega}^+) &= E_g[\text{var}_g(\omega^+ - \hat{\omega}^+|y^+)] + \text{var}_g[E_g(\omega^+ - \hat{\omega}^+|y^+)] \\ &= E_g[\text{var}_g(\omega^+|y^+) + \text{var}_g(\hat{\omega}^+|y^+)] + \text{var}_g[E_g(\omega^+|y^+) - E_g(\hat{\omega}^+|y^+)] \\ &= E_g[W + 0] + \text{var}_g[\hat{\omega}^+ - \hat{\omega}^+] = W. \end{aligned}$$

Koska kyseessä on lineaarinen ja gaussinen malli, niin  $\omega^+ - \hat{\omega}^+ \sim N(0, W)$ . Ollaan siis saatu vektori  $\omega^+ - \hat{\omega}^+$  jakaumasta  $N(0, W)$  muuttujasta  $y$  riippumattomasti, sillä simuloinnissa ei missään vaiheessa käytetty muuttujaa  $y$ .

Lasketaan seuraavaksi  $\hat{\omega} = E_g(\omega|y)$  häiriötermien tasoitusrekursiolla (lause 2.4). Merkitään  $\tilde{\omega} = \hat{\omega} + \omega^+ - \hat{\omega}^+$ . Koska muuttuja  $\omega^+ - \hat{\omega}^+$  on riippumaton muuttujasta

$y$  ja kun  $y$  oletetaan kiinnitetyksi, niin saadaan

$$\begin{aligned} E_g(\tilde{\omega}|y) &= E_g(\hat{\omega} + \omega^+ - \hat{\omega}^+|y) = E_g(\hat{\omega}|y) + E_g(\omega^+ - \hat{\omega}^+|y) \\ &= \hat{\omega} + E_g(\omega^+ - \hat{\omega}^+) = \hat{\omega} + 0 = \hat{\omega}. \end{aligned}$$

ja

$$\begin{aligned} \text{var}_g(\tilde{\omega}|y) &= \text{var}_g(\hat{\omega} + \omega^+ - \hat{\omega}^+|y) = \text{var}_g(\omega^+ - \hat{\omega}^+|y) \\ &= \text{var}_g(\omega^+ - \hat{\omega}^+) = W. \end{aligned}$$

Jälleen koska malli on lineaarinen ja gaussinen, niin  $\tilde{\omega} \sim N(\hat{\omega}, W)$ . Siis  $\tilde{\omega}$  on simulaatio jakaumasta  $g(\omega|y)$ .

Seuraava algoritmi kokoaa yhteen simuloinnin vaiheet. Algoritmin tuloksena saatava  $\tilde{\omega}$  on simulaatio jakaumasta  $g(\omega|y)$ .

*Algoritmi 1:* Havainnoilla ehdollistettu häiriötermien simulointi.

1. Simuloi  $\omega^+ = ((\epsilon^+)', (\eta^+)')$  jakaumasta  $g(\omega)$  ja  $\alpha_1^+$  jakaumasta  $N(a_1, P_1)$ . Generoi arvot  $y^+$  lineaarisesta ja gaussisesta mallista (määritelmä 2.1), jossa muuttujille  $\alpha_1$ ,  $\epsilon$  ja  $\eta$  on asetettu simuloitut arvot  $\alpha_1^+$ ,  $\epsilon^+$  ja  $\eta^+$ .
2. Laske  $\hat{\omega} = E_g(\omega|y)$  ja  $\hat{\omega}^+ = E_g(\omega^+|y^+)$  häiriötermien tasoitusrekursion (lause 2.4) avulla.
3. Laske  $\tilde{\omega} = \hat{\omega} + \omega^+ - \hat{\omega}^+$ .

### Havainnoilla ehdollistettu tilojen simulointi.

Tarkoituksena on simuloida tilavektorin  $\alpha$  arvoja ehdolla kaikki havainnot  $y$  eli simuloidaan siis arvoja jakaumasta  $g(\alpha|y)$ .

Seuraava algoritmi simuloi yhden arvon, merkitään  $\tilde{\alpha}$ , jakaumasta  $g(\alpha|y)$ . Simulointialgoritmin toimivuus voidaan perustella samalla tavalla kuin häiriötermien simulointiin kehitetty algoritmi 1.

*Algoritmi 2:* Havainnoilla ehdollistettu tilojen simulointi

1. Simuloi  $\omega^+ = ((\epsilon^+)', (\eta^+)')$  jakaumasta  $g(\omega)$  ja  $\alpha_1^+$  jakaumasta  $N(a_1, P_1)$ . Generoi arvot  $y^+$  lineaarisesta ja gaussisesta mallista (määritelmä 2.1), jossa muuttujille  $\alpha_1$ ,  $\epsilon$  ja  $\eta$  on asetettu simuloitut arvot  $\alpha_1^+$ ,  $\epsilon^+$  ja  $\eta^+$ .
2. Laske  $\hat{\alpha} = E_g(\alpha|y)$  ja  $\hat{\alpha}^+ = E_g(\alpha^+|y^+)$  tilojen tasoitusrekursion (lause 2.3) avulla.
3. Laske  $\tilde{\alpha} = \hat{\alpha} + \alpha^+ - \hat{\alpha}^+$ .

### 2.1.5 Parametrien estimointi

Edellä oleva teoria on rakennettu sille oletukselle, että kaikki mallin määräävät matriisit  $Z_t$ ,  $H_t$ ,  $T_t$ ,  $R_t$  ja  $Q_t$  ovat tunnettuja. Käytännön sovelluksissa yksi tai useampi näistä matriiseista yleensä riippuu tuntemattomista parametreista. Kootaan nämä parametrit parametrivektoriksi  $\psi$ . Parametrit estimoidaan suurimman uskottavuuden menetelmällä.

Uskottavuusfunktio on

$$L_g(\psi) = p(y|\psi) = p(y_1) \prod_{t=2}^n p(y_t|y_1, \dots, y_{t-1}).$$

Logaritminen uskottavuusfunktio on

$$\begin{aligned} \log L_g(\psi) &= \log p(y_1) + \sum_{t=2}^n \log p(y_t|y_1, \dots, y_{t-1}) \\ &= -\frac{np}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^n (\log |F_t| + v_t' F_t^{-1} v_t), \end{aligned}$$

missä  $v_t$  ja  $F_t$  saadaan Kalmanin suotimesta (lause 2.2). Uskottavuusfunktion lauseke on saatu johdettua tiedosta, että  $y_t|y_1, \dots, y_{t-1} \sim N(Z_t a_t, F_t)$ .

Suurimman uskottavuuden estimaattivektori  $\hat{\psi}$  on se parametrivektori  $\psi$ , joka maksimoi logaritmisen uskottavuusfunktion  $\log L_g(\psi)$ . Maksimointi suoritetaan numeerisesti.

### 2.1.6 Alustus

Kaikki edellä oleva teoria on olettanut, että ensimmäisen tilan  $\alpha_1$  jakauma  $N(a_1, P_1)$  tunnetaan eksaktisti eli  $a_1$  ja  $P_1$  ovat täysin tunnetut. Käytännön sovelluksissa näin ei kuitenkaan ole. Tilasarjan aloittamista ilman tunnettua alkutilan jakaumaa kutsutaan alustamiseksi.

Yksinkertaisin tapa suorittaa alustaminen on noudattaa Harvey'n ja Phillipsin (1979) [6] käyttämää approksimatiivista tekniikkaa. Tällöin tuntemattomat alkioit vektorissa  $a_1$  voidaan valita vapaasti ja tuntemattomat alkioit matriisissa  $P_1$  puolestaan valitaan hyvin suureksi. Menetelmä saattaa johtaa suuriin pyöristysvirheisiin, joten sen käyttöä ainoana menetelmänä ei suositella. Menetelmän yksinkertaisuudesta ja helposta ohjelmallisesta toteutuksesta johtuen sitä on kuitenkin käytetty tässä tutkielmassa.

## 2.2 Lineaarisista ja ei-gaussisista tila-avaruusmalleista

Tässä kappaleessa laajennetaan tila-avaruusmalleja siten, että mallien normaalisuudesta luovutaan. Yleinen lineaarinen ja ei-gaussinen tila-avaruusmalli voidaan määrittellä seuraavalla tavalla:

**Määritelmä 2.5** (Lineaarinen ja ei-gaussinen tila-avaruusmalli).

$$\begin{aligned} p(y_t | \alpha_1, \dots, \alpha_t, y_1, \dots, y_{t-1}) &= p(y_t | Z_t \alpha_t), \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, \quad \eta_t \sim p(\eta_t), \end{aligned}$$

kaikille  $t = 1, \dots, n$ , missä  $\eta_t$ :t ovat kaikki keskenään riippumattomia ja riippumattomia myös muuttujasta  $\alpha_1$ . Matriisi  $R_t$  koostuu  $m \times m$  -kokoisen yksikkömatriisin  $I_m$  sarakkeista. Merkitään lisäksi, että  $\theta_t := Z_t \alpha_t$ . Vektoreiden ja matriisien dimensiot on esitelty taulukossa 2.1.

Ensimmäinen yhtälö on jälleen nimeltään havaintoyhtälö; jälkimmäinen yhtälö on tilayhtälö. Muuttujaa  $\theta_t$  kutsutaan signaaliksi.

Erikoistapaus yleisestä lineaarisesta ja ei-gaussisesta tila-avaruusmallista on eksponenttisen perheen tila-avaruusmalli, jonka tilayhtälö on gaussinen.

**Määritelmä 2.6** (Eksponenttisen perheen tila-avaruusmalli). Havaintoyhtälö on muotoa

$$p(y_t | \theta_t) = \exp[y_t' \theta_t - b_t(\theta_t) + c_t(y_t)]$$

kaikille  $t = 1, \dots, n$ , missä  $\theta_t = Z_t \alpha_t$ . Havaintoyhtälössä  $b_t(\theta_t)$  on kahdesti differentioituva funktio ja  $c_t(y_t)$  on vain muuttujan  $y_t$  funktio. Tilayhtälö on muotoa

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \quad \eta_t \sim N(0, Q_t),$$

kaikille  $t = 1, \dots, n$ , missä  $\eta_t$ :t ovat kaikki keskenään riippumattomia ja riippumattomia myös muuttujasta  $\alpha_1$ . Lisäksi oletetaan, että  $\alpha_1 \sim N(a_1, P_1)$ . Matriisi  $R_t$  koostuu  $m \times m$  -kokoisen yksikkömatriisin  $I_m$  sarakkeista. Vektoreiden ja matriisien dimensiot on esitelty taulukossa 2.1.

Molemmissa määritelmässä 2.5 ja 2.6 matriisin  $R_t$  rakenteesta johtuen pätee, että  $\eta_t = R_t'(\alpha_{t+1} - T_t \alpha_t)$ .

Tavoitteena on laskea eksponenttisen perheen tila-avaruusmallille arvot  $\hat{\alpha}_t = E_p(\alpha_t | y)$  ja  $\text{var}_p(\alpha_t | y)$  sekä etsiä luottamusväli signaalille  $\theta_t$  ehdolla kaikki havainnot  $y$ . Tavoitteeseen päästään käyttämällä Monte Carlo -integrointia ja tärkeys-painotusta sekä linearisaatiota.

### 2.2.1 Monte Carlo -integrointi ja tärkeys-painotus

Olkoon  $p(\alpha | y)$  määritelmän 2.5 mukaisen lineaarisen ja ei-gaussisen mallin tilavektorin  $\alpha$  ehdollinen tiheysfunktio ehdolla kaikki havainnot  $y$ . Olkoon  $x$  jokin tilavektorin  $\alpha$  funktio. Halutaan selvittää odotusarvo

$$\bar{x} := E_p[x(\alpha) | y] = \int x(\alpha) p(\alpha | y) d\alpha.$$

Kun halutaan esimerkiksi laskea arvo  $E_p(\alpha_t | y)$ , valitaan tietenkin  $x(\alpha) = \alpha_t$ . Monte Carlo -integroinnissa simuloidaan  $s$  kappaletta arvoja jakaumasta  $p(\alpha | y)$  ja merkitään simuloituja arvoja  $\alpha^{(1)}, \dots, \alpha^{(s)}$ . Odotusarvon  $\bar{x}$  estimaattina  $\hat{x}$  voidaan käyttää keskiarvoa eli

$$\hat{x} = \hat{E}_p[x(\alpha) | y] = \frac{1}{s} \sum_{k=1}^s x(\alpha^{(k)}).$$

Jos havaintojen simuloiminen jakaumasta  $p(\alpha|y)$  on vaikeaa tai mahdotonta, käytetään apuna tärkeyspainotusta. Olkoon  $g(\alpha|y)$  jokin määritelmän 2.1 mukaisen lineaarisen ja gaussisen mallin tilavektorin  $\alpha$  ehdollinen jakauma ehdolla havainnot  $y$ . Tämän jakauman simulointihan onnistuu algoritmin 2 mukaisesti. Muokataan odotusarvon lauseketta seuraavasti:

$$\begin{aligned}\bar{x} = \mathbb{E}_p[x(\alpha)|y] &= \int x(\alpha)p(\alpha|y) \, d\alpha \\ &= \int x(\alpha) \frac{p(\alpha|y)}{g(\alpha|y)} g(\alpha|y) \, d\alpha \\ &= \mathbb{E}_g \left[ x(\alpha) \frac{p(\alpha|y)}{g(\alpha|y)} \middle| y \right].\end{aligned}$$

Odotusarvon estimaattia laskettaessa havainnot simuloidaan siis nyt jakaumasta  $g(\alpha|y)$ . Muokataan seuraavaksi jakaumien tiheysfunktioiden lausekkeet toiseen muotoon eli

$$p(\alpha|y) = \frac{p(\alpha, y)}{p(y)} = \frac{p(y|\alpha)p(\alpha)}{p(y)}$$

ja samoin

$$g(\alpha|y) = \frac{g(\alpha, y)}{g(y)} = \frac{g(y|\alpha)g(\alpha)}{g(y)}.$$

Sijoitetaan nämä odotusarvon  $\bar{x}$  lausekkeeseen ja saadaan

$$\bar{x} = \frac{g(y)}{p(y)} \mathbb{E}_g \left[ x(\alpha) \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right].$$

Muokataan lauseketta edelleen. Jos määritellään funktio  $x$  siten, että  $x(\alpha) = 1$ , niin saadaan

$$1 = \frac{g(y)}{p(y)} \mathbb{E}_g \left[ \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right].$$

Kun nyt otetaan osamäärä kahdesta edellisestä lausekkeesta, saadaan

$$\bar{x} = \frac{\bar{x}}{1} = \frac{\mathbb{E}_g \left[ x(\alpha) \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right]}{\mathbb{E}_g \left[ \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right]}.$$

Nyt tilat  $\alpha^{(1)}, \dots, \alpha^{(s)}$  simuloidaan jakaumasta  $g(\alpha|y)$ , jolloin odotusarvon  $\bar{x}$  estimaatti  $\hat{x}$  on

$$\hat{x} = \hat{\mathbb{E}}_p[x(\alpha)|y] = \frac{\hat{\mathbb{E}}_g \left[ x(\alpha) \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right]}{\hat{\mathbb{E}}_g \left[ \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right]} = \frac{\sum_{k=1}^s x(\alpha^{(k)}) \frac{p(y|\alpha^{(k)})p(\alpha^{(k)})}{g(y|\alpha^{(k)})g(\alpha^{(k)})}}{\sum_{k=1}^s \frac{p(y|\alpha^{(k)})p(\alpha^{(k)})}{g(y|\alpha^{(k)})g(\alpha^{(k)})}}. \quad (2.1)$$

Jos käsiteltävänä on määritelmän 2.6 mukainen eksponenttisen perheen malli, niin pätee  $g(\alpha) = p(\alpha)$ . Lisäksi yleisesti on voimassa  $p(y|\alpha) = p(y|\theta)$  ja  $g(y|\alpha) = g(y|\theta)$ . Tällöin odotusarvon  $\bar{x}$  estimaatti  $\hat{x}$  on

$$\hat{x} = \hat{\mathbb{E}}_p[x(\alpha)|y] = \frac{\hat{\mathbb{E}}_g \left[ x(\alpha) \frac{p(y|\theta)}{g(y|\theta)} \middle| y \right]}{\hat{\mathbb{E}}_g \left[ \frac{p(y|\theta)}{g(y|\theta)} \middle| y \right]} = \frac{\sum_{k=1}^s x(\alpha^{(k)}) \frac{p(y|\theta^{(k)})}{g(y|\theta^{(k)})}}{\sum_{k=1}^s \frac{p(y|\theta^{(k)})}{g(y|\theta^{(k)})}}, \quad (2.2)$$

missä  $\theta_t^{(k)} = Z_t \alpha_t^{(k)}$  kaikilla  $t = 1, \dots, n$ , ja  $\alpha^{(1)}, \dots, \alpha^{(s)}$  ovat simulointeja jakaumasta  $g(\alpha|y)$ . Otetaan lausekkeille (2.1) ja (2.2) käyttöön yhteinen merkintä

$$\hat{x} = \hat{E}_p[x(\alpha)|y] = \frac{\sum_{k=1}^s x(\alpha^{(k)}) z^{(k)}}{\sum_{k=1}^s z^{(k)}}, \quad (2.3)$$

missä siis yleisessä tapauksessa  $z^{(k)} = \frac{p(y|\alpha^{(k)})p(\alpha^{(k)})}{g(y|\alpha^{(k)})g(\alpha^{(k)})}$  ja eksponenttisen perheen tilanteessa  $z^{(k)} = \frac{p(y|\theta^{(k)})}{g(y|\theta^{(k)})}$ .

Odotusarvon  $\bar{x}$  estimaatin  $\hat{x}$  tarkkuuteen voidaan vaikuttaa simuloitien määrällä ja jakauman  $g(\alpha|y)$  valinnalla. Tarkkuus paranee, kun simuloitien määrä kasvaa. Lisäksi estimaatti on sitä tarkempi mitä samankaltaisempi simuloitava jakauma  $g(\alpha|y)$  on verrattuna jakaumaan  $p(\alpha|y)$ . Seuraavassa kappaleessa esitellään, miten jakauma  $g(\alpha|y)$  valitaan, jotta samankaltaisuus toteutuu.

## 2.2.2 Linearisaatio

Oletetaan aluksi, että  $p(\alpha|y)$  on määritelmän 2.5 mukaisesta lineaarisesta ja ei-gaussisesta mallista saatava jakauma. Tarkoituksena on löytää jakauman  $p(\alpha|y)$  kanssa samankaltainen, määritelmän 2.1 mukaisen lineaarisen ja gaussisen mallin perusteella muodostettava jakauma  $g(\alpha|y)$ . Merkitään  $\alpha^* = (\alpha', \alpha'_{n+1})'$ . Samankaltaisuus saadaan aikaan, kun jakauma  $g(\alpha|y)$  spesifioidaan siten, että jakaumilla  $g(\alpha^*|y)$  ja  $p(\alpha^*|y)$  on sama moodi. Merkitään moodia  $\check{\alpha}^*$ .

Etsitään aluksi jakauman  $g(\alpha^*|y)$  moodi. Moodi on se arvo  $\check{\alpha}^*$ , joka maksimoi jakauman  $g(\alpha^*|y)$ . Lineaarisen ja gaussisen mallin tapauksessa moodi löytyy gradientin nollakohdasta eli on ratkaistava yhtälö

$$\frac{\partial g(\alpha^*|y)}{\partial \alpha^*} = 0.$$

Koska logarimi on monotoninen muunnos, niin moodi saadaan myös yhtälön

$$\frac{\partial \log g(\alpha^*|y)}{\partial \alpha^*} = 0$$

ratkaisuna. Edelleen

$$\log g(\alpha^*|y) = \log \left[ \frac{g(\alpha^*, y)}{g(y)} \right] = \log g(\alpha^*, y) - \log g(y),$$

joten moodi saadaan ratkaisemalla yhtälö

$$\frac{\partial \log g(\alpha^*, y)}{\partial \alpha^*} = 0,$$

sillä  $\frac{\partial \log g(y)}{\partial \alpha^*} = 0$ . Logaritminen yhteisjakauma on

$$\begin{aligned} \log g(\alpha^*, y) &= \text{vakio} - \frac{1}{2}(\alpha_1 - a_1)' P_1^{-1}(\alpha_1 - a_1) \\ &\quad - \frac{1}{2} \sum_{t=1}^n (\alpha_{t+1} - T_t \alpha_t)' R_t Q_t^{-1} R_t' (\alpha_{t+1} - T_t \alpha_t) \\ &\quad - \frac{1}{2} \sum_{t=1}^n (y_t - Z_t \alpha_t)' H_t^{-1} (y_t - Z_t \alpha_t). \end{aligned}$$

Kun lasketaan osittaisderivaatat, saadaan yhtälöt

$$\begin{aligned} (d_t - 1)P_1^{-1}(\alpha_1 - a_1) - d_t R_{t-1} Q_{t-1}^{-1} R_{t-1}' (\alpha_t - T_{t-1} \alpha_{t-1}) \\ + T_t' R_t Q_t^{-1} R_t' (\alpha_{t+1} - T_t \alpha_t) + Z_t' H_t^{-1} (y_t - Z_t \alpha_t) = 0, \end{aligned} \quad (2.4)$$

kaikille  $t = 1, \dots, n$ , missä  $d_1 = 0$  ja  $d_2 = \dots = d_n = 1$ , ja lisäksi yhtälö

$$R_n Q_n R_n' (\alpha_{n+1} - T_n \alpha_n) = 0. \quad (2.5)$$

Näiden yhtälöiden ratkaisu on siis moodi  $\check{\alpha}^*$ . Lineaarisen ja gaussisen mallin tapauksessa moodi  $\check{\alpha}^*$  on sama kuin odotusarvo  $\hat{\alpha}^* = E(\alpha^*|y)$ , joka puolestaan saadaan ratkaistua Kalmanin suotimen (lause 2.2) ja tilojen tasoitusrekursion (lause 2.3) avulla.

Tarkastellaan seuraavaksi jakauman  $p(\alpha^*|y)$  moodin  $\check{\alpha}^*$  laskentaa. Oletetaan, että lineaarinen ja ei-gaussinen malli on siinä mielessä hyvin käyttäytyvä, että moodi saadaan ratkaisemalla yhtälö

$$\frac{\partial p(\alpha^*|y)}{\partial \alpha^*} = 0.$$

Kuten gaussisen mallin tapauksessa, moodi löytyy myös yhtälön

$$\frac{\partial \log p(\alpha^*, y)}{\partial \alpha^*} = 0$$

ratkaisuna. Nyt pätee

$$\log p(\alpha^*, y) = \text{vakio} + \log p(\alpha_1) - \sum_{t=1}^n [q_t(\eta_t) + h_t(y_t|\theta_t)],$$

missä  $q_t(\eta_t) = -\log p(\eta_t)$ ,  $\eta_t = R_t' (\alpha_{t+1} - T_t \alpha_t)$ ,  $h_t(y_t|\theta_t) = -\log p(y_t|\theta_t)$  ja  $\theta_t = Z_t \alpha_t$ . Kun lasketaan osittaisderivaatat, saadaan yhtälöt

$$\begin{aligned} (1 - d_t) \frac{\partial \log p(\alpha_1)}{\partial \alpha_1} - d_t R_{t-1} \frac{\partial q_{t-1}(\eta_{t-1})}{\partial \eta_{t-1}} \\ + T_t' R_t \frac{\partial q_t(\eta_t)}{\partial \eta_t} - Z_t' \frac{\partial h_t(y_t|\theta_t)}{\partial \theta_t} = 0, \end{aligned} \quad (2.6)$$

kaikille  $t = 1, \dots, n$ , missä  $d_1 = 0$  ja  $d_2 = \dots = d_n = 1$ , ja lisäksi yhtälö

$$R_n \frac{\partial q_n(\eta_n)}{\partial \eta_n} = 0. \quad (2.7)$$



Tehdään seuraavaksi lisäoletus, että  $p(\alpha^*|y)$  saadaan mallista, jossa tilayhtälö on gaussinen eli muotoa  $\alpha_{t+1} = T_t\alpha_t + R_t\eta_t$ , missä  $\eta_t \sim N(0, Q_t)$ . Näin on myös eksponenttisen perheen tapauksessa. Tällöin yhtälö (2.6) saadaan muotoon

$$(d_t - 1)P_1^{-1}(\alpha_1 - a_1) - d_t R_{t-1} Q_{t-1}^{-1} R'_{t-1}(\alpha_t - T_{t-1}\alpha_{t-1}) \\ + T'_t R_t Q_t^{-1} R'_t(\alpha_{t+1} - T_t\alpha_t) - Z'_t \frac{\partial h_t(y_t|\theta_t)}{\partial \theta_t} = 0, \quad (2.8)$$

ja yhtälö (2.7) muotoon

$$R_n Q_n R'_n(\alpha_{n+1} - T_n\alpha_n) = 0. \quad (2.9)$$

Tarkoituksena on nyt saada yhtälöt (2.8) ja (2.9) mahdollisimman samaan muotoon kuin yhtälöt (2.4) ja (2.5), jotta nähdään, miten jakauman  $g(\alpha|y)$  taustalla oleva lineaarinen ja gaussinen tila-avaruusmalli on spesifioitava, jotta jakaumalla  $g(\alpha^*|y)$  olisi sama moodi kuin jakaumalla  $p(\alpha^*|y)$ . Yhtälöt ovat muuten samaa muotoa, mutta yhtälöiden (2.8) ja (2.4) viimeiset termit eroavat. Tarkastellaan siis termejä

$$Z'_t H_t^{-1}(y_t - Z_t\alpha_t) \quad (2.10)$$

ja

$$-Z'_t \frac{\partial h_t(y_t|\theta_t)}{\partial \theta_t}. \quad (2.11)$$

Oletetaan, että moodille  $\check{\alpha}^*$  on saatu jonkinlainen arvio  $\check{\alpha}^{*(i)}$ . Tällöin merkitään  $\check{\theta}_t^{(i)} = Z_t \check{\alpha}_t^{*(i)}$  kaikilla  $t = 1, \dots, n$ . Määritellään

$$\dot{h}_t^{(i)} = \left. \frac{\partial h_t(y_t|\theta_t)}{\partial \theta_t} \right|_{\theta_t = \check{\theta}_t^{(i)}}$$

ja

$$\ddot{h}_t^{(i)} = \left. \frac{\partial^2 h_t(y_t|\theta_t)}{\partial \theta_t \partial \theta'_t} \right|_{\theta_t = \check{\theta}_t^{(i)}}.$$

Taylorin kehitelmän perusteella saadaan approksimaatio

$$\frac{\partial h_t(y_t|\theta_t)}{\partial \theta_t} \approx \dot{h}_t^{(i)} + \ddot{h}_t^{(i)}(\theta_t - \check{\theta}_t^{(i)}).$$

Kun sijoitetaan tämä approksimaatio lausekkeeseen (2.11), saadaan se muotoon

$$\begin{aligned} & -Z'_t \left( \dot{h}_t^{(i)} + \ddot{h}_t^{(i)} (\theta_t - \check{\theta}_t^{(i)}) \right) \\ & = -Z'_t \left( \dot{h}_t^{(i)} + \ddot{h}_t^{(i)} \theta_t - \ddot{h}_t^{(i)} \check{\theta}_t^{(i)} \right) \\ & = Z'_t \ddot{h}_t^{(i)} \left( \check{\theta}_t^{(i)} - \left[ \ddot{h}_t^{(i)} \right]^{-1} \dot{h}_t^{(i)} - \theta_t \right) \\ & = Z'_t \ddot{h}_t^{(i)} \left( \check{\theta}_t^{(i)} - \left[ \ddot{h}_t^{(i)} \right]^{-1} \dot{h}_t^{(i)} - Z_t \alpha_t \right). \end{aligned}$$

Kun valitaan

$$\check{H}_t^{(i)} = \left[ \ddot{h}_t^{(i)} \right]^{-1}$$

ja

$$\check{y}_t^{(i)} = \check{\theta}_t^{(i)} - \left[ \ddot{h}_t^{(i)} \right]^{-1} \dot{h}_t^{(i)},$$

niin termi (2.11) saadaan muotoon

$$Z_t' \left[ \check{H}_t^{(i)} \right]^{-1} \left( \check{y}_t^{(i)} - Z_t \alpha_t \right),$$

joka on samaa muotoa kuin termi (2.10). Saadaan lineaarinen ja gaussinen malli

$$\begin{aligned} \check{y}_t^{(i)} &= Z_t \alpha_t + \epsilon_t, & \epsilon_t &\sim N(0, \check{H}_t^{(i)}), \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, & \eta_t &\sim N(0, Q_t). \end{aligned} \quad (2.12)$$

Merkitään mallin (2.12) perusteella saatavaa havainnoilla ehdollistettua tilojen jakaumaa  $g(\alpha|y)^{(i)}$ . Jakauman  $g(\alpha^*|y)^{(i)}$  moodi, jota merkitään  $\check{\alpha}^{*(i+1)}$ , voidaan laskea Kalmanin suotimen (lause 2.2) ja tilojen tasoitusrekursion (lause 2.3) avulla eli  $\check{\alpha}^{*(i+1)} = E_g(\alpha^*|\check{y}^{(i)})$ , sillä kyseessä on lineaarinen ja gaussinen malli, jolloin havainnoilla ehdollistettu moodi ja odotusarvo ovat samat. Jos moodille edellä tehty arvio  $\check{\alpha}^{*(i)}$  on hyvä, niin on voimassa  $\check{\alpha}^{*(i)} \approx \check{\alpha}^{*(i+1)}$  ja valitaan  $g(\alpha|y)^{(i)}$  lopulliseksi jakaumaksi, jota myöhemmin simuloidaan. Jos näin ei ole, valitaan  $\check{\alpha}^{*(i+1)}$  uudeksi arvioksi moodille  $\check{\alpha}^*$ , toistetaan prosessi ja saadaan uusi jakauma  $g(\alpha|y)^{(i+1)}$ . Tätä jatketaan, kunnes peräkkäisillä malleilla saadut moodit ovat riittävän lähellä toisiaan.

Tehdään vielä lisäoletus, että käsiteltävänä on määritelmän 2.6 mukainen eksponenttisen perheen malli. Tällöin

$$h_t(y_t|\theta_t) = -\log p(y_t|\theta_t) = -[y_t' \theta_t - b_t(\theta_t) + c_t(y_t)].$$

Määritellään

$$\dot{b}_t^{(i)} = \left. \frac{\partial b_t(y_t|\theta_t)}{\partial \theta_t} \right|_{\theta_t = \check{\theta}_t^{(i)}}$$

ja

$$\ddot{b}_t^{(i)} = \left. \frac{\partial^2 b_t(y_t|\theta_t)}{\partial \theta_t \partial \theta_t'} \right|_{\theta_t = \check{\theta}_t^{(i)}},$$

jolloin  $\dot{h}_t^{(i)} = \dot{b}_t^{(i)} - y_t$  ja  $\ddot{h}_t^{(i)} = \ddot{b}_t^{(i)}$ . Täten eksponenttisen perheen tapauksessa saadaan

$$\check{H}_t^{(i)} = \left[ \ddot{b}_t^{(i)} \right]^{-1}$$

ja

$$\check{y}_t^{(i)} = \check{\theta}_t^{(i)} - \left[ \ddot{b}_t^{(i)} \right]^{-1} \left( \dot{b}_t^{(i)} - y_t \right).$$

### 2.2.3 Tasoitus tiloille

Oletetaan, että käsiteltävänä on määritelmän 2.5 mukainen yleinen lineaarinen ja ei-gaussinen tila-avaruusmalli. Halutaan siis laskea arvot  $\hat{\alpha}_t = E_p(\alpha_t|y)$  ja  $\text{var}_p(\alpha_t|y)$ .

Muodostetaan aluksi jakauma  $g(\alpha|y)$  kappaleen 2.2.2 mukaisesti. Simuloidaan arvot  $\alpha^{(1)}, \dots, \alpha^{(s)}$  tästä jakaumasta algoritmin 2 mukaisesti. Arvot  $\hat{\alpha}_t$  saadaan laskettua, kun valitaan  $x(\alpha) = \alpha_t$  lausekkeessa (2.3), jolloin saadaan

$$\hat{\alpha}_t \approx \frac{\sum_{k=1}^s \alpha_t^{(k)} z^{(k)}}{\sum_{k=1}^s z^{(k)}}.$$

Arvojen  $\text{var}_p(\alpha_t|y) = E_p(\alpha_t \alpha_t'|y) - \hat{\alpha}_t \hat{\alpha}_t'$  saamiseksi on laskettava termi  $E_p(\alpha_t \alpha_t'|y)$ , joka saadaan valitsemalla  $x(\alpha) = \alpha_t \alpha_t'$  kaavassa 2.3. Tällöin saadaan

$$E_p(\alpha_t \alpha_t'|y) \approx \frac{\sum_{k=1}^s \alpha_t^{(k)} \alpha_t^{(k)'} z^{(k)}}{\sum_{k=1}^s z^{(k)}},$$

josta edelleen

$$\text{var}_p(\alpha_t|y) \approx \frac{\sum_{k=1}^s \alpha_t^{(k)} \alpha_t^{(k)'} z^{(k)}}{\sum_{k=1}^s z^{(k)}} - \hat{\alpha}_t \hat{\alpha}_t'.$$

### 2.2.4 Luottamusväli signaaleille

Oletetaan, että käsiteltävänä on määritelmän 2.5 mukainen yleinen lineaarinen ja ei-gaussinen tila-avaruusmalli. Tarkoituksena on saada signaalille  $\theta_t$  luottamusväli ehdolla kaikki havainnot  $y$ .

Muodostetaan jälleen aluksi jakauma  $g(\alpha|y)$  kappaleen 2.2.2 mukaisesti ja simuloidaan arvot  $\alpha^{(1)}, \dots, \alpha^{(s)}$  tästä jakaumasta algoritmin 2 mukaisesti.

Nyt pitäisi etsiä  $100a$  prosentin kvantiili eli arvo  $\theta_{t,a}$  siten, että  $P(\theta_t < \theta_{t,a}|y) \approx a$ . Koska  $\theta_t = Z_t \alpha_t$ , niin määritellään indikaattorifunktio  $I_{t,a}(\alpha)$  siten, että

$$I_{t,a}(\alpha) = \begin{cases} 1 & \text{jos } Z_t \alpha_t < \theta_{t,a} \\ 0 & \text{muulloin.} \end{cases}$$

Nyt saadaan, että

$$P(\theta_t < \theta_{t,a}|y) = E[I_{t,a}(\alpha)|y].$$

Täten valitsemalla lausekkeessa (2.3)  $x(\alpha) = I_{t,a}(\alpha)$  voidaan arvolle  $P(\theta_t < \theta_{t,a}|y)$  laskea arvio kaavalla

$$P(\theta_t < \theta_{t,a}|y) \approx \frac{\sum_{k=1}^s I_{t,a}(\alpha^{(k)}) z^{(k)}}{\sum_{k=1}^s z^{(k)}}.$$

Järjestetään nyt arvot  $\theta_t^{(k)}$ ,  $\theta_t^{(k)} = Z_t \alpha_t^{(k)}$ , suuruusjärjestykseen pienimmästä suurimpaan ja arvot  $z^{(k)}$  edellistä vastaavaan järjestykseen. Merkitään järjestettyjä havaintoja  $\theta^{[k]}$  ja  $z^{[k]}$ . Nyt  $100a$  prosentin kvantiilin signaalille  $\theta_t$  antaa  $\theta_t^{[b]}$ , joka on valittu siten että

$$\frac{\sum_{j=1}^b z^{[j]}}{\sum_{j=1}^s z^{[j]}} \approx a.$$

100a prosentin luottamusväli  $(\theta_{t,(1-a)/2}, \theta_{t,(1+a)/2})$  on nyt

$$(\theta_t^{[b_1]}, \theta_t^{[b_2]}),$$

missä  $\theta_t^{[b_1]}$  on valittu siten, että

$$\frac{\sum_{j=1}^{b_1} z^{[k]}}{\sum_{j=1}^s z^{[k]}} \approx (1-a)/2$$

ja  $\theta_t^{[b_2]}$  siten, että

$$\frac{\sum_{j=1}^{b_2} z^{[k]}}{\sum_{j=1}^s z^{[k]}} \approx (1+a)/2.$$

## 2.2.5 Parametrien estimointi

Mallin määräävät matriisit  $Z_t$ ,  $T_t$ ,  $R_t$  ja  $Q_t$  voivat riippua tuntemattomista parametreista. Kuten lineaarisen ja gaussisen mallin tapauksessa kootaan parametrit parametrivektoriksi  $\psi$ . Parametrit estimoidaan suurimman uskottavuuden menetelmällä.

Uskottavuusfunktio on

$$L_p(\psi) = p(y|\psi) = \int p(\alpha, y) d\alpha.$$

Etsitään jakaumaa  $p(\alpha|y)$  approksimoiva, lineaarisen ja gaussisen mallin perusteella saatava ehdollinen jakauma  $g(\alpha|y)$  kappaleen 2.2.2 mukaisesti. Muokataan nyt uskottavuusfunktion lauseketta hieman:

$$\begin{aligned} L_p(\psi) &= \int \frac{p(\alpha, y)}{g(\alpha|y)} g(\alpha|y) d\alpha \\ &= \int \frac{p(\alpha, y)}{g(\alpha, y)/g(y)} g(\alpha|y) d\alpha \\ &= g(y) \int \frac{p(\alpha, y)}{g(\alpha, y)} g(\alpha|y) d\alpha \\ &= L_g(\psi) \mathbb{E}_g \left[ \frac{p(\alpha, y)}{g(\alpha, y)} \middle| y \right] \\ &= L_g(\psi) \mathbb{E}_g \left[ \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right], \end{aligned}$$

missä  $L_g(\psi)$  on approksimoivan lineaarisen ja gaussisen mallin uskottavuusfunktio, joka on määritelty kappaleessa 2.1.5. Edelleen logaritminen uskottavuusfunktio on

$$\log L_p(\psi) = \log L_g(\psi) + \log \mathbb{E}_g \left[ \frac{p(y|\alpha)p(\alpha)}{g(y|\alpha)g(\alpha)} \middle| y \right].$$

Tästä eteenpäin tässä kappaleessa oletetaan, että mallina on määritelmän 2.6 mukainen eksponenttisen perheen malli. Tällöin pätee  $p(\alpha) = g(\alpha)$ , ja logaritminen uskottavuusfunktio saa muodon

$$\log L_p(\psi) = \log L_g(\psi) + \log \mathbb{E}_g \left[ \frac{p(y|\theta)}{g(y|\theta)} \middle| y \right].$$

Merkitään

$$w = w(\theta) = \frac{p(y|\theta)}{g(y|\theta)},$$

ja tällöin

$$\log L_p(\psi) = \log L_g(\psi) + \log E_g(w|y). \quad (2.13)$$

Termille  $E_g(w|y)$  saadaan estimaatti  $\hat{E}_g(w|y)$  Monte Carlo -integroinnilla samaan tapaan kuin kappaleessa 2.2.1:

$$\hat{E}_g(w|y) = \hat{E}_g \left[ \frac{p(y|\theta)}{g(y|\theta)} \middle| y \right] = \frac{1}{s} \sum_{k=1}^s \frac{p(y|\theta^{(k)})}{g(y|\theta^{(k)})},$$

missä  $\theta_t^{(k)} = Z_t \alpha_t^{(k)}$  kaikilla  $t = 1, \dots, n$ , ja  $\alpha^{(1)}, \dots, \alpha^{(s)}$  ovat simulointeja jakaumasta  $g(\alpha|y)$ . Simuloinnit tehdään algoritmin 2 mukaisesti. Korvaamalla lausekkeessa (2.13) termi  $E_g(w|y)$  termillä  $\hat{E}_g(w|y)$  voidaan määrittellä estimoitu logaritminen uskottavuusfunktio seuraavasti:

$$\log \hat{L}_p(\psi) = \log L_g(\psi) + \log \hat{E}_g(w|y). \quad (2.14)$$

Suurimman uskottavuuden estimaatit saadaan maksimoimalla estimoitu logaritminen uskottavuusfunktio  $\log \hat{L}_p(\psi)$  parametrivektorin  $\psi$  suhteen. Maksimointi voidaan suorittaa sopivalla numeerisella menetelmällä.

Numeeriset menetelmät vaativat yleensä hyvän alkuarvon muuttujalle, jonka suhteen halutaan maksimoida. Durbin & Koopman (1997) [2] esittelevät niin sanotun logaritmissen nollauskottavuusfunktion, joka maksimoimalla saadaan parametrivektorille  $\psi$  hyvä alkuarvo. Seuraavassa kappaleessa kerrotaan, kuinka tämä logaritminen nollauskottavuusfunktio muodostetaan ja kuinka pelkästään sen avulla voidaan saada käytännössä riittävän tarkkoja estimaatteja parametreille.

## Logaritminen nollauskottavuusfunktio

Logaritminen nollauskottavuusfunktio saadaan, kun logaritmissen uskottavuusfunktion kaavassa (2.13) oleva funktio  $w$  korvataan Taylor-approksimaatiolla  $w_T$ . Logaritminen nollauskottavuusfunktio on siis

$$\log L_0(\psi) := \log L_g(\psi) + \log E_g(w_T|y). \quad (2.15)$$

Lasketaan aluksi kappaleen 2.2.2 mukaisesti muodostetun approksimoivan lineaarisen ja gaussisen mallin tilojen tasoitettut arvot

$$\hat{\alpha} = E_g(\alpha|y),$$

ja edelleen merkitään

$$\hat{\theta}_t = Z_t \hat{\alpha}_t,$$

kaikilla  $t = 1, \dots, n$ . Kehitetään funktiolle  $w$  Taylorin sarja pisteen  $\hat{\theta}$  suhteen, jolloin saadaan

$$w(\theta) = \sum_{j=0}^{\infty} \frac{1}{j!} d^j w(\hat{\theta}; \theta - \hat{\theta}),$$

missä  $d^j w(\hat{\theta}; \theta - \hat{\theta})$  on muutosta  $\theta - \hat{\theta}$  vastaava funktion  $w(\theta)$   $j$ . kertaluvun kokonaisdifferentiaali pisteessä  $\hat{\theta}$ . Oletetaan, että neljännen asteen Taylorin polynomi

$$w_T = w_T(\theta) = \sum_{j=0}^4 \frac{1}{j!} d^j w(\hat{\theta}; \theta - \hat{\theta})$$

on tarpeeksi tarkka approksimoimaan funktiota  $w$ .

Seuraavaksi muodostetaan tarvittavat kokonaisdifferentiaalit ja niiden perusteella Taylorin polynomi  $w_T$ . Määritellään

$$l = l(\theta) = \log w(\theta).$$

Muokataan termi  $l$  toiseen muotoon:

$$\begin{aligned} l(\theta) &= \log w(\theta) = \log \frac{p(y|\theta)}{g(y|\theta)} = \log p(y|\theta) - \log g(y|\theta) \\ &= \log \left[ \prod_{t=1}^n p(y_t|\theta_t) \right] - \log \left[ \prod_{t=1}^n g(y_t|\theta_t) \right] \\ &= \sum_{t=1}^n [\log p(y_t|\theta_t) - \log g(y_t|\theta_t)]. \end{aligned}$$

Laskemalla saadaan, että

$$\frac{\partial^{d_1+\dots+d_k} l(\theta)}{\partial^{d_1} \theta_1 \dots \partial^{d_k} \theta_k} = 0,$$

kun  $i$  ja  $j$ ,  $i \neq j$ , ovat sellaisia, että  $d_i > 0$  ja  $d_j > 0$ . Otetaan käyttöön merkinnät

$$\begin{aligned} w'_t &= w'_t(\theta) = \frac{\partial w(\theta)}{\partial \theta_t}, & w''_t &= w''_t(\theta) = \frac{\partial^2 w(\theta)}{\partial^2 \theta_t}, & \dots &, \\ l'_t &= l'_t(\theta) = \frac{\partial l(\theta)}{\partial \theta_t}, & l''_t &= l''_t(\theta) = \frac{\partial^2 l(\theta)}{\partial^2 \theta_t}, & \dots &. \end{aligned}$$

Nyt funktion  $w$  derivaatat voidaan ilmaista funktion  $l$  avulla:

$$\begin{aligned} w'_t &= w l'_t, \\ w''_t &= w \left[ l''_t + (l'_t)^2 \right], \\ w'''_t &= w \left[ l'''_t + 3l'_t l''_t + (l'_t)^3 \right], \\ w''''_t &= w \left[ l''''_t + 4l'_t l'''_t + 3(l''_t)^2 + 6l''_t (l'_t)^2 + (l'_t)^4 \right]. \end{aligned}$$

Edellä olevasta seuraa, että

$$\frac{\partial^{d_1+\dots+d_k} w(\theta)}{\partial^{d_1} \theta_1 \dots \partial^{d_k} \theta_k} = 0,$$

kun  $i$  ja  $j$ ,  $i \neq j$ , ovat sellaisia, että  $d_i > 0$  ja  $d_j > 0$ . Kokonaisdifferentiaalien muodostamisessa voidaan siis jättää huomiotta kaikki useamman kuin yhden muuttujan

suhteen otetut derivaatat ja tarkastellaan vain derivaattoja  $w'_t$ ,  $w''_t$ ,  $w'''_t$  ja  $w''''_t$ . Merkitään

$$\hat{w} = w(\hat{\theta})$$

ja

$$\hat{w}'_t = w'_t(\hat{\theta}), \quad \hat{w}''_t = w''_t(\hat{\theta}), \quad \dots$$

Samalla tavalla merkitään

$$\hat{l} = l(\hat{\theta})$$

ja

$$\hat{l}'_t = l'_t(\hat{\theta}), \quad \hat{l}''_t = l''_t(\hat{\theta}), \quad \dots$$

Kappaleessa 2.2.2 tehdyn linearisaation perusteella saadaan, että

$$\left. \frac{\partial \log g(y_t | \theta_t)}{\partial \theta_t} \right|_{\theta_t = \hat{\theta}_t} = \left. \frac{\partial \log p(y_t | \theta_t)}{\partial \theta_t} \right|_{\theta_t = \hat{\theta}_t},$$

$$\left. \frac{\partial^2 \log g(y_t | \theta_t)}{\partial^2 \theta_t} \right|_{\theta_t = \hat{\theta}_t} = \left. \frac{\partial^2 \log p(y_t | \theta_t)}{\partial^2 \theta_t} \right|_{\theta_t = \hat{\theta}_t},$$

joten  $\hat{l}'_t = 0$  ja  $\hat{l}''_t = 0$ . Tästä edelleen saadaan, että  $\hat{w}'_t = 0$ ,  $\hat{w}''_t = 0$ ,  $\hat{w}'''_t = \hat{w} \hat{l}'''_t$  ja  $\hat{w}''''_t = \hat{w} \hat{l}''''_t$ . Nyt saadaan muodostettua tarvittavat kokonaisdifferentiaalit

$$d^0 w(\hat{\theta}; \theta - \hat{\theta}) = \hat{w},$$

$$d^1 w(\hat{\theta}; \theta - \hat{\theta}) = 0,$$

$$d^2 w(\hat{\theta}; \theta - \hat{\theta}) = 0,$$

$$d^3 w(\hat{\theta}; \theta - \hat{\theta}) = \hat{w} \sum_{t=1}^n \hat{l}'''_t (\theta_t - \hat{\theta}_t)^3,$$

$$d^4 w(\hat{\theta}; \theta - \hat{\theta}) = \hat{w} \sum_{t=1}^n \hat{l}''''_t (\theta_t - \hat{\theta}_t)^4,$$

jolloin Taylor-approksimaatio saa muodon

$$w_T = \hat{w} \left[ 1 + \frac{1}{6} \sum_{t=1}^n \hat{l}'''_t (\theta_t - \hat{\theta}_t)^3 + \frac{1}{24} \sum_{t=1}^n \hat{l}''''_t (\theta_t - \hat{\theta}_t)^4 \right].$$

Seuraavana tehtävänä on laskea odotusarvo  $E_g(w_T | y)$ :

$$E_g(w_T | y) = \hat{w} \left[ 1 + \frac{1}{6} \sum_{t=1}^n \hat{l}'''_t E_g \left[ (\theta_t - \hat{\theta}_t)^3 | y \right] + \frac{1}{24} \sum_{t=1}^n \hat{l}''''_t E_g \left[ (\theta_t - \hat{\theta}_t)^4 | y \right] \right]. \quad (2.16)$$

Muuttujan  $\theta_t$  jakauma ehdolla muuttuja  $y$  on normaalin, kun taustalla on määritelmän 2.1 mukainen lineaarinen ja gaussinen malli. Tiedetään, että  $E_g(\theta_t | y) = \hat{\theta}_t$  ja lisäksi huomataan, että

$$\text{var}_g(\theta_t | y) = \text{var}_g(y_t - \epsilon_t | y) = \text{var}_g(\epsilon_t | y),$$

joka saadaan häiriötermien tasoitusrekursiolla (lause 2.4). Merkitään, että

$$(\theta_t|y)_g \sim N(\hat{\theta}_t, \text{var}_g(\epsilon_t|y)).$$

Koska kyseessä on normaalijakauma, niin kolmannelle keskusmomentille pätee

$$E_g \left[ (\theta_t - \hat{\theta}_t)^3 | y \right] = 0$$

ja neljännelle keskusmomentille puolestaan pätee

$$E_g \left[ (\theta_t - \hat{\theta}_t)^4 | y \right] = 3 [\text{var}_g(\epsilon_t|y)]^2.$$

Sijoitetaan saadut tulokset lausekkeeseen (2.16), jolloin saadaan

$$E_g(w_T|y) = \hat{w} \left[ 1 + \frac{1}{8} \sum_{t=1}^n \hat{l}_t''' \text{var}_g(\epsilon_t|y)^2 \right]. \quad (2.17)$$

Sijoitetaan lopuksi saatu tulos (2.17) lausekkeeseen (2.15), jolloin logaritminen nollauskottavuusfunktio saa muodon

$$\log L_0(\psi) = \log L_g(\psi) + \log \hat{w} + \log \left( 1 + \frac{1}{8} \sum_{t=1}^n \hat{l}_t''' [\text{var}_g(\epsilon_t|y)]^2 \right). \quad (2.18)$$

Durbin & Koopman (1997) [2] ovat empiirisesti tutkineet tämän logaritmisen nollauskottavuusfunktion (2.18) hyvyttä verrattuna estimoituun logaritmiseen uskottavuusfunktioon (2.14). He osoittavat, että jo logaritmisella nollauskottavuusfunktioilla saadaan hyvin tarkka arvio estimaateille, eivätkä estimaatit paljon muutu vaikka siirytään käyttämään estimoitua logaritmista uskottavuusfunktioita. Tällä perusteella tässä tutkielmassa estimoinnit on tehty käyttäen pelkästään logaritmista nollauskottavuusfunktioita.



# Luku 3

## Sovellus alkoholikuolleisuusaineistoon

Tarkoituksena on tutkia alkoholikuolleisuuden kehitystä Suomessa vuosina 1969–2007. Tutkimuskysymyksenä on ensinnäkin alkoholikuolleisuuden yleinen kehitys Suomessa sekä koko väestön osalta että eri sukupuoli- ja ikäryhmissä. Lisäksi tarkastellaan, voidaanko alkoholikuolleisuuden ja alkoholiolojen muutosten välillä nähdä jonkinlaista yhteyttä. Alkoholiolojen muutoksilla tarkoitetaan esimerkiksi muutoksia alkoholilainsäädännössä. Jos jokin oleellinen yhteys havaitaan, tästä havainnosta voi olla hyötyä pohdittaessa, miten alkoholipolitiikkaa pitää tulevaisuudessa kehittää, jotta alkoholikuolleisuus pienenesi.

### 3.1 Yhteiskunnallinen tausta

Seuraavassa esitellään alkoholikuolleisuuteen liittyviä yhteiskunnallisia taustoja sekä alkoholiolojen historian että suomalaisten alkoholiin liittyvien asenteiden kautta.

#### 3.1.1 Alkoholiolojen historiasta Suomessa

Seuraavan historiakatsauksen rungon kokoamisessa on käytetty lähteenä Stakesin julkaisua *Päihdetilastollinen vuosikirja 2007* [11]. Lakeihin liittyvien yksityiskohtien lähteenä on käytetty Internetissä olevaa FINLEX® – Valtion säädöstietopankki -sivustoa [5].

Itsenäisen Suomen alkoholihistoria voidaan karkeasti jakaa kolmeen osaan: kieltoain aika vuosina 1919–1932, ankaran holhouksen aika vuosina 1932–1968 ja vuonna 1969 voimaan tulleen keskiolutlain jälkeinen aika. Tämä tutkimus keskittyy siis selvittämään alkoholikuolleisuuden muutoksia aikana, jolloin keskiolutlaki on ollut voimassa.

Tammikuun alussa vuonna 1969 voimaan astuneen keskiolutlain tärkein sisältö oli lupa myydä III-veroluokan olutta elintarvikeliikkeissä. Lisäksi alkoholin myynti tuli lailliseksi myös maaseudulla, kun se oli ennen ollut sallittua vain kaupungeissa ja kauppaloissa.

Keskiolutlain voimaantulon jälkeen suomalaisessa alkoholipolitiikassa on tapahtunut monia muutoksia. Vuonna 1977 kaikki alkoholimainonta kiellettiin lailla. Vuon-

na 1994 alkoholin verotuksen perusteena olleet veroluokat poistettiin ja verotuksessa siirryttiin määräveroon.

Vuonna 1995 Suomi liittyi Euroopan Unioniin. Tällöin Alkolla ollut alkoholijuomien tuonti-, vienti-, valmistus- ja tukkumyyntimonopoli purettiin. Alkolle jäi yli 4.7-tilavuusprosenttisten alkoholijuomien vähittäismyyntimonopoli, mutta kaikkien alle 4.7-prosenttisten käymisteitse valmistettujen alkoholijuomien myynti sallittiin elintarvikeliikkeissä. Ruokakauppojen valikoimiin tulivat siis esimerkiksi alle 4.7-prosenttiset siiderit, kun ennen oli saanut myydä vain samaprosenttista olutta. Myös alle 22-prosenttisten mietojen alkoholijuomien mainonta sallittiin tietyin rajoituksin. Laissa lievennettiin lisäksi alkoholin tuontiin liittyviä tullimääräyksiä; esimerkiksi alkoholin tuontia rajoittaneet aikarajat poistettiin. Vuonna 1996 tullimääräyksiä jälleen kiristettiin, kun aikarajoituksia palautettiin; alle 20 tuntia kestävältä EU-alueen ulkopuolelle suuntautuvalta matkalta, johon lukeutuivat myös matkat Ahvenanmaalle, ei saanut tuoda alkoholia lainkaan.

Euroopan Unioniin liittymisestä lähtien Suomessa olivat olleet voimassa niin sanottu siirtymäajan rajoitukset, jotka olivat rajoittaneet alkoholin tuontia Suomeen EU:n sisältä määrällisillä rajoituksilla. Näitä rajoituksia vähitellen lievennettiin ja 1.1.2004 Suomi luopui kokonaan siirtymäajan rajoituksista EU:n sisällä. Tämä tarkoittaa sitä, että EU:n alueelta saa tuoda omaan käyttöön rajoituksetta alkoholijuomia. Viro liittyi Euroopan Unioniin 1.5.2004, jolloin näin ollen tuonti myös sieltä vapautui. Tuontia EU:n ulkopuolelta, mukaan lukien Ahvenanmaalta, rajoitetaan edelleen EU:n yhteisillä säännöillä ja lisäksi Suomen kansallisilla 20-tunnin aikarajoituksella ja oluen tuontirajoituksilla.

Maaliskuussa 2004 Suomessa alennettiin alkoholin verotusta. Alkoholijuomien valmisteveroja alennettiin keskimäärin 33 prosenttia; väkevien alkoholijuomien veroja alennettiin 44 prosenttia, välituotteiden veroja 40 prosenttia, viinien veroja 10 prosenttia ja oluiden veroja 32 prosenttia. Tällä ennakoitiin jo edellä mainittua Viron liittymistä EU:hun ja matkustajatuonnin vapautumista sieltä. Viron merkitys alkoholin tuontilähteenä perustui Viron selvästi Suomea alhaisempiin alkoholihintoihin ja Viron maantieteelliseen läheisyyteen.

Huhtikuun 2007 alusta voimaan tuli lakimuutos, joka kielsi alkoholin myynnin vähittäismyyntipaikoissa kello 21.00-9.00 välisenä aikana. Ennen myynti oli kiellettyä kello 22.00-7.00 välisenä aikana.

Myöhemmin nähdään, voidaanko alkoholikuolleisuuden ja vallitsevien alkoholio-  
lojen välillä nähdä jonkinlaista yhteyttä.

### 3.1.2 Suomalaisen alkoholiasenteista

Terveyden edistämisen keskus on tehnyt selvityksen suomalaisten alkoholiin liittyvistä asenteista. Selvitykseen kuuluvat haastattelut tehtiin toukokuussa 2008 ja vastaajien lukumäärä oli 1001. Tutkimuksen otanta toteutettiin monivaiheisena ositettuna otantana; kiintiöinnit tehtiin vastaajan iän, sukupuolen ja asuinpaikan mukaan. Seuraavassa esiteltävät selvityksen tulokset löytyvät Terveyden edistämisen keskuksen julkaisusta *Suomalaisten alkoholiasenteet* (2008) [9].

Suomalaisten alkoholiasenteet -selvityksessä 53 % vastaajista oli sitä mieltä, että

alkoholiveron nosto ei ole tarpeellista alkoholihaittojen vähentämiseksi. 45 % vastaajista puolestaan piti alkoholiveron nostoa tarpeellisena. 2 % vastaajista ei osannut sanoa mielipidettään.

Alkoholimainonnan rajoittaminen nykyisestä saa kannatusta suomalaisten keskuudessa. 59 % vastaajista oli sitä mieltä, että alkoholin mielikuvamainonta pitäisi kieltää ja sallia ainoastaan hinta- ja tuotetietojen esittely. Alkoholimainonnan rajoittamista ei nähnyt tarpeellisena 38 % vastaajista.

Suomalaisten alkoholiasenteet -selvityksessä kysymys siitä, pitäisikö vähittäismyyntiliikkeissä myytävien alkoholituotteiden prosenttimäärää laskea, sai vastaajien keskuudessa 26 prosentin kannatuksen. Vastustajia oli puolestaan 71 prosenttia.

Myöhemmin jälleen nähdään, ovatko kansalaisten mielipiteet yhteydessä siihen, mitkä keinot ovat historiassa vaikuttaneet alkoholikuolleisuuteen.

## 3.2 Aineistosta

Tutkielman aineistona on alkoholikuolleisuus Suomessa vuosina 1969–2007. Koko aineiston tarkastelemisen lisäksi aineistoa käytetään sukupuolen mukaan ja eri ikäryhmiin jaoteltuna.

### 3.2.1 Alkoholikuolleisuuden määrittely

Tässä tutkielmassa alkoholikuolleisuus määritellään alkoholiin kuolleiden määränä sataatuhatta henkeä kohti seuraavasti:

$$\text{alkoholikuolleisuus} = \frac{\text{alkoholiin kuolleiden määrä}}{\text{väestömäärä}} \times 100000.$$

Seuraavaksi esitellään, kuinka alkoholiin kuolleiden määrät ja väestömäärät saadaan.

#### Alkoholiin kuolleiden määrät

Alkoholiin kuolleiden määrät on saatu Tilastokeskuksen Kuolemansyytilastosta. Tämän kappaleen lähteenä on käytetty Tilastokeskuksen julkaisua *Kuolemansyyt 2006* [8] ja Kuolemansyytilaston kotisivuja [12].

Tässä tutkielmassa alkoholiin kuolleella henkilöllä tarkoitetaan henkilöä, jonka peruskuolemansyy on alkoholitauti tai tapaturmainen alkoholimyrkytys. Peruskuolemansyyllä puolestaan tarkoitetaan sitä tautia tai tapahtumaa, joka on pannut alulle välittömästi kuolemaan johtaneen sairaustilojen sarjan. Mukana ovat kaikki ne henkilöt, joilla kuolinhetkellä oli kotipaikka Suomessa, riippumatta henkilön kansalaisuudesta.

Kuolemansyytilaston laatiminen perustuu prosessiin, joka alkaa, kun henkilön kuoltua lääkäri kirjoittaa kuolintodistuksen. Tarvittaessa ennen kuolintodistuksen laatimista suoritetaan ruumiinavaus kuolemansyyntä selvittämiseksi. Kuolintodistus toimitetaan lopulta Tilastokeskukselle.

Tilastokeskuksessa kuolemansyyt luokitellaan ja luokittelussa tukeudutaan ensisijaisesti kuolintodistuksessa tekstinä annettuun diagnoosiin. Jos kuolintodistuksen tiedot ovat puutteelliset, ristiriitaiset tai vaikeasti luokiteltavat, käytetään apuna kuolintodistuksen tapahtumatietoja, lääketieteellistä asiantuntijaa tai kysytään lisätietoja kuolintodistuksen kirjoittajalta. Alkoholimyrkytyksissä käytetään lisätietona Helsingin Yliopiston oikeuslääketieteen laitoksella olevan oikeuskemian rekisterin tutkimustuloksia.

Tilastokeskuksessa kuolemansyyden luokittelu perustuu kansainvälisiin tai kansallisiin tautiluokituksiin, joista on muodostettu Tilastokeskuksessa lyhennettyjä luokituksia. Tässä tutkielmassa käytetään näistä lyhennytyistä luokituksista 54-luokkaista peruskuolemansyylluokitusta (jatkossa tähän luokitukseen viitataan lyhyesti nimellä peruskuolemansyylluokitus), ja sen luokasta alkoholitaudit ja tapaturmaiset alkoholimyrkytykset saatavia kuolleiden määriä. Määrät saadaan iän ja sukupuolen mukaan jaoteltuna. Tämä peruskuolemansyylluokitus on Tilastokeskuksessa laadittu siten, että se on ajallisesti vertailukelpoinen, vaikka käytetyt tautiluokitukset ovat vuosien välillä muuttuneet. Taulukossa 3.1 esitellään peruskuolemansyylluokituksen alkoholitaudit ja tapaturmaiset alkoholimyrkytykset -luokkaa vastanneet tautiluokitusten koodit eri vuosina. Vuodesta 1996 lähtien peruskuolemansyylluokitus on perustunut kansainväliseen ICD-10-tautiluokitukseen. ICD-10-luokituksen koodit ja niitä vastaavat alkoholitaudit ja tapaturmaiset alkoholimyrkytykset selväkielisinä niminä on esitelty taulukossa 3.2.

Taulukko 3.1: Peruskuolemansyylluokituksen alkoholitaudit ja tapaturmaiset alkoholimyrkytykset -luokkaa vastaavat tautiluokitusten koodit. Lähde: *Kuolemansyyt 2006*, Tilastokeskus [8].

Vuosi	Tautiluokitus	Peruskuolemansyylluokituksen alkoholitaudit ja tapaturmaiset alkoholimyrkytykset -luokkaa vastaavat tautiluokituksen koodit
1969–1986	Kansainvälinen tautiluokitus ICD-8	291, 303, 5710, 577, E860(miehet)
1987–1995	Kansallinen tautiluokitus 1987 (noudatti kansainvälistä tautiluokitusta ICD-9)	291, 303, 3050, 3575, 4255, 5353, 5710-5713, 5770D-5770F, 5771C-5771D, 7607A, 7795A, E851
1996–	Kansainvälinen tautiluokitus ICD-10	F10, G312, G4051, G621, G721, I426, K292, K70, K860, K8600, 0354, P043, X45

Taulukko 3.2: ICD-10-tautiluokituksen koodit ja niitä vastaavat alkoholitaudit ja tapaturmaiset alkoholimyrkytykset selväkielisinä niminä. Lähde: Stakesissa toimivan Sosiaali- ja terveydenhuollon luokituskeskuksen ylläpitämä ICD-10-luokituksen suomalainen versio [7].

ICD-10 -koodi	Selitys
F10	Alkoholin käytön aiheuttamat elimelliset aivo-oireyhtymät ja käyttäytymisen häiriöt.
G312	Alkoholin aiheuttama hermoston rappeutuminen; alkoholin aiheuttama (iso)aivorappeuma; alkoholin aiheuttama autonomisen hermoston toimintahäiriö; alkoholin aiheuttama aivosairaus; alkoholin aiheuttama pikkuaivorappeuma; alkoholin aiheuttama pikkuaivoataksia.
G4051	Alkoholin käyttöön liittyvät epileptiset kohtaukset.
G621	Alkoholin käyttöön liittyvä monihermosairaus.
I426	Alkoholisydänlihassairaus.
K292	Alkoholin aiheuttama mahatulehdus.
K70	Alkoholin aiheuttama maksasairaus.
K860 ja K8600	Alkoholin aiheuttama haimatulehdus, akuutti vaihe; alkoholin aiheuttaman (toistuvan) haimatulehduksen jälkitila; alkoholin aiheuttama pitkäaikainen haimatulehdus.
O354	Äidin hoito raskauden aikana hänen alkoholinväärinkäyttönsä aiheuttaman (epäillyn) sikiövaurion vuoksi; sikiön epämuodostuma, jonka epäillään olevan äidin alkoholinkäytön aiheuttama.
P043	Äidin alkoholinkäytön vaikutus sikiöön ja vastasyntyneeseen.
X45	Myrkytystapaturma tai muu altistuminen alkoholeille.

## Väestömäärät

Väestömäärät on saatu Tilastokeskuksen Väestörakennetilastosta. Tämän kappaleen lähteenä on käytetty Tilastokeskuksen Väestörakennetilaston kotisivuja [13].

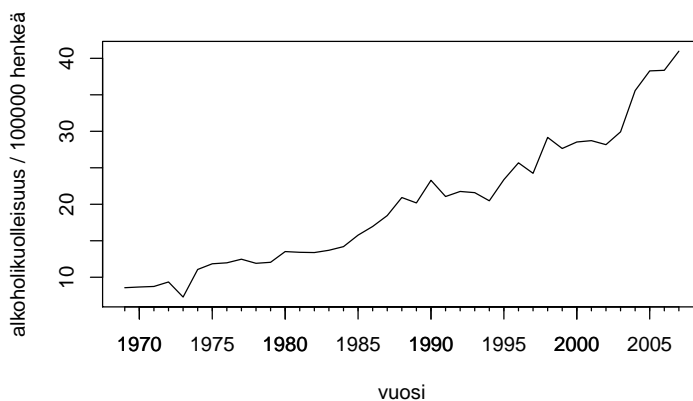
Tilastokeskuksen Väestörakennetilasto kuvaa Suomessa vuoden viimeisenä päivänä asunutta väestöä. Väestöön kuuluvat ne Suomen kansalaiset ja ulkomaalaiset, jotka asuvat vakituisesti Suomessa, vaikka olisivatkin tilapäisesti ulkomailla.

Väestörekisterikeskus ja maistraatit ylläpitävät Suomen väestötietojärjestelmää. Väestörekisterikeskus toimittaa Tilastokeskukselle väestön tiedot ja Tilastokeskus koottaa Väestörakennetilaston.

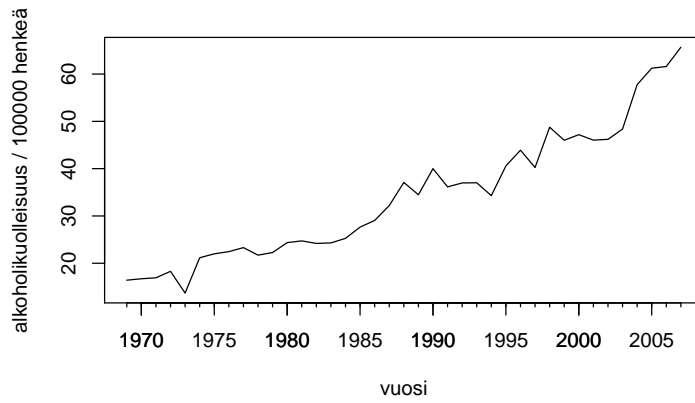
Tiedossa on täten väestömäärät iän ja sukupuolen mukaan jaoteltuna jokaisen vuoden viimeiselle päivälle. Tästä lasketaan vuosittaiset keski­väestömäärät tietylle vuodelle laskemalla aritmeettinen keskiarvo edellisen vuoden ja kyseessä olevan vuoden väestömääristä. Jatkossa termillä väestömäärä viitataan nimenomaan keski­väestömäärään.

### 3.2.2 Aineiston kuvailu

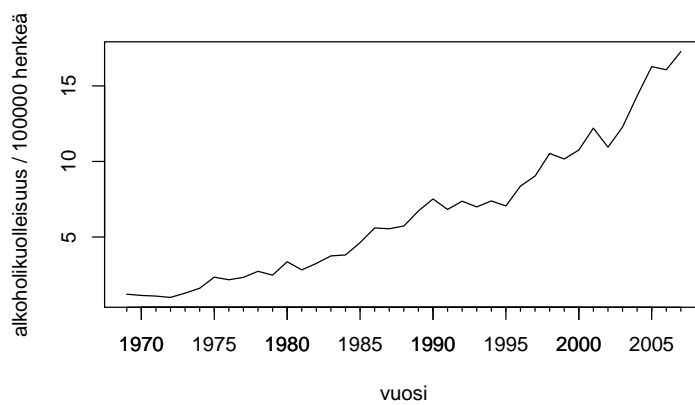
Alkoholikuolleisuuden kehitystä vuosina 1969–2007 koko väestön osalta havainnollistetaan kuvassa 3.1. Kuolleisuudessa on nouseva trendi. Kuvissa 3.2 ja 3.3 havainnollistetaan erikseen miesten ja naisten alkoholikuolleisuuden kehitystä. Naisten alkoholikuolleisuus on koko tarkastelujakson ajan ollut pienempi kuin miesten. Molempien sukupuolten tapauksessa kuolleisuudessa näkyy kuitenkin nouseva kehityssuunta.



Kuva 3.1: Alkoholikuolleisuus vuosina 1969–2007, koko väestö.



Kuva 3.2: Alkoholikuolleisuus vuosina 1969–2007, miehet.



Kuva 3.3: Alkoholikuolleisuus vuosina 1969–2007, naiset.

### 3.3 Mallin määrittely

Alkoholiperäisistä syistä kuolleiden määrä on prosessi, jonka muutoksista ajassa halutaan tietoa. Olkoon  $y_t$  hetkellä  $t$  tästä prosessista tehty havainto eli havaittu alkoholiin kuolleiden määrä. Kuolleiden määrä on suhteutettava kunkin hetken  $t$  väestömäärään  $u_t$ , jotta alkoholikuolleisuuden ajallinen vertailu onnistuu. Havaittu alkoholiin kuolleiden määrä  $y_t$  mallinnetaan Poisson-jakaumalla, jonka odotusarvo on  $\mu_t = u_t \lambda_t$ , missä muuttujaa  $\lambda_t$  kutsutaan intensiteetiksi. Muuttujan  $y_t$  pistetodennäköisyysfunktio on täten

$$p(y_t|\lambda_t) = \frac{(u_t \lambda_t)^{y_t}}{y_t!} e^{-u_t \lambda_t}$$

kaikilla  $t = 1, \dots, n$ .

Käytettäväksi malliksi on valittu Poisson-jakauma, koska sitä käytetään yleisesti tilanteissa, joissa populaatio on suuri, mutta mielenkiinnon kohteena oleva tapahtuma on harvinainen. Alkoholikuolleisuuden tilanteessa populaationa on Suomen väestö, eivätkä alkoholiperäiset kuolemat ole kovin yleisiä väestön kokoon nähden.

Parametroidaan nyt Poisson-jakauma toisin. Poisson-jakauman pistetodennäköisyysfunktio voidaan muuttaa muotoon

$$p(y_t|\lambda_t) = \exp[\log p(y_t|\lambda_t)] = \exp[y_t \log \lambda_t - u_t \lambda_t + (y_t \log u_t - \log(y_t!))].$$

Kun valitaan  $\theta_t = \log \lambda_t$ , jakauma saadaan muotoon

$$p(y_t|\theta_t) = \exp[y_t \theta_t - u_t \exp \theta_t + (y_t \log u_t - \log(y_t!))],$$

kaikilla  $t = 1, \dots, n$ . Merkitään seuraavaksi  $\theta_t = \nu_t$  ja määritellään yhteys:

$$\begin{aligned} \nu_{t+1} &= \nu_t + \phi_t + \xi_t, & \xi_t &\sim N(0, \sigma_\xi^2), \\ \phi_{t+1} &= \phi_t + \zeta_t, & \zeta_t &\sim N(0, \sigma_\zeta^2). \end{aligned}$$

Edellä olevaa mallia kutsutaan Poisson jakaumaan perustuvaksi lokaalisen lineaarisen trendin malliksi ja se voidaan esittää matriisimerkinnöin seuraavasti:

$$\begin{aligned} p(y_t|\theta_t) &= \exp[y_t \theta_t - u_t \exp \theta_t + (y_t \log u_t - \log(y_t!))], \\ \theta_t &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \nu_t \\ \phi_t \end{bmatrix}, \\ \begin{bmatrix} \nu_{t+1} \\ \phi_{t+1} \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \nu_t \\ \phi_t \end{bmatrix} + \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix}, \quad \begin{bmatrix} \xi_t \\ \zeta_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\xi^2 & 0 \\ 0 & \sigma_\zeta^2 \end{bmatrix}\right). \end{aligned} \tag{3.1}$$

Mallissa sekä trendin tason että kulmakertoimen sallitaan vaihtelevan ajassa. Malli on siis hyvin väljä ja sopii monenlaisiin aineistoihin. Tästä johtuen tämä malli on sovitettu myös kyseessä olevaan aineistoon.

Malli kuuluu määritelmän 2.6 mukaiseen eksponenttiseen perheeseen, jossa

$$b_t(\theta_t) = u_t \exp(\theta_t)$$



ja

$$c_t(y_t) = y_t \log u_t - \log(y_t!).$$

Mallin linearisaatio tehdään kappaleen 2.2.2 mukaisesti. Nyt saadaan  $\dot{b}_t^{(i)} = u_t \exp(\check{\theta}_t^{(i)})$  ja  $\ddot{b}_t^{(i)} = u_t \exp(\check{\theta}_t^{(i)})$ , mistä edelleen

$$\check{H}_t^{(i)} = \frac{\exp(-\check{\theta}_t^{(i)})}{u_t}$$

ja

$$\check{y}_t^{(i)} = \check{\theta}_t^{(i)} - 1 + \frac{y_t \exp(-\check{\theta}_t^{(i)})}{u_t}.$$

Logaritmisen nollassuuttavuusfunktion muodostamisessa tarvittava  $l_t''''(\theta)$  on puolestaan

$$l_t''''(\theta) = -u_t \exp(\theta_t).$$

Päämääränä on selvittää kuolleiden määrän muutoksia ajassa ehdolla kaikki tehdyt havainnot  $y$ . Lasketaan täten odotusarvo  $\hat{\mu}_t := E_p(\mu_t|y)$  ja keskivirhe  $se_p(\mu_t|y) = \sqrt{\text{var}_p(\mu_t|y)}$  sekä etsitään 95 prosentin luottamusväli  $(\mu_{t,0.025}, \mu_{t,0.975})$  muuttujalle  $\mu_t$  ehdolla kaikki havainnot  $y$ . Odotusarvon ja varianssin laskusääntöjä sekä deltamenettelmää avuksi käyttäen saadaan seuraavat kaavat:

$$\begin{aligned} \hat{\mu}_t &= E_p(u_t \lambda_t | y) \\ &= u_t E_p(\lambda_t | y) \\ &= u_t E_p(\exp(\theta_t) | y) \\ &\approx u_t \exp(E_p(\theta_t | y)) \\ &= u_t \exp(E_p(Z_t \alpha_t | y)) \\ &= u_t \exp(Z_t E_p(\alpha_t | y)) \\ &= u_t \exp(Z_t \hat{\alpha}_t) \end{aligned}$$

ja

$$\begin{aligned} \text{var}_p(\mu_t | y) &= \text{var}_p(u_t \lambda_t | y) \\ &= u_t^2 \text{var}_p(\lambda_t | y) \\ &= u_t^2 \text{var}_p(\exp(\theta_t) | y) \\ &\approx u_t^2 [\exp(E_p(\theta_t | y))]^2 \text{var}_p(\theta_t | y) \\ &= u_t^2 [\exp(Z_t \hat{\alpha}_t)]^2 \text{var}_p(Z_t \alpha_t | y) \\ &= u_t^2 [\exp(Z_t \hat{\alpha}_t)]^2 Z_t \text{var}_p(\alpha_t | y) Z_t'. \end{aligned}$$

Lisäksi 95 prosentin luottamusvälille ehdolla kaikki havainnot  $y$  pätee

$$\begin{aligned} (\mu_{t,0.025}, \mu_{t,0.975}) &= (u_t \lambda_{t,0.025}, u_t \lambda_{t,0.975}) \\ &= u_t (\lambda_{t,0.025}, \lambda_{t,0.975}) \\ &= u_t (\exp(\theta_{t,0.025}), \exp(\theta_{t,0.975})). \end{aligned}$$

Lopulliseksi tehtäväksi muotoutuu siis arvojen  $\hat{\alpha}_t = E_p(\alpha_t|y)$  ja  $\text{var}_p(\alpha_t|y)$  laskeminen kappaleen 2.2.3 mukaisesti sekä luottamusvälin  $(\theta_{t,0.025}, \theta_{t,0.975})$  muodostaminen kappaleen 2.2.4 mukaisesti.

### 3.4 Mallin sovitus

Alkoholikuolleisuusaineistoon sovitetaan Poisson-jakaumaan perustuva lokaalin lineaarisen trendin malli (3.1). Alustus suoritetaan valitsemalla

$$\begin{bmatrix} \nu_1 \\ \phi_1 \end{bmatrix} \sim N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 100000 & 0 \\ 0 & 100000 \end{bmatrix} \right).$$

Mallin sovittamisessa käytän kirjoittamiani R-koodeja, jotka ovat liitteessä A. Linearisaatiossa käytetään 20 iteraatiota ja tasoituksen teossa tiloille käytetään 1000 simulointia.

Valittu malli sovitetaan koko aineistoon sekä naisten ja miesten osa-aineistoon. Lisäksi sovitus on tehty kymmenvuotiskäryhmittäin 20–79 -vuotiaiden joukossa naisten ja miesten osa-aineistoissa erikseen.

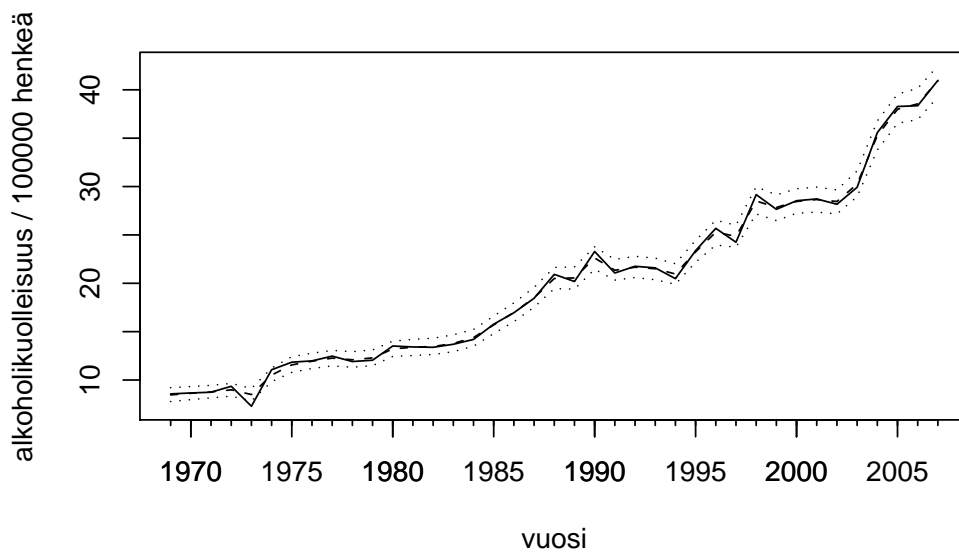
Taulukossa 3.3 on mallin tuntemattomien parametrien  $\sigma_\xi^2$  ja  $\sigma_\zeta^2$  estimaatit eri ryhmille.

Taulukko 3.3: Parametrien  $\sigma_\xi^2$  ja  $\sigma_\zeta^2$  estimaatit.

Ryhmä	$10000 \cdot \hat{\sigma}_\xi^2$	$10000 \cdot \hat{\sigma}_\zeta^2$
kaikki	47.17	0.00
miehet	53.49	0.00
miehet 20–29	245.00	0.01
miehet 30–39	104.99	0.01
miehet 40–49	58.77	0.00
miehet 50–59	54.07	0.00
miehet 60–69	16.60	0.00
miehet 70–79	0.02	0.02
naiset	22.39	0.38
naiset 20–29	351.74	0.28
naiset 30–39	0.15	2.52
naiset 40–49	0.06	0.83
naiset 50–59	75.94	0.42
naiset 60–69	13.51	0.00
naiset 70–79	72.91	0.02

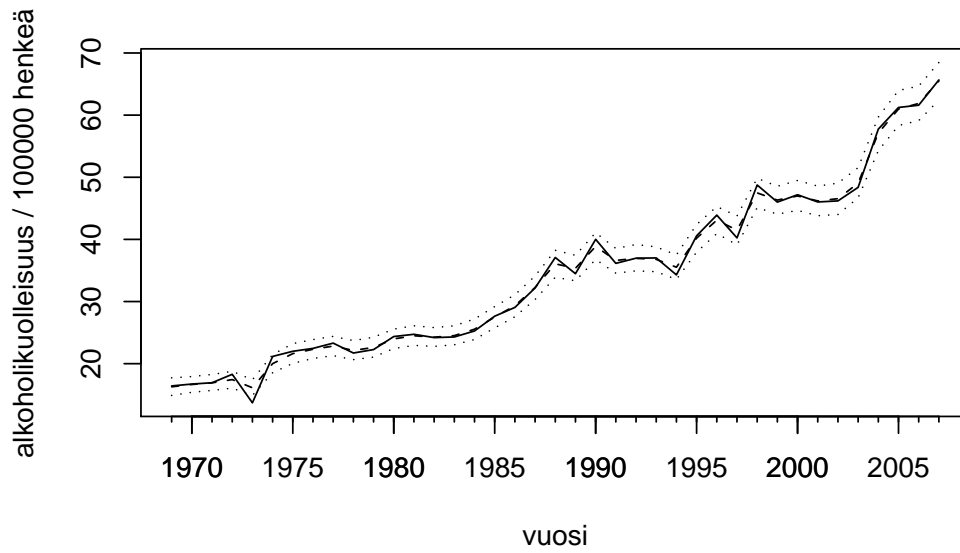
Koko väestön osalta tasoitettu odotusarvo ja 95 prosentin luottamusväli yhdessä alkuperäisen sarjan kanssa on piirretty kuvaan 3.4. Tässä kuvassa ja jatkossa kaikissa muissakin kuvissa sarjojen arvot on luvun 3.2.1 alkoholikuolleisuuden määritelmän

mukaisesti suhteutettu lukuun 100000. Lisäksi kuvat on piirretty niin, että kuvasta toiseen asteikko pystyakselilla vaihtelee. Samalla tavalla on saatu miesten alkoholi-kuolleisuutta kuvaava kuva 3.5 ja naisten alkoholi-kuolleisuutta esittelevä kuva 3.6. Kaikissa näissä tapauksissa tasoitus lieventää muutamaa jyrkkää kuolleisuusvaihtelua, mutta selvästi nähdään, että alkoholi-kuolleisuus on lisääntynyt vuodesta 1969 vuoteen 2007. Vuonna 2004 sekä koko väestöä että miesten osaryhmää tarkasteltaessa nähdään, että kuolleisuus on kasvanut voimakkaasti tämän yhden vuoden aikana. Naisilla vastaavaa yhtä voimakasta kasvua ei näy.

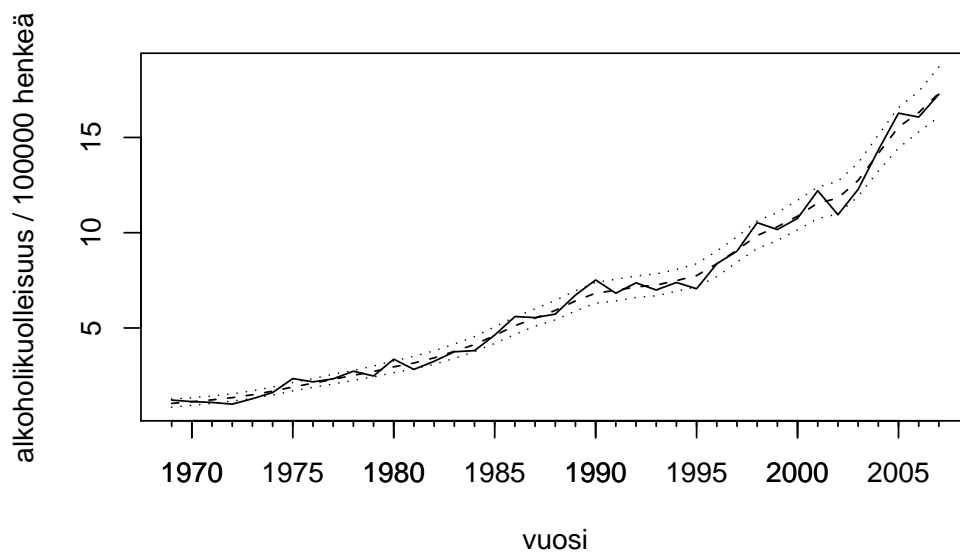


Kuva 3.4: Koko väestön alkoholi-kuolleisuus, tasoitettu odotusarvo ja luottamusväli.

Seuraavaksi tarkastellaan tuloksia kymmenvuotiskäryhmiin jaoteltuna eri sukupuolille erikseen. Tarkastellaan aluksi miehiä (kuva 3.7). 20–29-vuotiaiden ryhmässä kuolleisuus on vuodesta 1969 vuoteen 1975 kasvanut voimakkaasti, jonka jälkeen kuolleisuus on hitaammalla tahdilla vähentynyt aina vuoteen 1992 asti. Sen jälkeen kuolleisuus on jälleen kasvanut, muttei läheskään 1970-luvun puolivälin tasolle. 30–39-vuotiaiden ryhmässä kuolleisuus on pääpiirteissään kasvanut vuodesta 1969 vuoteen 1990 asti. Tämän jälkeen kuolleisuudessa on ollut jyrkkä lasku vuoteen 1994 asti, jonka jälkeen kuolleisuus on vuosikymmenen lopuksi taas noussut korkeammalle tasolle. 2000-luvun alussa kuolleisuus on hetkeksi jälleen vähentynyt, mutta vuodesta 2004 lähtien kuolleisuus on ollut kasvusuunnassa. Ikäryhmässä 40–49-vuotiaat kuolleisuus on ollut hyvin tasaista vuodesta 1969 vuoteen 1984 asti, jonka jälkeen kuolleisuus on noussut voimakkaasti vuoteen 1988 asti. Tasaisemman ajanjakson jälkeen kuolleisuus on vuodesta 1995 alkaen jälleen kasvanut vuoteen 1998 asti. Sen jälkeen kuolleisuus on vuoden tai kahden välein vuoroin kasvanut ja vähentynyt, muttei kovin voimakkaasti. 50–59-vuotiaiden miesten joukossa alkoholi-kuolleisuus on kasvanut lähes lineaarisesti vuodesta 1969 aina vuoteen 2003 asti. Vuonna 2004 kuolleisuus on lähtenyt



Kuva 3.5: Miesten alkoholikuolleisuus, tasoitettu odotusarvo ja luottamusväli.



Kuva 3.6: Naisten alkoholikuolleisuus, tasoitettu odotusarvo ja luottamusväli.

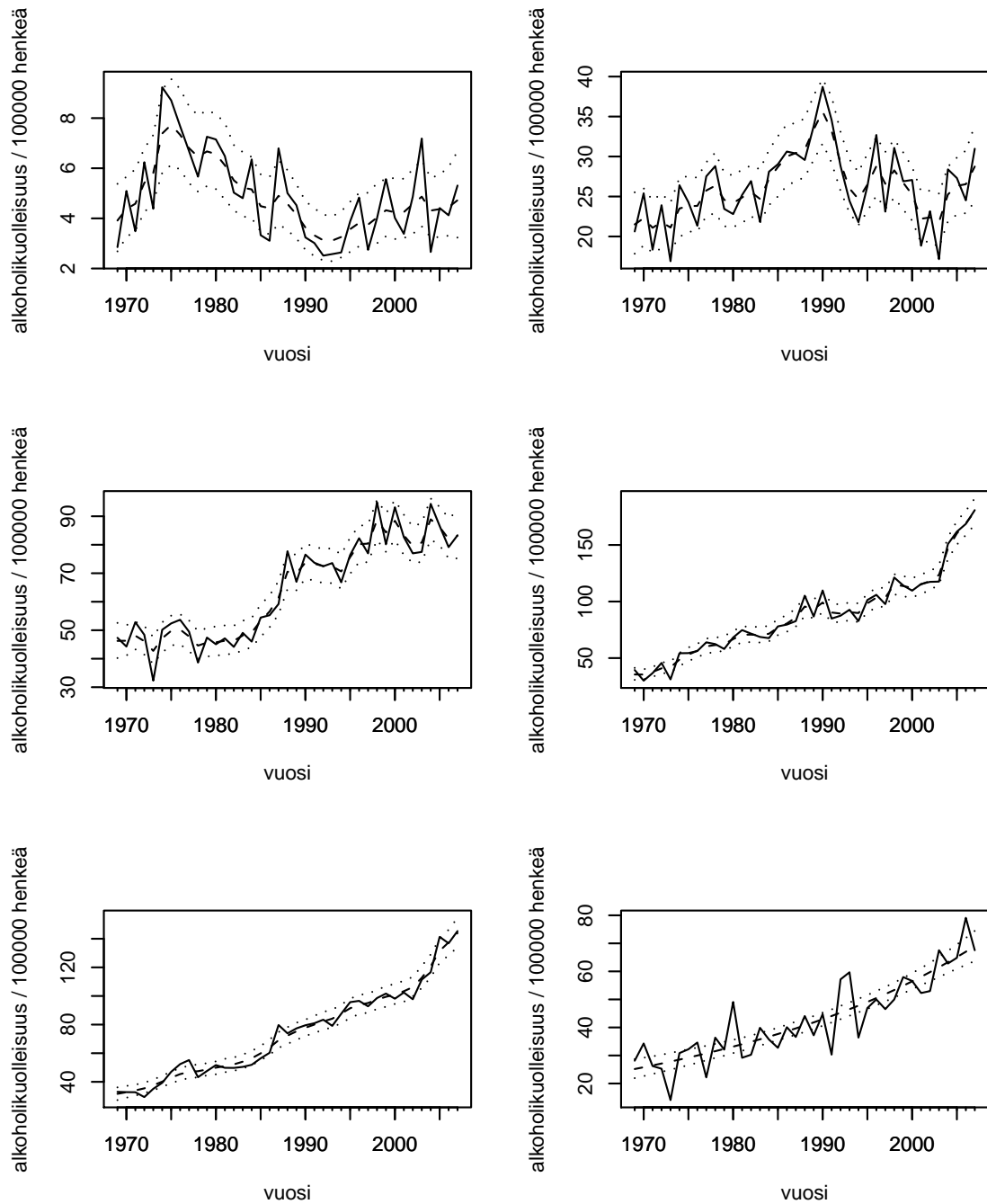
aikaisempaan verrattuna todella jyrkkään kasvuun ja kasvu on jatkunut lähes yhtä voimakkaana aina vuoteen 2007 asti. 60–69-vuotiaiden ikäryhmän alkoholikuolleisuuskehitys on ollut suoraviivaisesti kasvavaa vuodesta 1969 vuoteen 2003. Vuodesta 2004 lähtien kasvu on jyrkentynt. 70–79-vuotiaiden miesten ryhmää analysoitaessa nähdään, että kuolleisuuden kasvu on ollut hyvin tasaista koko tarkastelujakson eli vuosien 1969–2007 aikana.

Siirrytään tarkastelemaan naisten ikäryhmittäistä alkoholikuolleisuutta (kuva 3.8). Ikäryhmässä 20–29-vuotiaat alkoholikuolleisuus ei ole tarkastelujaksolla vuosina 1969–2007 juurikaan vaihdellut. 30–39-vuotiaiden ikäryhmässä kuolleisuus on kasvanut tasaisesti vuodesta 1969 vuoteen 1995, jonka jälkeen se on loivasti lähtenyt laskuun. 40–49-vuotiaiden ryhmässä kuolleisuus on kasvanut tasaisesti koko tarkastelujakson ajan. 50–59-vuotiaiden ikäryhmässä kuolleisuuden kasvu on myös ollut tasaista vuoteen 2003 asti, jonka jälkeen kahden vuoden ajan kasvu on ollut kiihtyvää. Vuosina 2006 ja 2007 kasvua ei enää juurikaan ole tapahtunut. 60–69-vuotiaiden naisten ikäryhmässä koko tarkastelujakson ajan alkoholikuolleisuuden kasvu on ollut lievästi kiihtyvää. 70–79-vuotiaiden naisten ikäryhmässä vuodesta 1969 vuoteen 1990 kuolleisuus on kasvanut tasaisesti, vuosina 1991–1995 kuolleisuus on pysynyt vakiona ja jälleen vuodesta 1996 lähtien vuoteen 2007 asti kuolleisuus on kasvanut lähes lineaarisesti, mutta nopeampaa tahtia kuin ennen vuotta 1991.

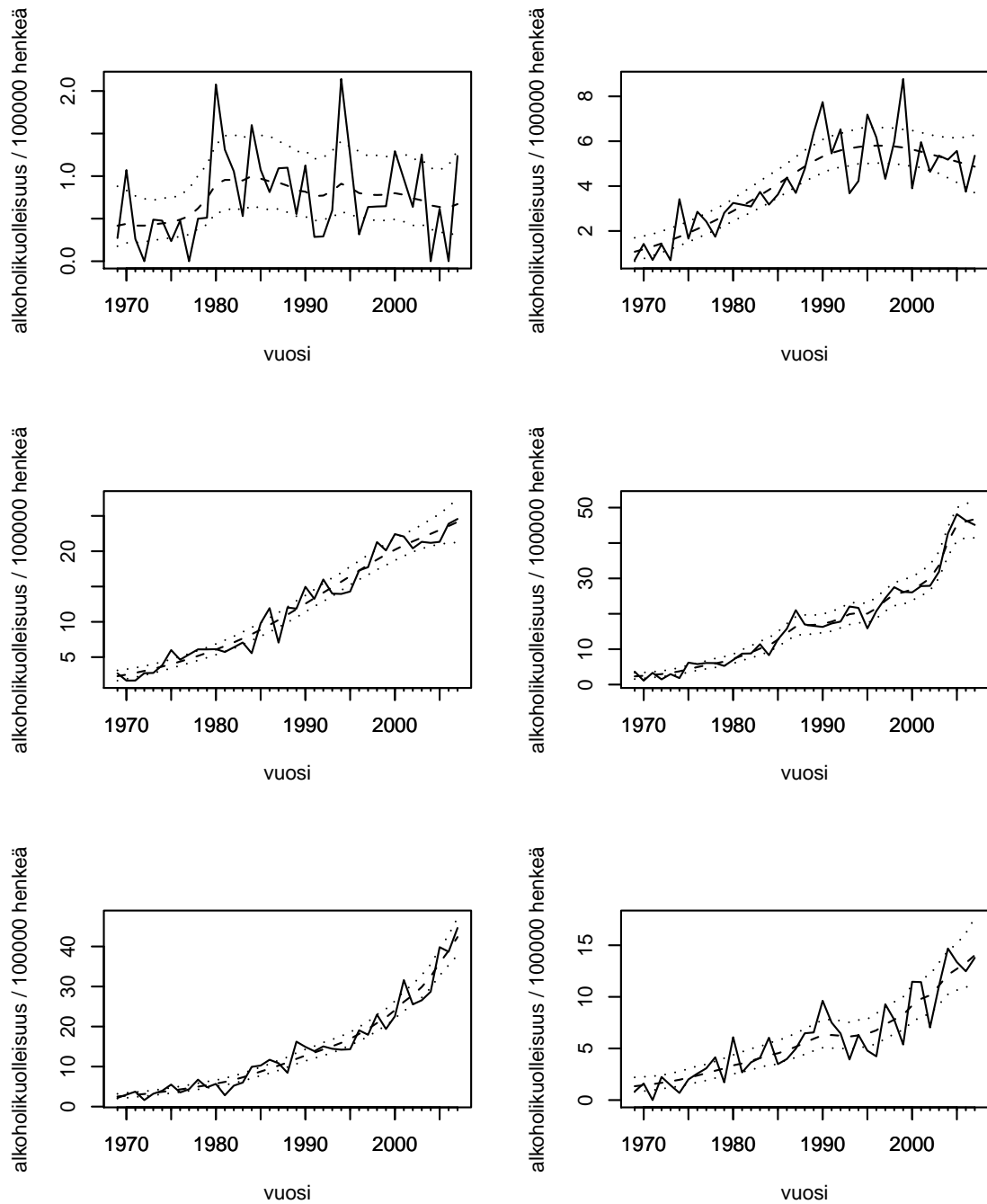
Tarkastellaan vielä lähemmin tasoitettuja odotusarvoja, variansseja ja luottamusvälejä viimeiselle käytössä olevalle vuodelle 2007. Tulokset esitellään taulukossa 3.4. On huomattava, että tässä taulukossa kuolleiden määriä ei ole suhteutettu lukuun 100000.

Taulukko 3.4: Kuolleiden määrät sekä tasoitetut odotusarvot, keskivirheet ja 95 prosentin luottamusvälit vuodelle 2007.

Ryhmä	kuolleiden määrä	odotusarvo	keskivirhe	95%:n luottamusväli
kaikki	2167	2163.42	43.85	(2076.71, 2246.27)
miehet	1701	1696.35	39.80	(1616.70, 1773.77)
miehet 20–29	18	16.01	2.90	(10.96, 22.63)
miehet 30–39	103	95.59	8.05	(81.04, 112.64)
miehet 40–49	316	315.44	15.71	(285.23, 347.33)
miehet 50–59	719	714.52	23.70	(668.43, 760.28)
miehet 60–69	411	407.16	15.06	(379.14, 437.17)
miehet 70–79	112	114.35	4.51	(105.63, 123.46)
naiset	466	467.46	18.12	(434.01, 504.82)
naiset 20–29	4	2.17	0.79	(1.01, 4.21)
naiset 30–39	17	15.45	2.07	(11.74, 19.93)
naiset 40–49	91	89.42	5.63	(78.70, 101.79)
naiset 50–59	180	186.47	11.08	(165.39, 209.88)
naiset 60–69	136	129.10	7.13	(115.42, 143.76)
naiset 70–79	31	31.57	3.57	(25.66, 39.72)



Kuva 3.7: Miesten alkoholikuolleisuus, tasoitettu odotusarvo ja luottamusväli ikäryhmissä 20–29, 30–39, 40–49, 50–59, 60–69 ja 70–79.



Kuva 3.8: Naisten alkoholikuolleisuus, tasoitettu odotusarvo ja luottamusväli ikäryhmissä 20–29, 30–39, 40–49, 50–59, 60–69 ja 70–79.

### 3.5 Johtopäätökset

Yleisesti voidaan todeta, että vuodesta 1969 vuoteen 2007 suomalaisten alkoholikuolleisuus on kasvanut. 40–79-vuotiaiden joukossa kehitys on pääpiirteissään ollut suoraan kasvavaa sekä miesten että naisten ryhmissä, mutta nuorempien 20–29- ja 30–39-vuotiaiden ikäryhmissä kuolleisuuden kehitys on ollut vaihtelevaa.

Tarkastellaan nyt, voidaanko havaita yhteyksiä kappaleessa 3.1.1 mainittujen alkoholiolojen muutosten ja alkoholikuolleisuuden välillä. Ainoa selkeä yhteys, joka voidaan havaita, on alkoholikuolleisuuden jyrkkä kasvu vuosina 2004 ja 2005. Tämä jyrkkä kasvu on selkeimmin näkyvissä 50–59- ja 60–69-vuotiaiden miesten sekä 50–59-vuotiaiden naisten ryhmissä. Vaikutus on niin selvä, että se näkyy koko väestönkin alkoholikuolleisuudessa. Vuonna 2004 suoritettiin sekä alkoholin veronalennus että matkustajatuonti Virossa vapautui, joten näillä tapahtumilla on siis ilmeisesti ollut alkoholikuolleisuutta lisäävä vaikutus.

Seuraavaksi tarkastellaan luvussa 3.1.2 mainittujen suomalaisten alkoholisientien yhteyksiä siihen, mitkä keinot ovat historiassa vaikuttaneet alkoholikuolleisuuteen. Niukka enemmistö suomalaisista (53 %) oli sitä mieltä, että alkoholiveron nosto ei ole tarpeellista alkoholihaittojen vähentämiseksi. Ilmeisesti kuitenkin kuolleisuus voisi vähentyä alkoholiveroa nostamalla, ainakin jos vuoden 2004 alkoholikuolleisuuden jyrkän kasvun voitaisiin tietää aiheutuneen enemmän veronalennuksesta kuin matkustajatuonnin vapautumisesta. Enemmistö suomalaisista on sitä mieltä, että alkoholin mielikuvamainonta pitäisi kieltää ja sallia ainoastaan hinta- ja tuotetietojen esittely. Vuonna 1977 alkoholimainonta kiellettiin lailla ja vuonna 1995 mietojen alkoholijuomien mainonta sallittiin jälleen tietyin rajoituksin. Vuoden 1977 kiellolla ei näytä olleen alkoholikuolleisuuteen vaikutusta. Vuodesta 1995 alkaen on 40–49-vuotiaiden miesten alkoholikuolleisuudessa jyrkähkö kasvu, mutta on mahdotonta sanoa, onko sillä yhteyttä mainonnan sallimiseen, koska vuonna 1995 tapahtui alkoholioloissa myös monia muita muutoksia.

Edellisten yhteyksien lisäksi alkoholikuolleisuudesta voidaan tehdä myös seuraava huomio. 20–29-vuotiaiden miesten ikäryhmässä kuolleisuus on ollut suurinta 1970-luvun puolivälistä 1980-luvun alkuvuosiin. 30–39-vuotiaiden ikäryhmässä kuolleisuus on puolestaan ollut suurinta vuodesta 1985 alkaen 1990-luvun alkupuolelle. Tästä voidaan päätellä, että vuosina 1945–1955 syntyneiden miesten sukupolvi eli niin sanottu suuret ikäluokat ovat kuolleet nuoruusvuosinaan alkoholiin enemmän kuin muut sukupolvet. Koska suurempi kuolleisuus ei käytännössä voi johtua muusta kuin suuremmasta alkoholin kulutuksesta, niin voidaan tehdä päätelmä, että suuret ikäluokat ovat myös käyttäneet nuoruusvuosinaan alkoholia enemmän kuin muut sukupolvet.

Taloudellisilla suhdanteillakin saattaa olla alkoholikuolleisuuteen vaikutusta. 40–49-vuotiaiden miesten ryhmässä alkoholikuolleisuuden jyrkin kasvu ajoittuu 1980-luvun lopun nousukauteen. Myös 30–39-vuotiaiden miesten alkoholikuolleisuudessa on suurin piikki nousukauden loppuaikoina. 1990-luvun laman aikaan puolestaan 20–29-vuotiaiden miesten alkoholikuolleisuus oli koko tarkastelujakson matalimmalla tasolla. 90-luvun laman aikaan myös 40–49- ja 50–59-vuotiaiden miesten ja 70–79-vuotiaiden naisten ryhmissä kuolleisuuden kasvu on hetkeksi pysähtynyt.



# Viitteet

- [1] Anderson, T. W. : *An Introduction to Multivariate Statistical Analysis*. Toinen painos. John Wiley & Sons. New York. 1984.
- [2] Durbin, J.; Koopman, S. J.: Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models. *Biometrika*, Vol. 84, No. 3, s. 669–684. 1997.
- [3] Durbin, J.; Koopman, S. J.: *A simple and efficient simulation smoother for state space time series analysis*. 2001.  
URL: <http://www.ssfpack.com/dkbook/dksimsmo.pdf>.
- [4] Durbin, J.; Koopman, S. J.: *Time Series Analysis by State Space Methods*. Oxford University Press. Oxford. 2001.
- [5] FINLEX® - Valtion säädöstietopankki.  
URL: <http://www.finlex.fi/fi/>.
- [6] Harvey, A. C.; Phillips, G. D. A.: The estimation of regression models with autoregressive-moving average disturbances. *Biometrika*, 66, s. 49–58. Oxford University Press. 1979.
- [7] ICD-10 -luokituksen suomalainen versio.  
URL: <http://sty.stakes.fi/FI/luokituskeskus/index.htm>.
- [8] *Kuolemansyyt 2006*. Tilastokeskus. 2007.
- [9] Laari, Liisa; Opari, Päivi; Varamäki, Ritva: *Suomalaisten alkoholiasenteet*. Terveyden edistämisen keskuksen julkaisuja 3/2008. 2008.
- [10] R Development Core Team: R, A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 2008.  
URL: <http://www.R-project.org>.
- [11] *Päihdetilastollinen vuosikirja 2007*. Stakes. 2007.
- [12] Tilastokeskuksen Kuolemansyytilaston kotisivu.  
URL: <http://www.tilastokeskus.fi/til/ksyyt/index.html>.
- [13] Tilastokeskuksen Väestörakennetilaston kotisivu.  
URL: <http://www.tilastokeskus.fi/til/vaerak/index.html>.

# Liite A

## R-ohjelmiston koodit

Seuraavassa on tutkielmassa käytetyt R-koodit. R-ohjelmistosta löytyy lisätietoa viitteestä [10].

```
#Eksponenttisen perheen tila-avaruusmallien analysointifunktioita

# Funktio: gYtCondThetat
gYtCondThetat <- function(y_t,theta_t,H_t) {
  p_y_t_cond_theta_t_H_t <- dnorm(y_t,mean=theta_t,sd=sqrt(H_t))
  return(p_y_t_cond_theta_t_H_t)
}

# Funktio: mleExpFamilyStart
mleExpFamilyStart <- function(a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start,initial,omax) {
  o <- optim(par=initial,
            fn=minuslogLExpFamilyStart,
            method="BFGS",
            y=y,a1=a1,P1=P1,Z=Z,T=T,R=R,Q=Q,itermax=itermax,alpha_tilde_start=alpha_tilde_start,
            control=list(maxit=omax))
  return(o)
}

# Funktio: minuslogLExpFamilyStart
minuslogLExpFamilyStart <- function(psi,a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start) {
  return((-1)*logLExpFamilyStart(psi,a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start))
}

# Funktio: logLExpFamilyStart
logLExpFamilyStart <- function(psi,a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start) {
  n <- length(y)

  if (is.null(Z)) {
    Z <- Zf(psi)
  }

  if (is.null(T)) {
    T <- Tf(psi)
  }

  if (is.null(R)) {
    R <- Rf(psi)
  }

  if (is.null(Q)) {
```

```

    Q <- Qf(psi)
  }

  line <- linearisation(a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start)

  varepsilon <- disturbancesmoother(a1,P1,line$y_tildes_final,
    Z,line$H_tildes_final,T,R,Q)$varepsilon

  theta_hat <- statesmoother(a1,P1,line$y_tildes_final,Z,line$H_tildes_final,T,R,Q)$theta_hat

  logLg <- logLg(a1,P1,line$y_tildes_final,Z,line$H_tildes_final,T,R,Q)

  w_hat <- w(theta_hat,y,line$y_tildes_final,line$H_tildes_final)

  sumterm <- 0
  for (t in 1:n) {
    sumterm <- sumterm + l4(t,theta_hat[[t]]*(varepsilon[[t]])^2)
  }

  return(as.vector(logLg + log(w_hat) + log(1+(1/8)*sumterm)))
}

# Funktio: w
w <- function(theta,y,y_tilde,H_tilde) {
  theta_v <- matrixlistToVector(theta)
  y_v <- matrixlistToVector(y)
  y_tilde_v <- matrixlistToVector(y_tilde)
  H_tilde_v <- matrixlistToVector(H_tilde)
  product <- 1
  for (t in 1:n) {
    product <- product*(pYtCondThetat(t,y_v[t],theta_v[t])/
      gYtCondThetat(y_tilde_v[t],theta_v[t],H_tilde_v[t]))
  }
  return(product)
}

# Funktio: logLg
logLg <- function(a1,P1,y,Z,H,T,R,Q) {
  n <- length(y)
  sf <- statefilter(a1,P1,y,Z,H,T,R,Q)
  v <- sf$v
  F <- sf$F
  sumterm <- 0
  for (t in 1:n) {
    sumterm <- sumterm + as.vector(log(det(F[[t]])) + t(v[[t]])%*%solve(F[[t]])%*%v[[t]])
  }
  return(-(n/2)*log(2*pi) - (1/2)*sumterm)
}

# Funktio: statesmootherExpFamily
statesmootherExpFamily <- function(a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start,nsim,simul=NULL) {
  n <- length(y)
  m <- length(T[[1]][1,])

  line <- linearisation(a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start)

  if (is.null(simul)) {
    simu <- simulate(a1,P1,line$y_tildes_final,Z,line$H_tildes_final,T,R,Q,nsim)
  }
  else {
    simu <- simul
  }
}

```

```

alpha_simuls_1 <- simu$alpha_simuls_1
theta_simuls_1 <- simu$theta_simuls_1
theta_simuls_v <- simu$theta_simuls_v

alpha_hat_1 <- makeMatrixlist(n,m,1)
alpha_var_1 <- makeMatrixlist(n,m,m)

division <- vector(length=nsim)
for (k in 1:nsim) {
  division[k] <- w(theta_simuls_1[[k]],y,line$y_tildes_final,line$H_tildes_final)
}

denominator <- sum(division)

for (t in 1:n) {

  numerator1 <- 0
  numerator2 <- 0

  for (k in 1:nsim) {
    numerator1 <- numerator1 + alpha_simuls_1[[k]][[t]]*division[k]
    numerator2 <- numerator2 + (alpha_simuls_1[[k]][[t]])**%
      t(alpha_simuls_1[[k]][[t]])*division[k]
  }

  alpha_hat_1[[t]] <- numerator1/denominator
  alpha_var_1[[t]] <- numerator2/denominator - alpha_hat_1[[t]]**%t(alpha_hat_1[[t]])
}

theta_hat_1<-makeMatrixlist(n,1,1)
for (t in 1:n) {
  theta_hat_1[[t]] <- Z[[t]]**%alpha_hat_1[[t]]
}
theta_hat_v <- matrixlistToVector(theta_hat_1)

theta_var_1<-makeMatrixlist(n,1,1)
for (t in 1:n) {
  theta_var_1[[t]] <- Z[[t]]**%alpha_var_1[[t]]**%t(Z[[t]])
}
theta_var_v <- matrixlistToVector(theta_var_1)

theta_simuls_v2 <- vector("list",length=n)
for (t in 1:n) {
  theta_simuls_v2[[t]] <- vector(length=nsim)
  for (k in 1:nsim) {
    theta_simuls_v2[[t]][k] <- theta_simuls_v[[k]][t]
  }
}

theta_coin_down_v <- vector(length=n)
theta_coin_up_v <- vector(length=n)

for (t in 1:n) {
  rank_theta <- rank(theta_simuls_v2[[t]])
  division2 <- vector(length=n)
  for (j in 1:nsim) {
    division2[j] <- division[which(rank_theta==j)]
  }
  theta2 <- sort(theta_simuls_v2[[t]])
  quantiles <- cumsum(division2)/sum(division2)
  m1 <- length(quantiles[quantiles<=0.025])
  m2 <- length(quantiles[quantiles<=0.975])
  theta_coin_down_v[t] <- theta2[m1]
  theta_coin_up_v[t] <- theta2[m2]
}

```

```

out <- list(alpha_hat_l=alpha_hat_l,
            theta_hat_l=theta_hat_l,theta_hat_v=theta_hat_v,
            alpha_var_l=alpha_var_l,
            theta_var_l=theta_var_l,theta_var_v=theta_var_v,
            theta_coin_down_v=theta_coin_down_v,theta_coin_up_v=theta_coin_up_v,
            simul=simu)
invisible(out)
}

# Funktio: simulate
simulate <- function(a1,P1,y,Z,H,T,R,Q,nsim) {

n <- length(y)
m <- length(T[[1]][1,])
r <- length(Q[[1]][1,])

nw <- n+n*r

W <- matrix(c(0),nrow=nw,ncol=nw)

for (t in 1:n) {
  W[t,t] <- H[[t]][1,1]
}

for (t in 1:n) {
  for (t1 in 1:r) {
    for (t2 in 1:r) {
      W[n+(t-1)*r+t1,n+(t-1)*r+t2]<-Q[[t]][t1,t2]
      W[n+(t-1)*r+t2,n+(t-1)*r+t1]<-Q[[t]][t2,t1]
    }
  }
}

alpha_simuls<-vector("list",length=nsim)
for (k in 1:nsim) {
  alpha_simuls[[k]] <- makeMatrixlist(n,m,1)
}

theta_simuls <- vector("list",length=nsim)
for (k in 1:nsim) {
  theta_simuls[[k]] <- makeMatrixlist(n,1,1)
}

epsilon_pluss<-vector("list",length=nsim)
for (k in 1:nsim) {
  epsilon_pluss[[k]] <- makeMatrixlist(n,1,1)
}

eta_pluss<-vector("list",length=nsim)
for (k in 1:nsim) {
  eta_pluss[[k]] <- makeMatrixlist(n,r,1)
}

y_pluss<-vector("list",length=nsim)
for (k in 1:nsim) {
  y_pluss[[k]] <- makeMatrixlist(n,1,1)
}

y_pluss2<-vector("list",length=nsim)
for (k in 1:nsim) {
  y_pluss2[[k]] <- makeMatrixlist(n,1,1)
}

alpha_pluss<-vector("list",length=nsim)
for (k in 1:nsim) {
  alpha_pluss[[k]] <- makeMatrixlist(n+1,m,1)
}

```

```

}

alpha_pluss2<-vector("list",length=nsim)
for (k in 1:nsim) {
  alpha_pluss2[[k]] <- makeMatrixlist(n+1,m,1)
}

alpha_plus_hats<-vector("list",length=nsim)
for (k in 1:nsim) {
  alpha_plus_hats[[k]] <- makeMatrixlist(n,m,1)
}

alpha_hat <- statesmoothen(a1,P1,y,Z,H,T,R,Q)$alpha_hat

for (k in 1:nsim) {

  w_plus <- normsimu(0,W)

  epsilon_pluss[[k]] <- vectorToMatrixlist(w_plus[1:n],n)

  eta_pluss[[k]] <- vectorToMatrixlist(w_plus[-(1:n)],n)

  alpha_pluss[[k]][[1]] <- normsimu(a1,P1)

  for (t in 1:n) {
    y_pluss[[k]][[t]] <- Z[[t]]**alpha_pluss[[k]][[t]] + epsilon_pluss[[k]][[t]]
    alpha_pluss[[k]][[t+1]] <- T[[t]]**alpha_pluss[[k]][[t]]
      + R[[t]]**eta_pluss[[k]][[t]]
  }

  alpha_pluss[[k]] <- alpha_pluss[[k]][1:n]

  alpha_plus_hats[[k]] <- statesmoothen(a1,P1,y_pluss[[k]],Z,H,T,R,Q)$alpha_hat

  for (t in 1:n) {
    alpha_simuls[[k]][[t]]<- alpha_hat[[t]] - alpha_plus_hats[[k]][[t]] + alpha_pluss[[k]][[t]]
    theta_simuls[[k]][[t]] <- Z[[t]]**alpha_simuls[[k]][[t]]
  }

}

alpha_simuls_v <- vector("list",length=nsim)
for (k in 1:nsim) {
  alpha_simuls_v[[k]] <- matrixlistToVector(alpha_simuls[[k]])
}

theta_simuls_v <- vector("list",length=nsim)
for (k in 1:nsim) {
  theta_simuls_v[[k]] <- matrixlistToVector(theta_simuls[[k]])
}

out <- list(alpha_simuls_l=alpha_simuls,alpha_simuls_v=alpha_simuls_v,
           theta_simuls_l=theta_simuls,theta_simuls_v=theta_simuls_v)
invisible(out)
}

# Funktio: linearisation
linearisation <- function(a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start) {

  n <- length(y)
  m <- length(T[[1]][1,])
  r <- length(Q[[1]][1,])

  alpha_tildes <- vector("list",length=(itermax+2))
  for (iter in 1:(itermax+2)) {
    alpha_tildes[[iter]] <- makeMatrixlist(n,m,1)
  }
}

```

```

}

theta_tildes <- vector("list",length=(itermax+1))
for (iter in 1:(itermax+1)) {
  theta_tildes[[iter]] <- makeMatrixlist(n,1,1)
}

H_tildes <- vector("list",length=(itermax+1))
for (iter in 1:(itermax+1)) {
  H_tildes[[iter]] <- makeMatrixlist(n,1,1)
}

y_tildes <- vector("list",length=(itermax+1))
for (iter in 1:(itermax+1)) {
  y_tildes[[iter]] <- makeMatrixlist(n,1,1)
}

alpha_tildes[[1]] <- alpha_tilde_start

for (iter in 1:(itermax+1)) {

  for (t in 1:n) {

    theta_tildes[[iter]][[t]] <- Z[[t]]%*%alpha_tildes[[iter]][[t]]
    linearisation_iter_temp <- linearisationIter(t,theta_tildes[[iter]][[t]],y[[t]])
    H_tildes[[iter]][[t]] <- linearisation_iter_temp$H_iter_t
    y_tildes[[iter]][[t]] <- linearisation_iter_temp$y_iter_t

  }

  ss <- statesmooher(a1,P1,y_tildes[[iter]],Z,H_tildes[[iter]],T,R,Q)
  alpha_tildes[[iter+1]] <- ss$alpha_hat

}

out <- list(y_tildes_all=y_tildes,H_tildes_all=H_tildes,
           y_tildes_final=y_tildes[[itermax+1]],H_tildes_final=H_tildes[[itermax+1]],
           theta_tildes=theta_tildes, theta_tildes_final=theta_tildes[[itermax+1]],
           alpha_tildes=alpha_tildes[1:(itermax+1)], alpha_tildes_final=alpha_tildes[[itermax+1]])
invisible(out)
}

# Funktio: statefilter
statefilter <- function(a1,P1,y,Z,H,T,R,Q) {

  n <- length(y)
  m <- length(T[[1]][1,])
  r <- length(R[[1]][1,])

  a <- makeMatrixlist(n+1,m,1)
  P <- makeMatrixlist(n+1,m,m)
  v <- makeMatrixlist(n,1,1)
  F <- makeMatrixlist(n,1,1)
  K <- makeMatrixlist(n,m,1)
  L <- makeMatrixlist(n,m,m)
  M <- makeMatrixlist(n,m,1)

  a[[1]] <- a1
  P[[1]] <- P1

  for (t in 1:n) {
    v[[t]] <- y[[t]] - Z[[t]]%*%a[[t]]
    F[[t]] <- Z[[t]]%*%P[[t]]%*%t(Z[[t]]) + H[[t]]
    K[[t]] <- T[[t]]%*%P[[t]]%*%t(Z[[t]])%*%solve(F[[t]])
    L[[t]] <- T[[t]] - K[[t]]%*%Z[[t]]
  }
}

```

```

    a[[t+1]] <- T[[t]]%*%a[[t]] + K[[t]]%*%v[[t]]
    P[[t+1]] <- T[[t]]%*%P[[t]]%*%t(L[[t]]) + R[[t]]%*%Q[[t]]%*%t(R[[t]])
  }

  out <- list(v=v,F=F,K=K,L=L,a=a,P=P)
  invisible(out)
}

# Funktio: statesmoothen
statesmoothen<-function(a1,P1,y,Z,H,T,R,Q) {

  n <- length(y)
  m <- length(a1[,1])
  r <- length(R[[1]][1,])

  sf <- statefilter(a1,P1,y,Z,H,T,R,Q)
  v <- sf$v
  F <- sf$F
  L <- sf$L
  a <- sf$a
  P <- sf$P

  alpha_hat <- makeMatrixlist(n,m,1)

  r <- makeMatrixlist(n,m,1)
  N <- makeMatrixlist(n,m,m)
  V <- makeMatrixlist(n,m,m)

  theta_hat <- makeMatrixlist(n,1,1)

  r[[n]] <- matrix(c(0),nrow=m,ncol=1)
  N[[n]] <- matrix(c(0),nrow=m,ncol=m)

  for (t in n:2) {
    r[[t-1]] <- t(Z[[t]])%*%solve(F[[t]])%*%v[[t]] + t(L[[t]])%*%r[[t]]
    N[[t-1]] <- t(Z[[t]])%*%solve(F[[t]])%*%Z[[t]] + t(L[[t]])%*%N[[t]]%*%L[[t]]
    alpha_hat[[t]] <- a[[t]] + P[[t]]%*%r[[t-1]]
    V[[t]] <- P[[t]] - P[[t]]%*%N[[t-1]]%*%P[[t]]
    theta_hat[[t]] <- Z[[t]]%*%alpha_hat[[t]]
  }

  r0 <- t(Z[[1]])%*%solve(F[[1]])%*%v[[1]] + t(L[[1]])%*%r[[1]]
  N0 <- t(Z[[1]])%*%solve(F[[1]])%*%Z[[1]] + t(L[[1]])%*%N[[1]]%*%L[[1]]
  alpha_hat[[1]] <- a[[1]] + P[[1]]%*%r0
  V[[1]] <- P[[1]] - P[[1]]%*%N0%*%P[[1]]
  theta_hat[[1]] <- Z[[1]]%*%alpha_hat[[1]]

  out <- list(r=r,r0=r0,N=N,N0=N0,alpha_hat=alpha_hat,V=V,theta_hat=theta_hat)
  invisible(out)
}

# Funktio: disturbancesmoothen
disturbancesmoothen <- function(a1,P1,y,Z,H,T,R,Q) {

  n <- length(y)
  m <- length(a1[,1])
  r <- length(R[[1]][1,])

  epsilon_hat <- makeMatrixlist(n,1,1)
  eta_hat <- makeMatrixlist(n,r,1)
  varepsilon <- makeMatrixlist(n,1,1)
  varetay <- makeMatrixlist(n,r,r)

  sf <- statefilter(a1,P1,y,Z,H,T,R,Q)

```



```

sm <- statesmoothe(a1,P1,y,Z,H,T,R,Q)

F <- sf$F
v <- sf$v
K <- sf$K

r <- sm$r
N <- sm$N

for (t in n:1) {
  epsilon_hat[[t]] <- H[[t]]%*(solve(F[[t]])%*v[[t]] - t(K[[t]])%*x[[t]])
  varepsilony[[t]] <- H[[t]] - H[[t]]%*(solve(F[[t]]) + t(K[[t]])%*N[[t]]%*K[[t]])%*H[[t]]
  eta_hat[[t]] <- Q[[t]]%*t(R[[t]])%*r[[t]]
  varetay[[t]] <- Q[[t]] - Q[[t]]%*t(R[[t]])%*N[[t]]%*R[[t]]%*Q[[t]]
}

out <- list(epsilon_hat=epsilon_hat,varepsilony=varepsilony,eta_hat=eta_hat,varetay=varetay)
invisible(out)
}

# Funktio: normsimu
normsimu <- function(mu,var) {
  dime <- length(var[,1])
  B <- t(chol(var))
  simu <- as.matrix(mu + B%*rnorm(dime,mean=0,sd=1),nrow=1,ncol=dime)
  return(simu)
}

# Funktio: makeMatrixlist
makeMatrixlist <- function(listlength,numberofrow,numberofcol) {
  matrixlist<-vector("list",length=listlength)
  for (t in 1:listlength) {
    matrixlist[[t]] <- matrix(c(0),nrow=numberofrow,ncol=numberofcol)
  }
  return(matrixlist)
}

# Funktio: matrixlistToVector
matrixlistToVector <- function(listor) {
  n <- length(listor)
  numberofrow <- length(listor[[1]][,1])
  vectorn <- NULL
  for (t in 1:n) {
    for (r in 1:numberofrow) {
      vectorn<-c(vectorn,listor[[t]][r,1])
    }
  }
  return(vectorn)
}

# Funktio: vectorToMatrixlist
vectorToMatrixlist <- function(vectoror,n) {
  numberofrow <- length(vectoror)/n
  listn <- makeMatrixlist(n,numberofrow,1)
  temp <- 1
  for (t in 1:n) {
    for (r in 1:numberofrow) {
      listn[[t]][r,1] <- vectoror[temp]
      temp <- temp + 1
    }
  }
  return(listn)
}

```

```

#####

#Esimerkki ajojonosta.
#Aineistona 20-29-vuotiaiden naisten ikäryhmä.

y_v <- kn2029 #alkoholiin kuolleiden määrä
u_v <- vln2029 #väestömäärä
a1_va <- 0
P1_va <- 100000
itermax <- 20
omax <- 1000
a_t_s_va <- 0
initial <- c(-1,-1)
nsim <- 1000
ts_start <- 1969
m <- 2
r <- 2

n <- length(y_v)

y_ts <- ts(y_v,start=ts_start,end=ts_start+n-1)
u_ts <- ts(u_v,start=ts_start,end=ts_start+n-1)

ts.plot((y_ts/u_ts)*100000)

Qf <- function(psi) {
  Q<-vector("list",length=n)
  for (t in 1:n) {
    Q[[t]] <- matrix(c(exp(psi[1]),0,0,exp(psi[2])),nrow=r,ncol=r)
  }
  return(Q)
}

u <- u_v

l4 <- function(t,theta_t) {
  return(-u[t]*exp(theta_t))
}

linearisationIter <- function(t,theta_t,y_t) {
  H_iter_t <- (exp(-1*theta_t))/u[t]
  y_iter_t <- theta_t - 1 + (y_t*exp(-1*theta_t))/u[t]
  out <- list(H_iter_t=H_iter_t, y_iter_t=y_iter_t)
  invisible(out)
}

pYtCondThetat <- function(t,y_t,theta_t) {
  mu_t <- u[t]*exp(theta_t)
  p_y_t_cond_mu_t <- dpois(y_t,mu_t)
  return(p_y_t_cond_mu_t)
}

y <- vectorToMatrixlist(y_v,n)

a1<-matrix(c(a1_va),nrow=m,ncol=1)

P1<-matrix(c(P1_va,0,0,P1_va),nrow=m,ncol=m)

Z <- vector("list",length=n)
for (t in 1:n) {
  Z[[t]] <- matrix(c(1,0),nrow=1,ncol=m)
}

T <- vector("list",length=n)
for (t in 1:n) {

```

```

T[[t]] <- matrix(c(1,0,1,1),nrow=m,ncol=m)
}

R<-vector("list",length=n)
for (t in 1:n) {
  R[[t]] <- matrix(c(1,0,0,1),nrow=m,ncol=r)
}

alpha_tilde_start<-vector("list",length=n)
for (t in 1:n) {
  alpha_tilde_start[[t]] <- matrix(c(a_t_s_va),nrow=m,ncol=1)
}

o <- mleExpFamilyStart(a1,P1,y,Z,T,R,NULL,itermax,alpha_tilde_start,initial,omax)

psi_hat <- o$par

Q<-vector("list",length=n)
for (t in 1:n) {
  Q[[t]] <- matrix(c(exp(psi_hat[1]),0,0,exp(psi_hat[2])),nrow=r,ncol=r)
}

level <- statesmoootherExpFamily(a1,P1,y,Z,T,R,Q,itermax,alpha_tilde_start,nsim)

mu_hat_v <- u_v*exp(level$theta_hat_v)
mu_var_v <- (u_v^2)*((exp(level$theta_hat_v))^2)*level$theta_var_v
mu_coin_down_v <- u_v*exp(level$theta_coin_down_v)
mu_coin_up_v <- u_v*exp(level$theta_coin_up_v)

mu_hat_ts <- ts(mu_hat_v,start=ts_start,end=ts_start+n-1)
mu_var_ts <- ts(mu_var_v,start=ts_start,end=ts_start+n-1)
mu_coin_down_ts <- ts(mu_coin_down_v,start=ts_start,end=ts_start+n-1)
mu_coin_up_ts <- ts(mu_coin_up_v,start=ts_start,end=ts_start+n-1)

```