VAHINKOPROSESSIN MATEMAATTINEN MALLI JA SEN MAXIMUM LIKELIHOOD ESTIMOINTI

VAKUUTUSYHTIÖIDEN KESKUSLIITTO
VAKUUTUSKIRJASTO

Hely Salomaa

SHV-tutkielma

Joulukuu 1992

ABSTRACT

Hely Salomaa:

Maximum Likelihood Estimation of a Mathematical Model for the Claims Process

The loss reserve estimation is considered as the statistical estimation problem of a simple mathematical model proposed for the run-off data of settled claims. The entries of the data matrix are assumed to be independent random variables following a compound Poisson distribution with means, variances and skewnesses depending on few parameters. The compound Poisson distribution is approximated by the normal and gamma distributions and by the Wilson-Hilferty approximation. The parameters of the model are estimated from the observed run-off triangle using the maximum likelihood method. The likelihood function is derived for all three distributions mentioned above.

The estimate for the loss reserve is obtained as the sum of the expected values of the unobserved entries of the run-off data matrix. The variance estimate of the loss reserve estimator is calculated as well. If the normal distribution is assumed the distribution of the loss reserve estimator is also normal allowing confidence intervals to be calculated.

The connection of the chain ladder method to the model is also considered. Finally, two numerical examples are given to illustrate the estimation process.

SISÄLTÖ

1. JOHDANTO 1

2. RUN-OFF PROSESSI 3
2.1 Matemaattinen malli 3
2.2 Likelihood funktio 5

3. YKSINKERTAISTETTU MALLI 8
3.1 Mallin määrittely 8
3.2 Normaalijakauma 9
3.3 Gammajakauma 11
3.4 Wilson-Hilferty approksimaatio 13

4. CHAIN LADDER MENETELMÄ JA PARAMETRIESTIMAATTIEN ALKUARVOT 16
4.1 Chain ladder menetelmä 16
4.2 Parametriestimaattien alkuarvot 17

5. NUMEERISIA ESIMERKKEJÄ 20
5.1 Havaintojen generointi 20
5.2 Parametrien estimointi 21
5.3 Mallin ja havaintojen yhteensopivuus 23
5.4 Korvausvastuun estimointi 24

6. JOHTOPÄÄTÖKSET 27

LÄHDELUETTELO 28

LIITE 29

1. JOHDANTO

Vahinkoprosessin mallintamisella riskiteoriaan perustuen on pitkä perinne. Alan perusoppikirja Beard, Pentikäinen ja Pesonen (1984) käsittelee asiaa varsin laajasti. Riskiteoriaan perustuvan tarkastelun kohteena on yleensä ollut tilivuosittainen tai johonkin muuhun aikayksikköön kohdistuva kokonaisvahinkomeno ja tarkastelussa on lähdetty siitä, että vahinko korvataan samana vuonna jona se sattuu. Käytännössä tiettynä vuonna sattuneiden vahinkojen korvausten maksaminen saattaa kuitenkin tarkasteltavan vakuutusliikkeen luonteesta riippuen jakaantua hyvinkin monien vuosien osalle (ns. run-off-ilmiö). Vahinkomenon kohdistamiseksi oikeaan tilivuoteen on tilivuoden päättyessä arvioitava, paljonko siihen mennessä sattuneista vahingoista tullaan maksamaan korvauksia tulevina vuosina eli on arvioitava korvausvastuun määrä tilivuoden päättyessä. Korvausvastuu muodostuu tunnetuista, vielä maksamatta olevista vahingoista (outstanding claims) ja niistä ennen tilivuoden päättymistä sattuneista vahingoista, joista ei ole vielä saatu edes ilmoitusta (IBNR). Korvausvastuun arvioimista varten on kehitetty lukuisa joukko erityyppisiä menetelmiä. Rantala (1984) johti riskiteoriaa hyväksi käyttäen IBNR-laskukaavan, jossa tiettyä kantavuotta vastaava IBNR lasketaan ko. kantavuoteen liittyvän maksutulon ja tarkasteluhetkeen mennessä maksettujen korvausten painotettuna summana. Kertoimet on johdettu niin että IBNR-virheen varianssi minimoituu. Kertoimet riippuvat vahinkoprosessin parametreista, jotka oletetaan tunnetuiksi. Saatu IBNR-estimaattori on harhaton edellyttäen että maksutulo on todellinen riskimaksutulo.

Pentikäinen ja Rantala (1986, 1990) käsittelivät eri kantavuosiin liittyvien vahinkomenojen run-off-prosessia riskiteoreettisen mallin pohjalta. Edellisessä artikkelissa johdettiin kaavat mm. tilivuoden korvauskulujen varianssille ja tilivuoden loppua vastaavan korvausvastuun virheen eli run-off-virheen varianssille. Korvausvastuun laskennassa käytetty kaava, jossa korvausvastuu estimoidaan maksutulon ja maksettujen korvausten painotettuna summana, kattaa useat yleisesti käytetyt korvausvastuun laskentasäännöt. Korvausvastuun laskentasääntöjä vertailtiin ensin em. varianssien perusteella olettaen kaikki mallin parametrit sekä varausmenetelmän kertoimet tunnetuiksi. Toiseksi

run-off-virhettä ja sen vaikutusta korvauskuluihin tutkittiin samoilla oletuksilla tehtyjen simulointien perusteella. Vuoden 1990 artikkelissa run-off-virheen simulointitarkastelua jatkettiin muuten samoilla periaatteilla, mutta nyt varausmenetelmän kertoimet estimoitiin laskentahetkeen mennessä kertyneiden maksettujen korvausten perusteella. Vastaava run-off-malli on myös osana Pentikäisen, Bonsdorffin, Pesosen, Rantalan ja Ruohosen (1989) vakuutusyhtiön solvenssitarkastelun pohjana olevaa mallia.

Tässä työssä tarkastellaan mahdollisuuksia Pentikäisen ja Rantalan (1986) esittämän mallin parametrien maximum likelihood estimointiin tarkasteluhetkeen mennessä kertyneen vahinkohistorian perusteella. Mallin estimointi antaa mahdollisuuden myös tulevaisuudessa maksettavien korvausten arviointiin, ts. se tuottaa yhden korvausvastuun laskentamenetelmän. Kohdassa 2 kuvataan lyhyesti Pentikäisen ja Rantalan (1986) esittämä malli sekä tarkastellaan siihen liittyvää maximum likelihood estimointiprobleemaa. Kohdassa 3 esitetään yksinkertaistettu malli ja johdetaan likelihood funktio erilaisilla jakaumaoletuksilla. Kohdassa 4 tutkitaan korvausvastuun laskennassa käytetyn perinteisen chain ladder menetelmän yhteyttä yksinkertaistettuun malliin. Kohdassa 5 tarkastellaan esimerkinomaisesti yksinkertaistetun mallin maximum likelihood estimointia keinotekoisesti generoidusta datasta. Saatujen estimaattien perusteella lasketaan piste- ja väliestimaatit korvausvastuulle. Esimerkkien pohjalta tehdyt johtopäätökset on esitetty kohdassa 6.

2. RUN-OFF PROSESSI

2.1 Matemaattinen malli

Olkoon n_i vuonna i sattuneiden vahinkojen kokonaislukumäärä ja X_i niistä maksettava vahinkosumma. Vahingoista n_{i1} kpl maksetaan samana vuonna, n_{i2} kpl seuraavana vuonna jne. Vuonna $i+k-1$ maksetaan n_{ik} kpl vuonna i sattuneita vahinkoja vahinkosumman ollessa X_{ik}. Oletetaan että kaikki yhteen kantavuoteen liittyvät vahingot on maksettu t vuodessa. Tarkastellaan vuosina 1,2,...,t sattuneita vahinkoja. Esitetään kaikkiin t kantavuoteen liittyvät, t vuotena maksettavat vahinkosummat matriisina **X**:

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & X_{13} & \cdots & X_{1,t-2} & X_{1,t-1} & X_{1t} \\ X_{21} & X_{22} & X_{23} & \cdots & X_{2,t-2} & X_{2,t-1} & X_{2t} \\ X_{31} & X_{32} & X_{33} & \cdots & X_{3,t-2} & X_{3,t-1} & X_{3t} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ X_{t-2,1} & X_{t-2,2} & X_{t-2,3} & \cdots & X_{t-2,t-2} & X_{t-2,t-1} & X_{t-2,t} \\ X_{t-1,1} & X_{t-1,2} & X_{t-1,3} & \cdots & X_{t-1,t-2} & X_{t-1,t-1} & X_{t-1,t} \\ X_{t1} & X_{t2} & X_{t3} & \cdots & X_{t,t-2} & X_{t,t-1} & X_{tt} \end{bmatrix}$$

Jos prosessia tarkastellaan vuoden t lopussa, niin silloin tunnetaan muuttujien X_{ik}, $i = 1,...,t$; $k = 1,...,t-i+1$ arvot eli tavanomainen run-off-kolmio ja tehtävänä on estimoida jäljellä olevien muuttujien arvot, joiden summa on korvausvastuun määrä vuoden t lopussa:

$$C_t = \sum_{i=2}^{t} \sum_{k=t-i+2}^{t} X_{ik}.$$

Lukumäärät n_i oletetaan Mixed Poisson-prosesseiksi odotusarvona:

$$E(\boldsymbol{n}_i) = n\, l_n(i)\, \boldsymbol{q}_i,$$

missä n on vakio, $l_n(i)$ on volyymin muutosta kuvaava indeksi ($l_n(1) = 1$) ja $\boldsymbol{q}_i$ on struktuurimuuttuja. Struktuurimuuttuja voi olla deterministinen funktio tai stokastinen aikasarja, jonka kuitenkin edellytetään olevan heikosti stationaarinen, ts.

$$E(\boldsymbol{q}_i) = 1,\ Var(\boldsymbol{q}_i) = \sigma_q^2,\ Cov(\boldsymbol{q}_i, \boldsymbol{q}_{i+j}) = \gamma_j.$$

Muuttujat $\boldsymbol{X}_{ik}$, $k = 1,\ldots,t$ ovat ehdollisesti riippumattomia ehdolla $\boldsymbol{q}_i$ jakaumanaan Mixed Compound Poisson, niin että vahinkojen lukumäärän odotusarvo on

$$E(\boldsymbol{n}_{ik}) = E(\boldsymbol{n}_i\, p_k) = n\, l_n(i)\, \boldsymbol{q}_i\, p_k,$$

missä p_k on todennäköisyys sille, että vuonna i sattunut vahinko maksetaan vuonna $i+k-1$. Aikaisemmin oletettiin, että kaikki vahingot maksetaan t vuoden kuluessa, joten

$$\sum_{k=1}^{t} p_k = 1.$$

Todennäköisyydet p_k siis määräävät vahinkojen selviämisjakauman, joka voi vaihdella hyvinkin paljon tarkasteltavan vakuutusliikkeen luonteesta riippuen. Vuonna i sattuneesta vahingosta vuonna $i+k$-1 maksettavan summan $\boldsymbol{Z}_{ik}$ jakaumasta oletetaan, että se on aina saman muotoinen ainoastaan skaalauksen muuttuessa niin että

$$E(\boldsymbol{Z}_{ik}) = m\, l_m(i)\, g_m(k),$$

missä m on yksittäisen vahingon määrän odotusarvo vuonna 1 sattuneelle ja samana vuonna maksettavalle vahingolle, $l_m(i)$ on indeksi joka kuvaa yksittäisen vahingon odotusarvon muutosta vahingon sattumisvuoden mukana ja $g_m(k)$ indeksi joka kuvaa odotusarvon muutosta riippuen siitä miten kauan vahingon sattumisesta on kulunut vahingon maksamisajankohtaan. Indeksit skaalataan niin, että $l_m(1) = g_m(1) = 1$. Jos vuonna 1 sattuneen ja maksetun vahingon jakauman j. momentti on a_j, niin jakauman muodon säilyessä samana

$$E(\boldsymbol{Z}_{ik}^j) = a_j\, l_m(i)^j\, g_m(k)^j, \qquad j = 1,2,3,\ldots$$

Edellä tehtyjen oletusten perusteella on

$$E(\boldsymbol{X}_{ik}) = nm\, l_n(i)\, l_m(i)\, g_m(k)\, p_k$$

$$Var(\boldsymbol{X}_{ik}) = na_2\, l_n(i)\, l_m(i)^2\, g_m(k)^2\, p_k + n^2m^2 l_n(i)^2\, l_m(i)^2\, g_m(k)^2\, p_k{}^2\sigma_q^2.$$

2.2 Likelihood funktio

Tarkastellaan edellä kuvattua prosessia vuonna t, jolloin muuttujien $\boldsymbol{X}_{ik}$, $i = 1,...,t$; $k = 1,...,t\text{-}i\text{+}1$ arvot on havaittu. Merkitään näiden muuttujien yhteisjakauman tiheysfunktiota $f(\mathbf{X})$:llä ja muuttujien $\boldsymbol{X}_{ik}$, $k = 1,...,t\text{-}i\text{+}1$ yhteisjakauman tiheysfunktiota $f(\mathbf{X}_i)$:llä. Satunnaismuuttujat $\boldsymbol{X}_{ik}$, $k = 1,...,t\text{-}i\text{+}1$ ovat siis keskenään riippumattomia, kun q_i:n arvo on kiinnitetty, ts.

$$f(\mathbf{X}_i|\boldsymbol{q}_i) = \prod_{k=1}^{t-i+1} f(\boldsymbol{X}_{ik}|\, \boldsymbol{q}_i).$$

Muuttujan $\boldsymbol{X}_{ik}$ ehdollinen jakauma ehdolla $\boldsymbol{q}_i$ on Compound Poisson, jolle

$$E(\boldsymbol{X}_{ik}|\, \boldsymbol{q}_i) = nm\, l_n(i)\, l_m(i)\, g_m(k)\, p_k\, \boldsymbol{q}_i,$$

$$Var(\boldsymbol{X}_{ik}|\, \boldsymbol{q}_i) = na_2\, l_n(i)\, l_m(i)^2\, g_m(k)^2\, p_k\, \boldsymbol{q}_i \qquad \text{ja}$$

$$\gamma(\boldsymbol{X}_{ik}|\, \boldsymbol{q}_i) = n^{-1/2}\, a_2^{-3/2}\, a_3\, l_n(i)^{-1/2}\, p_k^{-1/2}\, \boldsymbol{q}_i^{-1/2}.$$

Jakaumaa voidaan approksimoida esim. normaalijakaumalla, NP-approksimaatiolla, gammajakaumalla jne, kts. Beard, Pentikäinen, Pesonen (1984).

Merkitään $\mathbf{q} = (\boldsymbol{q}_1,...,\boldsymbol{q}_t)'$. Silloin on

$$f(\mathbf{X}|\mathbf{q}) = \prod_{i=1}^{t} \prod_{k=1}^{t-i+1} f(\boldsymbol{X}_{ik}|\, \boldsymbol{q}_i).$$

Olkoon muuttujien $\boldsymbol{q}_i$ yhteisjakauman tiheysfunktio $f(\mathbf{q})$. Silloin

$$f(\mathbf{X}) = \int f(\mathbf{X}|\mathbf{q}) f(\mathbf{q}) d\mathbf{q} = E_{\mathbf{q}}(f(\mathbf{X}|\mathbf{q})).$$

Pentikäinen ja Rantala (1986) valitsivat $\boldsymbol{q}$-prosessiksi ensimmäisen kertaluvun autoregressiivisen aikasarjan

$$\boldsymbol{q}_i - 1 = a(\boldsymbol{q}_{i-1} - 1) + \varepsilon_i,$$

missä $0 \leq a < 1$ ja $\varepsilon_i \sim N(0,\sigma_\varepsilon^2)$. Tällöin $\mathbf{q}$:n asymptoottinen jakauma on t-ulotteinen normaalijakauma, jolle

$$E(\boldsymbol{q}_i) = 1, \qquad Var(\boldsymbol{q}_i) = \frac{\sigma_\varepsilon^2}{1-a^2} \qquad i = 1,\ldots,t$$

$$Cov(\boldsymbol{q}_i, \boldsymbol{q}_{i+j}) = a^j \frac{\sigma_\varepsilon^2}{1-a^2}, \qquad i = 1,\ldots,t,\ j = 1,\ldots,t-i.$$

Likelihood funktio f($\mathbf{X}$) on tässä tapauksessa moniulotteista normaalijakaumaa noudattavien satunnaismuuttujien $\boldsymbol{q}_i$ funktion $h(\mathbf{q}) = f(\mathbf{X}|\mathbf{q})$ odotusarvo. Yksinkertaisimmillaankin, eli valittaessa $\boldsymbol{X}_{ik}$:n ehdolliseksi jakaumaksi ehdolla $\boldsymbol{q}_i$ normaalijakauma, $f(\mathbf{X}|\mathbf{q})$ on suhteellisen monimutkainen muuttujien $\boldsymbol{q}_i$ funktio, joten sen odotusarvon laskeminen merkitsee t-ulotteisen integraalin numeerista laskentaa.

Jos aikasarjan sijasta oletetaan, että $\boldsymbol{q}_i$-muuttujat ovat keskenään riippumattomia ja noudattavat jotain spesifioitua jakaumaa, esim. gammajakaumaa, jolloin lukumääräprosessi $\boldsymbol{n}_i$ on Polya-prosessi (kts. Beard, Pentikäinen, Pesonen (1984)), niin tilanne yksinkertaistuu jonkin verran. Silloin

$$f(\mathbf{X}) = \prod_{i=1}^{t} \int \prod_{k=1}^{t-i+1} f(\boldsymbol{X}_{ik}|\boldsymbol{q}_i) f(\boldsymbol{q}_i) d\boldsymbol{q}_i.$$

Nyt on siis laskettava t kpl yksiulotteisia integraaleja likelihood funktion arvon määräämiseksi annetuilla muuttujien $\boldsymbol{X}_{ik}$ arvoilla. Likelihood funktion maksimointi mallin parametrien suhteen tunnetuilla $\boldsymbol{X}_{ik}$-arvoilla näyttää tässäkin tapauksessa varsin työläältä jollei suorastaan mahdottomalta.

Koska $\boldsymbol{q}_i$-muuttujien olettaminen riippumattomiksi jo sinänsä aika radikaalisti rajoittaa prosessia, niin luutavasti ei menetetä enää paljoakaan, jos luovutaan niiden satunnai-

suudesta ja kiinnitetään ne ykkösiksi. Silloin muuttujat $\boldsymbol{X}_{ik}$, $i = 1,...,t$; $k = 1,...,t\text{-}i\text{+}1$ ovat kaikki keskenään riippumattomia ja niiden jakauma on Compound Poisson, jolle

$$E(\boldsymbol{X}_{ik}) = nm\, l_n(i)\, l_m(i)\, g_m(k)\, p_k,$$

$$Var(\boldsymbol{X}_{ik}) = na_2\, l_n(i)\, l_m(i)^2\, g_m(k)^2\, p_k$$

$$\gamma(\boldsymbol{X}_{ik}) = n^{-1/2}\, a_2^{-3/2}\, a_3\, l_n(i)^{-1/2}\, p_k^{-1/2}.$$

Tällaisella menettelyllä päädytään malliin, jonka parametrien estimointi maximum likelihood menetelmällä on mahdollista.

3. YKSINKERTAISTETTU MALLI

3.1 Mallin määrittely

Oletetaan, että satunnaismuuttujat X_{ik}, $i = 1,...,t$; $k = 1,...,t-i+1$ ovat keskenään riippumattomia ja X_{ik}:n jakauma on sama kaikilla i:n ja k:n arvoilla, kuitenkin niin, että jakauman parametrit m_{ik}, σ^2_{ik} ja γ_{ik} riippuvat i:stä ja k:sta seuraavasti

$$
\begin{aligned}
m_{ik} &= m\, r_n^{i-1}\, r_m^{i-1}\, r_c^{k-1}\, p_k, \\
\sigma^2_{ik} &= v^2\, r_n^{i-1}\, r_m^{2(i-1)}\, r_c^{2(k-1)}\, p_k, \qquad (3.1) \\
\gamma_{ik} &= g\, r_n^{-(i-1)/2}\, p_k^{-1/2}.
\end{aligned}
$$

Parametrit on johdettu em. Compound Poisson jakauman odotusarvosta, varianssista ja vinoudesta olettamalla, että volyymi kasvaa vuosittain vakioprosentilla ($l_n(i) = r_n^{i-1}$) ja yksittäisen vahingon odotusarvo kasvaa vakioprosentilla sekä kantavuoden mukana että maksuvuoden mukana ($l_m(i) = r_m^{i-1}$, $g_m(k) = r_c^{k-1}$). Koska vahinkojen lukumääriä ei havaita, ei vakiota n ole mahdollista estimoida erikseen, joten yksinkertaistetun mallin parametrit on valittu niin, että nm korvataan parametrilla m, na^2 parametrilla v^2 ja $n^{-1/2}a_2^{-3/2}a_3$ parametrilla g. Itse jakaumaksi voidaan valita esim. normaalijakauma, gammajakauma tai Compound Poisson-jakauman NP- tai Wilson-Hilferty-approksimaatio. Jos valitaan normaalijakauma, niin parametri γ_{ik} ei ole mukana mallissa. Seuraavassa johdetaan mallin likelihood-funktio normaalijakauman, gammajakauman ja Wilson-Hilferty-approksimaation tapauksessa.

Todennäköisyyksiä p_k, $k = 1,...,t$ voitaisiin periaatteessa pitää estimoitavina parametreina, mutta käytännössä se ei ole mahdollista parametrien lukumäärän ollessa liian suuri käytettävissä olevien havaintojen lukumäärään nähden. Siksi todennäköisyydet kannattaakin mallintaa. Todennäköisyysjakaumaksi voitaisiin valita esim. Poisson-jakauma, jolloin

$$p_k = e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!}.$$

Todennäköisyysjakauman valinta mallia havaintoihin sovitettaessa riippuu siitä, millaisesta vakuutusliikkeestä on kyse. Poisson-jakauma ei tuota kovin monipuolista selviämisjakaumien parvea. Tässä työssä on Poisson-jakauman ohella käytetty yleistettyä negatiivista binomijakaumaa, jolle

$$p_k = \frac{\Gamma(h+k-1)}{\Gamma(h)(k-1)!}\left(\frac{\upsilon}{1+\upsilon}\right)^h \frac{1}{(1+\upsilon)^{k-1}}.$$

Tässä jakaumassa on kaksi parametria, joten sen tuottama todennäköisyysjakaumien valikoima on Poisson-jakaumaa laajempi. Rajatapauksessa ($h \to \infty$) se tuottaa Poisson-jakauman.

3.2 Normaalijakauma

Oletetaan, että $\boldsymbol{X}_{ik} \sim N(m_{ik}, \sigma_{ik}^2)$, missä m_{ik} ja σ_{ik}^2 ovat (3.1):n mukaiset. Mallin parametreina ovat m, v^2, r_n, r_m, r_c sekä todennäköisyyksien p_k parametrit (λ jos jakaumana Poisson tai h ja υ jos jakaumana negatiivinen binomijakauma). Jakauman vinouteen liittyvä parametri g ei ole mukana tässä mallissa. Muuttujan $\boldsymbol{X}_{ik}$ tiheysfunktio on siis

$$f(\boldsymbol{X}_{ik}) = (2\pi)^{-1/2}\, \sigma_{ik}^{-1}\, exp\left\{-\frac{1}{2}\frac{(\boldsymbol{X}_{ik} - m_{ik})^2}{\sigma_{ik}^2}\right\}.$$

Koska $\boldsymbol{X}_{ik}$:t ovat keskenään riippumattomia, niin likelihood funktio on

$$L = \prod_{i=1}^{t} \prod_{k=1}^{t-i+1} f(\boldsymbol{X}_{ik}),$$

jolloin sen logaritmi l puolestaan on

$$l = \sum_{i=1}^{t} \sum_{k=1}^{t-i+1} ln\, f(\boldsymbol{X}_{ik}) = vakio - \frac{1}{2}\sum_i \sum_k ln\, \sigma_{ik}^2 - \frac{1}{2}\sum_i \sum_k \frac{(\boldsymbol{X}_{ik} - m_{ik})^2}{\sigma_{ik}^2}.$$

Kun tähän sijoitetaan suureiden m_{ik} ja σ_{ik}^2 lausekkeet (3.1), saadaan

$$l = vakio - \frac{1}{4}t(t+1)\ ln\ v^2 - \frac{1}{12}t(t-1)(t+1)(ln\ r_n + 2\ ln\ r_m + 2\ ln\ r_c) - \frac{1}{2}\sum_{k=1}^{t}(t-k+1)ln\ p_k$$

$$-\frac{1}{2}\frac{1}{v^2}\sum_{i=1}^{t}\sum_{k=1}^{t-i+1}\frac{1}{r_n^{i-1}}\left[\frac{X_{ik}}{r_m^{i-1}r_c^{k-1}}\right]^2\frac{1}{p_k} + \frac{m}{v^2}\sum_i\sum_k\frac{X_{ik}}{r_m^{i-1}r_c^{k-1}} - \frac{1}{2}\frac{m^2}{v^2}\frac{1}{1-r_n}\sum_k p_k(1-r_n^{t-k+1}).$$

Estimoitavien parametrien lukumäärän pienentämiseksi inflaatiotekijät $r_m$ja r_c on oletettu tunnetuiksi kohdan 5 esimerkkitapauksissa. Jos näin menetellään, niin alkuperäiset havainnot X_{ik} voidaan jakaa tekijällä $r_m^{i-1}r_c^{k-1}$, jolloin saadaan muunnetut havainnot Z_{ik}. Merkitään vielä

$$Z.. = \sum_{i=1}^{t}\sum_{k=1}^{t-i+1} Z_{ik}\ (= \text{tunnettu vakio})$$

sekä $c_1 = t(t+1)/4$ ja $c_2 = t(t-1)(t+1)/12$ likelihood funktion logaritmi saadaan muotoon

$$l = vakio - c_1\ ln\ v^2 - c_2\ ln\ r_n - \frac{1}{2}\sum_{k=1}^{t}(t-k+1)ln\ p_k$$

$$-\frac{1}{2}\frac{1}{v^2}\sum_{i=1}^{t}\sum_{k=1}^{t-i+1}\frac{Z_{ik}^2}{r_n^{i-1}p_k} + \frac{m}{v^2}Z.. - \frac{1}{2}\frac{m^2}{v^2}\frac{1}{1-r_n}\sum_{k=1}^{t}p_k(1-r_n^{t-k+1}).$$

Tehtävänä on siis maksimoida funktio l parametrien m, v^2, r_n ja p_k:n parametri(e)n suhteen.

Kun parametrit on estimoitu, niitä käyttäen voidaan laskea arviot tuleville X_{ik}:n arvoille, $i = 2,...,t$; $k = t\text{-}i\text{+}2,...,t$. Koska X_{ik}:t ovat normaalisti jakautuneita, niin

(3.2) $$E(X_{ik}) = m_{ik} \quad \text{ja} \quad Var(X_{ik}) = \sigma_{ik}^2.$$

Näiden estimaatit saadaan korvaamalla parametrit kaavoissa (3.1) estimaateillaan. Korvausvastuu C_t on keskenään riippumattomien normaalisti jakautuneiden satunnaismuuttujien summana normaalisti jakautunut odotusarvon ja varianssin ollessa

$$(3.3) \qquad m_c = E(\boldsymbol{C}_t) = \sum_{i=2}^{t} \sum_{k=t-i+2}^{t} E(\boldsymbol{X}_{ik}) \quad ja \quad \sigma_c^2 = Var(\boldsymbol{C}_t) = \sum_{i=2}^{t} \sum_{k=t-i+2}^{t} Var(\boldsymbol{X}_{ik}).$$

Korvausvastuun piste-estimaattina voidaan käyttää sen odotusarvon maximum likelihood estimaattia $\hat{m}_c$, joka saadaan kaavasta (3.3) korvaamalla odotusarvot $E(\boldsymbol{X}_{ik})$ estimaateillaan. Normaalijakaumaan perustuen korvausvastuun estimaatille voidaan määrätä myös 1-α -tason luottamusväli:

$$\hat{m}_c \pm z_{1-\alpha/2}\, \hat{\sigma}_c,$$

missä $\Phi(z_{1-\alpha/2}) = 1-\alpha/2$ sekä $\hat{m}_c$ ja $\hat{\sigma}_c^2$ lasketaan kaavoista (3.3) korvaamalla $E(\boldsymbol{X}_{ik})$ ja $Var(\boldsymbol{X}_{ik})$ estimaateillaan.

3.3 Gammajakauma

Seuraavaksi oletetaan, että $\boldsymbol{X}_{ik}$ on gammajakautunut odotusarvona m_{ik}, varianssina σ_{ik}^2 ja vinoutena γ_{ik}, ts. muuttujan $\boldsymbol{X}_{ik}$ jakaumafunktio on

$$F(\boldsymbol{X}_{ik}) = \Gamma\left(\frac{4}{\gamma_{ik}^2} + \frac{2}{\gamma_{ik}} \frac{\boldsymbol{X}_{ik} - m_{ik}}{\sigma_{ik}}, \frac{4}{\gamma_{ik}^2}\right), \qquad \left(\boldsymbol{X}_{ik} \geq m_{ik} - 2\frac{\sigma_{ik}}{\gamma_{ik}}\right)$$

missä m_{ik}, σ_{ik}^2 ja γ_{ik} ovat (3.1):n mukaiset sekä

$$\Gamma(x,h) = \frac{1}{\Gamma(h)} \int_0^x e^{-z}\, z^{h-1}\, dz \qquad (x \geq 0,\ h \geq 0)\ .$$

Mallin parametrit ovat muuten samat kuin normaalijakauman tapauksessa, mutta nyt mukana on myös jakauman vinouteen liittyvä parametri g. Korvausvastuun $\boldsymbol{C}_t$ odotusarvo ja varianssi saadaan nytkin vastaavasti satunnaismuuttujien $\boldsymbol{X}_{ik}$, $i=2,\ldots,t$; $k=t-i+2,\ldots,t$ odotusarvojen ja varianssien summana, mutta $\boldsymbol{C}_t$:n jakauma ei enää ole normaalijakauma.

Muuttujan $\boldsymbol{X}_{ik}$ tiheysfunktio on

$$f(\boldsymbol{X}_{ik}) = 2\sigma_{ik}^{-1}\gamma_{ik}^{-1}\Gamma(4/\gamma_{ik}^2)^{-1}\left\{\frac{4}{\gamma_{ik}^2}\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(\boldsymbol{X}_{ik}-m_{ik})+1\right]\right\}^{\frac{4}{\gamma_{ik}^2}-1} \times$$
$$exp\left\{-\frac{4}{\gamma_{ik}^2}\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(\boldsymbol{X}_{ik}-m_{ik})+1\right]\right\},$$

jolloin likelihood funktion logaritmiksi saadaan

$$l = vakio - \frac{1}{2}t(t+1)\ ln(vg) - \frac{1}{6}t(t-1)(t+1)(ln\, r_m + ln\, r_c) - \sum_{i=1}^{t}\sum_{k=1}^{t-i+1} ln\,\Gamma(\frac{4}{g^2}r_n^{i-1}p_k)$$
$$+\sum_{i=1}^{t}\sum_{k=1}^{t-i+1}\left(\frac{4}{g^2}r_n^{i-1}p_k - 1\right) ln\left[\frac{2}{vg}\frac{\boldsymbol{X}_{ik}}{r_m^{i-1}r_c^{k-1}} + \frac{2}{g}\left(\frac{2}{g}-\frac{m}{v}\right)r_n^{i-1}p_k\right]$$
$$-\frac{2}{vg}\sum_i\sum_k\frac{\boldsymbol{X}_{ik}}{r_m^{i-1}r_c^{k-1}} - \frac{2}{g}\left(\frac{2}{g}-\frac{m}{v}\right)\sum_i\sum_k r_n^{i-1}p_k.$$

Samoin kuin normaalijakauman tapauksessa oletetaan inflaatiotekijät $r_m$ja r_c tunnetuiksi ja siirrytään käyttämään muunnettuja havaintoja $\boldsymbol{Z}_{ik} = \boldsymbol{X}_{ik}/(r_m^{i-1}r_c^{k-1})$. Vastaavasti kuin edellä $\boldsymbol{Z}$.. on havaintojen $\boldsymbol{Z}_{ik}$ summa. Jos nyt merkitään $c_1 = t(t+1)/2$ niin funktio l saadaan muotoon

$$l = vakio - c_1\, ln(vg) - \sum_i\sum_k ln\,\Gamma(\frac{4}{g^2}r_n^{i-1}p_k)$$
$$+\sum_i\sum_k\left(\frac{4}{g^2}r_n^{i-1}p_k - 1\right) ln\left[\frac{2}{vg}\boldsymbol{Z}_{ik} + \frac{2}{g}\left(\frac{2}{g}-\frac{m}{v}\right)r_n^{i-1}p_k\right]$$
$$-\frac{2}{vg}\boldsymbol{Z}_{..} - \frac{2}{g}\left(\frac{2}{g}-\frac{m}{v}\right)\sum_i\sum_k r_n^{i-1}p_k.$$

Tämän funktion maksimointi parametrien m, v^2, g, r_n ja p_k:n parametri(e)n suhteen on työläämpää kuin normaalijakauman tapauksessa, mutta ainakin esimerkkitapauksissa se pystyttiin suorittamaan.

3.4 Wilson-Hilferty approksimaatio

Oletetaan, että

$$X_{ik} = m_{ik} + \frac{1}{3}\sigma_{ik}\left\{\left(\frac{\gamma_{ik}}{6}\right)^2\left[Y_{ik} - \frac{\gamma_{ik}}{6} + \frac{6}{\gamma_{ik}}\right]^3 - \frac{6}{\gamma_{ik}}\right\}, \tag{3.4}$$

missä $Y_{ik} \sim N(0,1)$ ja m_{ik}, σ_{ik}^2 sekä γ_{ik} ovat (3.1):n mukaiset. Tässäkin mallissa X_{ik}:n jakauma oletetaan siis vinoksi. Satunnaismuuttujan X_{ik} kertymäfunktio on nyt

$$\begin{aligned} P\{X_{ik} \le x\} &= P\left\{Y_{ik} \le \frac{6}{\gamma_{ik}}\left[\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(x - m_{ik}) + 1\right]^{1/3} + \left(\frac{\gamma_{ik}}{6}\right)^2 - 1\right]\right\} \\ &= \Phi\left(\frac{6}{\gamma_{ik}}\left[\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(x - m_{ik}) + 1\right]^{1/3} + \left(\frac{\gamma_{ik}}{6}\right)^2 - 1\right]\right), \end{aligned}$$

missä Φ on standardoidun normaalijakauman kertymäfunktio. Satunnaismuuttujan X_{ik} tiheysfunktioksi saadaan

$$\begin{aligned} f(X_{ik}) = (2\pi)^{-1/2}\sigma_{ik}^{-1}&\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(X_{ik} - m_{ik}) + 1\right]^{-2/3} \times \\ exp&\left\{-\frac{1}{2}\left(\frac{6}{\gamma_{ik}}\right)^2\left[\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(X_{ik} - m_{ik}) + 1\right]^{1/3} + \left(\frac{\gamma_{ik}}{6}\right)^2 - 1\right]^2\right\}. \end{aligned}$$

Likelihood funktion logaritmi on silloin

$$\begin{aligned} l = vakio &- \frac{1}{2}\sum_{i=1}^{t}\sum_{k=1}^{t-i+1} ln\ \sigma_{ik}^2 - \frac{2}{3}\sum_{i=1}^{t}\sum_{k=1}^{t-i+1} ln\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(X_{ik} - m_{ik}) + 1\right] \\ &- \frac{1}{2}\sum_{i=1}^{t}\sum_{k=1}^{t-i+1}\left(\frac{6}{\gamma_{ik}}\right)^2\left\{\left[\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(X_{ik} - m_{ik}) + 1\right]^{1/3} + \left(\frac{\gamma_{ik}}{6}\right)^2 - 1\right\}^2. \end{aligned}$$

Sijoittamalla suureiden m_{ik}, σ_{ik}^2 ja γ_{ik} lausekkeet (3.1):stä saadaan

$$\frac{1}{2}\frac{\gamma_{ik}}{\sigma_{ik}}(\boldsymbol{X}_{ik}-m_{ik})+1=\frac{1}{2}\frac{g}{v}\left(\frac{1}{r_n^{i-1}}\frac{\boldsymbol{X}_{ik}}{r_m^{i-1}r_c^{k-1}}\frac{1}{p_k}-m\right)+1.$$

Samoin kuin edellisissä malleissa, nytkin oletetaan, että inflaatiotekijät $r_m$ja r_c ovat tunnettuja ja siirrytään muunnettuihin havaintoihin

$$\boldsymbol{Z}_{ik}=\frac{\boldsymbol{X}_{ik}}{r_m^{i-1}r_c^{k-1}}.$$

Silloin funktio l voidaan kirjoittaa muodossa

$$l=vakio-c_1\,ln\,v^2-c_2\,ln\,r_n-\frac{1}{2}\sum_{k=1}^{t}(t-k+1)\,ln\,p_k-\frac{2}{3}\sum_{i=1}^{t}\sum_{k=1}^{t-i+1}ln\left[\frac{1}{2}\frac{g}{v}\left(\frac{\boldsymbol{Z}_{ik}}{r_n^{i-1}p_k}-m\right)+1\right]$$
$$-\frac{1}{2}\left(\frac{6}{g}\right)^2\sum_{i=1}^{t}\sum_{k=1}^{t-i+1}r_n^{i-1}\left\{\left[\frac{1}{2}\frac{g}{v}\left(\frac{\boldsymbol{Z}_{ik}}{r_n^{i-1}p_k}-m\right)+1\right]^{1/3}+\left(\frac{g}{6}\right)^2\frac{1}{r_n^{i-1}p_k}-1\right\}^2 p_k,$$

missä kertoimet c_1 ja c_2 ovat samat kuin normaalijakauman yhteydessä. Tämän funktion maksimointi on vaikeusasteeltaan samaa luokkaa kuin gammajakaumaan perustuvaa mallia vastaavan likelihood funktion maksimointi.

Wilson-Hilferty jakauman tapauksessa satunnaismuuttujan $\boldsymbol{X}_{ik}$ odotusarvo ja varianssi poikkeavat hieman luvuista m_{ik} ja σ_{ik}^2, sillä määritelmän mukaan on

$$E(\boldsymbol{X}_{ik})=m_{ik}+\frac{1}{3}\sigma_{ik}\left\{\left(\frac{\gamma_{ik}}{6}\right)^2 E\left[\boldsymbol{Y}_{ik}-\frac{\gamma_{ik}}{6}+\frac{6}{\gamma_{ik}}\right]^3-\frac{6}{\gamma_{ik}}\right\}.$$

Koska satunnaismuuttuja $\boldsymbol{Y}_{ik}$ noudattaa standardoitua normaalijakaumaa, niin

$$E(\boldsymbol{Y}_{ik}^{2j-1})=0,\quad E(\boldsymbol{Y}_{ik}^{2j})=1\cdot3\cdot5\cdots(2j-1),\qquad j=1,2,\ldots$$

jolloin $\boldsymbol{X}_{ik}$:n odotusarvo on

$$E(\boldsymbol{X}_{ik}) = m_{ik} - \frac{1}{3}\sigma_{ik}\left(\frac{\gamma_{ik}}{6}\right)^5.$$

Vastaavasti saadaan $\boldsymbol{X}_{ik}$:n varianssiksi

$$E[\boldsymbol{X}_{ik} - E(\boldsymbol{X}_{ik})]^2 = E\left[\frac{1}{3}\sigma_{ik}\left\{\left(\frac{\gamma_{ik}}{6}\right)^2\left[\boldsymbol{Y}_{ik} - \frac{\gamma_{ik}}{6} + \frac{6}{\gamma_{ik}}\right]^3 - \frac{6}{\gamma_{ik}} + \left(\frac{\gamma_{ik}}{6}\right)^5\right\}\right]^2$$
$$= \sigma_{ik}^2\left\{\left(\frac{\gamma_{ik}}{6}\right)^8 - \frac{1}{3}\left(\frac{\gamma_{ik}}{6}\right)^4 + 1\right\}.$$

Korjaustekijän merkitys kummassakin lausekkeessa on mitätön, paitsi siinä tapauksessa että vinous on erittäin suuri.

Korvausvastuun odotusarvo ja varianssi voidaan estimoida kaavoista (3.3) korvaamalla $E(\boldsymbol{X}_{ik})$ ja $Var(\boldsymbol{X}_{ik})$ yo. lausekkeilla.

4. CHAIN LADDER MENETELMÄ JA PARAMETRIESTIMAATTIEN ALKUARVOT

4.1 Chain ladder menetelmä

Seuraavassa tarkastellaan korvausvastuun estimointia perinteisellä chain ladder menetelmällä, kun run-off-kolmion havainnot ovat peräisin kohdan 3 yksinkertaistetusta mallista.

Muodostetaan maksettujen korvausten kumulatiivinen run-off-kolmio:

$$\begin{bmatrix} S_{11} & S_{12} & \dots & S_{1,t-1} & S_{1t} \\ S_{21} & S_{22} & \dots & S_{2,t-1} & \\ \vdots & \vdots & & & \\ S_{t-1,1} & S_{t-1,2} & & & \\ S_{t1} & & & & \end{bmatrix}$$

missä $S_{ij} = \sum_{k=1}^{j} X_{ik}$. Chain ladder menetelmän perusmuodossa kehitysvuosittaiset kertoimet ovat

$$\hat{f}_k = \sum_{i=1}^{t-k} S_{i,k+1} \Big/ \sum_{i=1}^{t-k} S_{ik}.$$

Chain ladder menetelmän mukainen arvio vuoden i kokonaiskorvausmenolle on

$$\hat{S}_{it} = S_{i,t-i+1} \prod_{k=t-i+1}^{t-1} \hat{f}_k$$

jolloin korvausvastuun estimaatiksi vuoden t lopussa saadaan

$$\hat{C}_t = \sum_{i=2}^{t} \left(\hat{S}_{it} - S_{i,t-i+1} \right).$$

Vastaavasti korvausmenon lisäyksille $X_{i,t-i+j}$, $j=2,\ldots,i$ saadaan arviot

$$\hat{X}_{i,t-i+j} = \hat{S}_{i,t-i+j} - \hat{S}_{i,t-i+j-1} = S_{i,t-i+1}\left(\hat{f}_{t-i+j-1} - 1\right) \prod_{k=t-i+1}^{t-i+j-2} \hat{f}_k .$$

Jos oletetaan, että satunnaisvaihtelua ei esiinny eli $X_{ik} = m_{ik}$, niin silloin

$$f_k = \frac{\sum_{i=1}^{t-k}\sum_{j=1}^{k+1} m_{ij}}{\sum_{i=1}^{t-k}\sum_{j=1}^{k} m_{ij}} = \frac{\sum_{i=1}^{t-k}\sum_{j=1}^{k+1} m\, r_n^{i-1}\, r_m^{i-1}\, r_c^{j-1}\, p_j}{\sum_{i=1}^{t-k}\sum_{j=1}^{k} m\, r_n^{i-1}\, r_m^{i-1}\, r_c^{j-1}\, p_j} = \frac{\sum_{j=1}^{k+1} r_c^{j-1} p_j}{\sum_{j=1}^{k} r_c^{j-1} p_j}$$

joten

$$\hat{X}_{i,t-i+j} = \frac{\sum_{j=1}^{t-i+j-1} r_c^{j-1} p_j}{\sum_{j=1}^{t-i+1} r_c^{j-1} p_j} \; \frac{r_c^{t-i+j-1} p_{t-i+j}}{\sum_{j=1}^{t-i+j-1} r_c^{j-1} p_j} \sum_{j=1}^{t-i+1} m\, r_n^{i-1}\, r_m^{i-1}\, r_c^{j-1}\, p_j = m\, r_n^{i-1}\, r_m^{i-1}\, r_c^{t-i+j-1}\, p_{t-i+j} = m_{i,t-i+j} .$$

Chain ladder menetelmä tuottaa siis tarkalleen oikean korvausvastuun määrän C_t, jos $X_{ik} = m_{ik}$, $i=1,\ldots,t$; $k=1,\ldots,t$ ja m_{ik} on (3.1):n mukainen. Näin ollen chain ladder menetelmän voidaan perustellusti katsoa tuottavan erään approksimaation kohdassa 3 esitetyn yksinkertaistetun mallin mukaisen korvausvastuun estimaatille.

4.2 Parametriestimaattien alkuarvot

Chain ladder menetelmän yhteydessä tarkasteltuja kehitysvuosittaisia kertoimia voidaan käyttää hyväksi määrättäessä alkuarvoja kohdan 3 mallien mukaisille maximum likelihood estimaateille. Kohdan 4.1 kumulatiivisen run-off-kolmion sijasta määrätään kertoimet $\hat{f}_k^s$ nyt havaintojen X_{ik}, $i=1,\ldots,t$; $k=1,\ldots,t-i+1$ muodostamasta run-off-kolmiosta:

$$\hat{f}_k^s = \sum_{i=1}^{t-k} X_{i,k+1} \Big/ \sum_{i=1}^{t-k} X_{ik}, \qquad k = 1,\ldots,t-1.$$

Kun nämä kertoimet määrätään muuttujien X_{ik} odotusarvojen muodostamasta run-off-kolmiosta, niin saadaan

$$f_k^s = r_c\, p_{k+1} / p_k, \qquad k = 1,\ldots,t-1.$$

Korvataan kertoimet f_k^s havaituilla kertoimilla $\hat{f}_k^s$ ja lasketaan todennäköisyyksille p_k estimaatit rekursiivisesti kaavoista

$$\hat{p}_{k+1} = \hat{p}_k \, \hat{f}_k^s \, / \, r_c, \qquad k = 1,\ldots,t-1.$$

Kun vielä asetetaan ehto $\sum_{k=1}^{t} \hat{p}_k = 1$, voidaan määrätä $\hat{p}_1$:

$$\hat{p}_1 = \left[1 + \sum_{k=1}^{t-1} \left(\prod_{j=1}^{k} \hat{f}_j^s \right) / \, r_c^k \right]^{-1}.$$

Mikäli r_c tunnetaan, niin eo. kaavoista saadaan suoraan arviot todennäköisyyksille $p_1,\ldots,p_t$. Sovittamalla näihin lukuihin mallin mukainen todennäköisyysjakauma (esim. Poisson-jakauma tai yleistetty negatiivinen binomijakauma) saadaan alkuarvot todennäköisyysjakauman parametriestimaateille.

Seuraavaksi muodostetaan rivikertoimet

$$\hat{f}_i^r = \sum_{k=1}^{t-i} \boldsymbol{X}_{i+1,k} \, / \sum_{k=1}^{t-i} \boldsymbol{X}_{ik}, \qquad i = 1,\ldots,t-1,$$

jotka odotusarvojen mukaisesta run-off-kolmiosta laskettuina ovat

$$f_i^r = r_n \, r_m, \qquad i = 1,\ldots,t-1.$$

Kun r_m oletetaan tunnetuksi, niin r_n:lle saadaan estimaatti esim. seuraavasti

$$\hat{r}_n = \left[\prod_{i=1}^{t-1} \hat{f}_i^r \right]^{1/(t-1)} / \, r_m.$$

Parametrin m alkuarvoksi voidaan ottaa lukujen

$$\boldsymbol{Y}_{ik} = \boldsymbol{X}_{ik} \, / \left(\hat{r}_n^{i-1} \, r_m^{i-1} \, r_c^{k-1} \, \hat{p}_k \right), \qquad i = 1,\ldots,t; \; k = 1,\ldots,t-i+1$$

keskiarvo. Parametrin v^2 alkuarvona taas voidaan käyttää lukujen

$$V_{ik} = X_{ik} \,/\, \left(\hat{r}_n^{(i-1)/2}\, r_m^{i-1}\, r_c^{k-1}\, \hat{p}_k^{1/2}\right) - \hat{m}\, \hat{r}_n^{(i-1)/2}\, \hat{p}_k^{1/2}, \qquad i = 1,\ldots,t;\; k = 1,\ldots,t-i+1$$

neliöiden keskiarvoa. Näin voidaan laskea alkuarvot normaalijakaumamallin parametreille. Vinousparametrille g alkuarvon määrääminen on hankalampaa.

5. NUMEERISIA ESIMERKKEJÄ

5.1 Havaintojen generointi

Havaintojen generointi, likelihood-funktion maksimointi ja muut laskelmat suoritettiin Mathematica-ohjelmiston versioilla 2.0 ja 2.1 (Wolfram (1991)). Havaintoja generoitiin Wilson-Hilferty jakauman mukaisesta mallista tunnetuilla parametrien arvoilla. Selviämistodennäköisyyksiä lukuunottamatta parametrien arvot otettiin Pentikäisen ja Rantalan (1986) artikkelista: $m = 60$, $v^2 = 10$, $g = 10^{-3/2}$, $r_n = 1.03$, $r_m = 1.06$ ja $r_c = 1.08$. Selviämistodennäköisyydet pyrittiin valitsemaan niin, että ne vastaisivat suhteellisen hyvin Poisson-jakaumaa tai yleistettyä negatiivista binomijakaumaa, ja päädyttiin seuraaviin kahteen selviämisjakaumaan:

1. $t = 6$ ja todennäköisyydet ovat 0.55, 0.25, 0.11, 0.06, 0.02, 0.01.
2. $t = 10$ ja todennäköisyydet ovat 0.06, 0.16, 0.23, 0.22, 0.16, 0.09, 0.04, 0.02, 0.01, 0.01.

Näistä ensimmäinen, jota myös Pentikäinen, Bonsdorff, Pesonen, Rantala ja Ruohonen (1989) käyttivät esimerkkinä, vastaa hyvin yleistettyä negatiivista binomijakaumaa, ja toinen taas Poisson-jakaumaa.

Kummassakin esimerkkitapauksessa generoitiin $t \times t$-matriisi $\mathbf{X} = \{X_{ik}\}_{t \times t}$. Aluksi määrättiin matriisit $\mathbf{M} = \{m_{ik}\}_{t \times t}$, $\mathbf{S} = \{\sigma_{ik}^2\}_{t \times t}$ ja $\boldsymbol{G} = \{\gamma_{ik}\}_{t \times t}$ annetuilla parametrien m, v^2, g, r_n, r_m ja r_c sekä todennäköisyyksien p_k, $k=1,\ldots,t$ arvoilla. Seuraavaksi generoitiin t^2 riippumatonta havaintoa standardoidusta normaalijakaumasta ja muodostettiin niistä $t \times t$-matriisi $\mathbf{Y}$. Matriisin $\mathbf{X}$ alkiot laskettiin sitten kaavasta (3.4). Saadut matriisit $\mathbf{X}_{6\times6}$ ja $\mathbf{X}_{10\times10}$ on esitetty liitteessä.

5.2 Parametrien estimointi

Havaittuina X_{ik}:n arvoina käytettiin sekä $\mathbf{X}_{6\times6}$:n että $\mathbf{X}_{10\times10}$:n tapauksessa alkioita X_{ik}, $i=1,\ldots,t$; $k=1,\ldots,t-i+1$. Havaintoihin sovitettiin normaalijakaumamalli, gammajakaumamalli ja Wilson-Hilferty malli selviämistodennäköisyysjakaumana tapauksessa $t=6$ yleistetty negatiivinen binomijakauma ja tapauksessa $t=10$ Poisson-jakauma. Mallin parametrien estimoimiseksi maksimoitiin mallia vastaavan likelihood-funktion logaritmi havaituilla X_{ik}:n arvoilla tuntemattomien parametrien m, v^2, g, r_n sekä h ja υ tai λ suhteen. Inflaatioparametrit r_m ja r_c oletettiin tunnetuiksi. Taulukoissa 5.1 ja 5.2 on esitetty sekä kohdassa 4 esitetyllä menettelyllä lasketut parametriestimaattien alkuarvot että eri malleja vastaavat maximum likelihood estimaattien arvot. Parametrien h ja υ (taulukko 5.1) sekä λ (taulukko 5.2) alkuarvot on saatu sovittamalla alkuestimaatit-sarakkeessa annettuihin todennäköisyyksiin $\hat{p}_k$ vastaavasti yleistetty negatiivinen binomijakauma (taulukko 5.1) tai Poisson-jakauma (taulukko 5.2). Ko. jakaumat eivät siis tuota tarkalleen alkuestimaatit-sarakkeen todennäköisyyksiä. Sen sijaan muissa sarakkeissa ole-

Taulukko 5.1 Parametrien alkuestimaatit ja maximum likelihood estimaatit tapauksessa $t=6$.

Estimaatti	Alkuestimaatit	Normaali	Gamma	Wilson-Hilferty
$\hat{m}$	58.984	62.900	63.226	63.012
$\hat{v}^2$	15.901	15.000	14.744	14.771
$\hat{g}$	0.036		0.057	0.034
$\hat{r}_n$	1.053	1.025	1.024	1.025
$\hat{h}$	1.319	1.169	1.138	1.158
$\hat{\upsilon}$	1.546	1.365	1.311	1.344
$\hat{p}_1$	0.521	0.526	0.525	0.525
$\hat{p}_2$	0.264	0.260	0.258	0.259
$\hat{p}_3$	0.117	0.119	0.119	0.119
$\hat{p}_4$	0.054	0.053	0.054	0.054
$\hat{p}_5$	0.027	0.023	0.024	0.024
$\hat{p}_6$	0.017	0.010	0.011	0.010

Taulukko 5.2 Parametrien alkuestimaatit ja maximum likelihood estimaatit tapauksessa t=10.

Estimaatti	Alkuestimaatit	Normaali	Gamma	Wilson-Hilferty
$\hat{m}$	59.354	62.579	62.672	62.514
$\hat{v}^2$	9.214	13.833	11.290	11.358
$\hat{g}$	0.036		0.065	0.062
$\hat{r}_n$	1.039	1.018	1.016	1.018
$\hat{\lambda}$	2.973	2.859	2.836	2.859
$\hat{p}_1$	0.065	0.057	0.059	0.057
$\hat{p}_2$	0.166	0.164	0.166	0.164
$\hat{p}_3$	0.229	0.234	0.236	0.234
$\hat{p}_4$	0.202	0.223	0.223	0.223
$\hat{p}_5$	0.157	0.160	0.158	0.160
$\hat{p}_6$	0.085	0.091	0.090	0.091
$\hat{p}_7$	0.042	0.043	0.042	0.043
$\hat{p}_8$	0.024	0.018	0.017	0.018
$\hat{p}_9$	0.009	0.006	0.006	0.006
$\hat{p}_{10}$	0.022	0.002	0.002	0.002

vat todennäköisyyksien p_k estimaatit ovat ko. todennäköisyysmallista parametrien maximum likelihood estimaattien arvoilla lasketut todennäköisyydet. Tässä tapauksessa, kun mallin parametrien todelliste arvot tunnetaan, parametrin g estimaatin alkuarvoksi otettiin parametrin tunnettu arvo $10^{-3/2} = 0.036$.

Kummassakin taulukossa eri malleja vastaavat estimoidut selviämistodennäköisyysjakaumat ovat hyvin lähellä toisiaan ja vastaavat suhteellisen hyvin niitä jakaumia, joista havaitut run-off-kolmiot on generoitu. Parametrin m estimaatitkaan eivät juuri poikkea toisistaan. Oleellisin ero mallien välillä näkyy parametrin v^2 estimaatissa, joka normaalijakaumamallille on kummassakin esimerkkitapauksessa suurempi kuin gammajakauma- ja Wilson-Hilferty-malleille, joissa on mukana myös vinousparametri. Kuten voisi odottaakin, niin vinous-parametrin puuttuminen normaalijakaumamallista näyttää kasvattavan varianssiestimaattia. Kaiken kaikkiaan gammajakaumamallia ja Wilson-Hilferty

mallia vastaavat parametriestimaatit ovat esimerkkitapauksissa hyvin lähellä toisiaan. Kun verrataan saatuja parametriestimaattien arvoja havaitut run-off-kolmiot tuottaneen mallin parametreihin, niin suuremman havaintojen lukumäärän ansiosta tapauksessa t=10 parametrien m *ja* v^2 estimaatit ovat lähempänä todellisia parametrin arvoja kuin tapauksessa t=6.

5.3 Mallin ja havaintojen yhteensopivuus

Saaduilla parametrien maximum likelihood estimaattien arvoilla laskettiin sitten sovitettua mallia vastaavat odotusarvot $\hat{E}(\boldsymbol{X}_{ik})$, i=1,...,t; k=1,...,t, jotka on annettu liitteessä. Mallin ja havaintojen yhteensopivuutta tutkittiin tarkastelemalla odotusarvoestimaattien $\hat{E}(\boldsymbol{X}_{ik})$ poikkeamia havaituista arvoista $\boldsymbol{X}_{ik}$, i=1,...,t; k=1,...,t-i+1. Vertailun helpottamiseksi poikkeamista laskettiin kolme tunnuslukua: poikkeamien summa D_s, poikkeamien neliöiden summa D_{ns} ja keskimääräinen suhteellinen poikkeama D_r:

$$D_s = \sum_{i=1}^{t} \sum_{k=1}^{t-i+1} (\boldsymbol{X}_{ik} - \hat{E}(\boldsymbol{X}_{ik}))$$

$$D_{ns} = \sum_{i=1}^{t} \sum_{k=1}^{t-i+1} (\boldsymbol{X}_{ik} - \hat{E}(\boldsymbol{X}_{ik}))^2$$

$$D_r = \frac{2}{t(t+1)} \sum_{i=1}^{t} \sum_{k=1}^{t-i+1} |\boldsymbol{X}_{ik} - \hat{E}(\boldsymbol{X}_{ik})| / \boldsymbol{X}_{ik}.$$

Tunnuslukujen arvot eri malleille kummassakin esimerkkitapauksessa on esitetty taulukossa 5.3. Kummassakaan esimerkkitapauksessa eri mallien tunnusluvut eivät juuri

Taulukko 5.3 Tunnusluvut mallien ja havaittujen run-off-kolmioiden yhteensopivuudelle.

	t=6			t=10		
	Normaali	Gamma	Wilson-H	Normaali	Gamma	Wilson-H
D_s	0.021	0.020	0.102	-0.598	0.221	-0.639
D_{ns}	141.476	141.328	141.298	211.483	208.349	211.804
D_r	0.132	0.130	0.131	0.135	0.135	0.135

poikkea toisistaan. Keskimääräinen suhteellinen poikkeama on kaikille malleille n. 13.1% tapauksessa t=6 ja 13.5% tapauksessa t=10. Kaiken kaikkiaan näyttää siltä, että esimerkkitapauksissa kaikkien kolmen mallin sopivuus havaittuun run-off-kolmioon on samaa luokkaa, joten vinousparametrin mukanaolo ei vaikuta erityisen tarpeelliselta.

5.4 Korvausvastuun estimointi

Lopuksi laskettiin sovitettua mallia vastaava korvausvastuun estimaatti $\hat{C}_t = \sum_{i=2}^{t} \sum_{k=t-i+2}^{t} \hat{E}(X_{ik})$ ja verrattiin sitä "todelliseen" korvausvastuun arvoon $C_t = \sum_{i=2}^{t} \sum_{k=t-i+2}^{t} X_{ik}$. Korvausvastuun varianssin estimaattia varten laskettiin ensin estimaatit satunnaismuuttujien X_{ik}, i=2,...,t; k=t-i+2,...,t variansseille. Korvausvastuun varianssiestimaatti on näiden estimaattien summa.

Koska sovitetun mallin todennäköisyydet $\hat{p}_{t+1}$, $\hat{p}_{t+2}$,... eivät ole nollia, joskin ne ovat pieniä lukuja, niin korvausvastuun estimaatiksi pitäisi oikeastaan ottaa summa $\hat{\hat{C}}_t = \sum_{i=1}^{t} \sum_{k=t-i+2}^{\infty} \hat{E}(X_{ik})$ ja vastaavasti varianssiestimaattia laskettaessa mukaan pitäisi ottaa myös muuttujien X_{ik}, i=1...,t; k=t+1... varianssien estimaatit. Esimerkkitapauksessa t=6 laskettiin $\hat{\hat{C}}_6$ ja vastaava varianssiestimaatti ottamalla mukaan vastaavasti muuttujien X_{ik} odotusarvoestimaatit ja varianssiestimaatit k:n arvoon 15 asti. Vastaavasti $\hat{\hat{C}}_{10}$ laskettiin k:n maksimiarvona 25. Wilson-Hilferty mallin kohdalla odotusarvon kaavassa oleva vinoudesta γ_{ik} (ja näin ollen p_k:n käänteisluvusta) riippuva korjaustermi tulee suureksi k:n kasvaessa alkaen tuottaa itseisarvoltaan yhä suurempia negatiivisia odotusarvoja jopa niin että $\hat{\hat{C}}_{10}$ kokonaisuudessaan tuli negatiiviseksi, kun k:n maksimiarvo oli 25. Tästä on pääteltävä, että Wilson-Hilferty-mallin tapauksessa ei yleisesti ottaen ole järkevää käyttää ei-havaittujen X_{ik}:n arvojen estimaatteina odotusarvoja vaan lukuja m_{ik}.

Korvausvastuun odotusarvon ja keskihajonnan estimaatit on esitetty taulukoissa 5.4 ja 5.5. Vertailun vuoksi mukaan on otettu myös chain ladder menetelmän tuottama korvausvastuun estimaatti. Tässä tarkasteltaville malleille on ominaista, että korvausvastuun odotusarvo ja hajonta yleensä kasvavat vahinkojen selviämisajan jakauman

Taulukko 5.4 Korvausvastuun odotusarvon ja hajonnan estimaatit tapauksessa t=6. E($\boldsymbol{C}_6$)=82.100, Dev($\boldsymbol{C}_6$)=4.606 ja $\boldsymbol{C}_6$=77.039.

Estimaatti	Normaali	Gamma	Wilson-Hilferty	Chain-ladder
$\hat{C}_6$	86.888	87.557	87.200	88.223
$\hat{\sigma}_{C_6}$	5.668	5.629	5.631	
95 % luottamusväli	75.779-97.997			
$\hat{\hat{C}}_6 \ (k_{max}=15)$	93.118	94.319	88.525	
$\hat{\hat{\sigma}}_{C_6} \ (k_{max}=15)$	5.925	5.904	6.329	
95 % luottamusväli	81.505-104.731			

pidentyessä. Suhteessa odotusarvoon hajonta sen sijaan pienenee johtuen siitä,että suhteen

$$\frac{Dev(\boldsymbol{C}_t)}{E(\boldsymbol{C}_t)} = \frac{v}{m} \frac{\sqrt{\sum\sum r_n^{i-1} r_m^{2(i-1)} r_c^{2(k-1)} p_k}}{\sum\sum r_n^{i-1} r_m^{i-1} r_c^{k-1} p_k}$$

osoittajassa ja nimittäjässä olevat summalausekkeet ovat samaa suuruusluokkaa ja summattavien lukumäärän kasvaessa osoittajassa olevan summan neliöjuuri kasvaa nimittäjässä olevaa summaa hitaammin. Tapauksessa t=6 kaikki estimaatit ovat suurempia ja tapauksessa t=10 pienempiä kuin toteutunut korvausvastuun arvo ja korvausvastuun odotusarvo, kuitenkin lähempänä odotusarvoa kuin toteutunutta arvoa. Kummassakin tapauksessa chain-ladder-estimaatti on suurin kaikista estimaateista ja näin ollen jälkimmäisessä lähimpänä sekä toteutunutta korvausvastuun arvoa että odotusarvoa. Tapauksessa t=6 kaikki neljä korvausvastuun estimaattia ovat hyvin lähellä toisiaan eikä hajontaestimaateissakaan ole juuri eroja. Taulukon 5.5 maximum likelihood estimaateista normaalijakaumaestimaatti ja Wilson-Hilferty estimaatti poikkeavat suunnilleen saman verran toteutuneesta korvausvastuun arvosta ja korvausvastuun odotusarvosta gammajakaumaestimaatin poikkeaman ollessa selvästi suurempi. Normaalijakaumamallin hajontaestimaatti on jonkin verran muita suurempi.

Normaalimallin yhteydessä korvausvastuulle voidaan määrätä väliestimaatteja sen pe-

Taulukko 5.5 Korvausvastuun odotusarvon ja hajonnan estimaatit tapauksessa t=10. E(C_{10})=464.718, Dev(C_{10})=12.924 ja C_{10}=486.922.

Estimaatti	Normaali	Gamma	Wilson-Hilferty	Chain-ladder
$\hat{C}_{10}$	429.384	421.648	429.937	456.508
$\hat{\sigma}_{C_{10}}$	14.214	12.708	12.895	
95 % luottamusväli	401.525-457.243			
$\hat{\hat{C}}_{10}$ $(k_{max}=25)$	430.926	423.092		
$\hat{\hat{\sigma}}_{C_{10}}$ $(k_{max}=25)$	14.251	12.739		
95 % luottamusväli	402.996-458.856			

rusteella että C_t on normaalisti jakautunut. Taulukoihin on laskettu 95 % luottamusvälit; yhtä hyvin voitaisiin määrätä raja, jonka alapuolella korvausvastuun arvo on annetulla todennäköisyydellä. Korvausvastuun odotusarvo ja toteutunut arvo sisältyvät estimoituun 95 % luottamusväliin tapauksessa t=6, kun taas tapauksessa t=10 kyseinen väliestimaatti sijoittuu kokonaisuudessaan sekä odotusarvon että toteutuneen arvon alapuolelle.

Muuttujien X_{ik}, $k>t$ odotusarvot ovat merkittävästi nollaa suurempia tapauksessa t=6, mikä näkyy taulukossa 5.4 niin, että estimaatti $\hat{\hat{C}}_6$ on selvästi estimaattia $\hat{C}_6$ suurempi. Sen sijaan tapauksessa t=10 vastaavien odotusarvojen merkitys on vähäinen. Vastaavien varianssiestimaattien vaikutus korvausvastuun hajonnan estimaattiin on vähäinen kummassakin tapauksessa.

6. JOHTOPÄÄTÖKSET

Vaikka kahden esimerkkitapauksen perusteella ei kovin pitkälle meneviä johtopäätöksiä voikaan tehdä, niin jotakin kuitenkin voitaneen jo niidenkin perusteella sanoa. Ensinnäkin näyttäisi siltä, että normaalijakaumamalli pärjää suhteellisen hyvin kahteen muuhun verrattuna. Vinousparametrin estimoiminen tämäntyyppisistä havainnoista saattaa olla suhteellisen epävarmaa, joten sen ottaminen mukaan malliin ei antane sanottavaa hyötyä. Laskennallisesti normaalijakaumamallin estimointi on huomattavasti yksinkertaisempi molempiin muihin malleihin verrattuna. Gammajakauma- ja Wilson-Hilferty mallien estimoinnissa laskennalliset vaikeudet kasvavat koko ajan *t*:n kasvaessa. Laskenta käy mahdottomaksi kun *t* on suuri (>10) ja selviämistodennäköisyysjakaumana on yleistetty negatiivinen binomijakauma. Näin ollen käytännössä mahdolliseksi soveltaa jää ainoastaan normaalijakaumamalli. Normaalijakaumamallin lisäetuna on, että korvausvastuun jakauma on silloin normaalijakauma, mikä mahdollistaa väliestimaattien laskemisen. Estimaattien tarkkuuden parantamiseksi havaintojen lukumäärää voisi lisätä laajentamalla havaittua run-off-kolmiota niin että mukaan tulee vanhoja vuosia enemmän kuin *t* kpl vaikka kaikki korvaukset oletettaisiinkin tulevan maksetuiksi *t* vuoden kuluessa.

Harkittaessa mallin kehittämistä kannattaisi vinouden sijasta pyrkiä ottamaan mukaan muuttujien keskinäinen riippuvuus. Tähän antaisi ehkä mahdollisuuden havaitun run-off-kolmion tarkastelu reaalisaationa moniulotteisesta normaalijakaumasta, jossa muuttujien odotusarvot ja varianssit voisivat olla samat kuin kohdassa 3 ja muuttujien väliset kovarianssit mallinnettaisiin vastaavalla tavalla.

LÄHDELUETTELO

Beard, R. E., Pentikäinen, T. and Pesonen, E. (1984) *Risk Theory.* Chapman & Hall, London.

Pentikäinen, T., Bonsdorff, H., Pesonen, M., Rantala, J. and Ruohonen, M, (1989) *Insurance Solvency and Financial Strength.* Finnish Training and Publishing Company Ltd, Helsinki.

Pentikäinen,T. and Rantala, J. (1986) Run-off Risk as a Part of Claims Fluctuation. *ASTIN BULLETIN* Vol. 16, No.2.

Pentikäinen,T. and Rantala, J. (1990) On the Evaluation of the Run-off Inaccuracies in Claims Reserving. ASTIN Qolloquium, Monte Carlo.

Rantala, J. (1984) An Application of Stochastic Control Theory to Insurance Business. *Acta Universitatis Tamperensis,* **A**, 164.

Wolfram, S. (1991) *Mathematica.* Addison-Wesley, Redwood City, California.

Liite 1.

1. $t = 6$

Generoitu matriisi $\mathbf{X}_{6\times6}$:

31.505	16.722	8.381	3.160	1.832	1.246
36.978	18.393	11.956	4.912	2.486	0.506
46.260	18.972	9.392	6.692	1.488	0.943
40.701	23.567	7.609	5.903	1.772	1.702
40.980	29.665	11.595	7.171	2.281	1.273
49.225	21.985	12.173	5.743	2.284	0.219

Odotusarvojen estimaatit:

Normaalijakaumamalli; $\hat{m}$=62.900, $\hat{v}^2$=15.000, $\hat{r}_n$=1.025, $\hat{h}$=1.169 ja $\hat{\upsilon}$=1.365.

31.686	16.921	8.382	4.044	1.925	0.909
34.436	18.389	9.109	4.395	2.092	0.988
37.424	19.985	9.900	4.777	2.274	1.073
40.672	21.719	10.759	5.191	2.471	1.167
44.201	23.604	11.692	5.641	2.686	1.268
48.037	25.652	12.707	6.131	2.919	1.378

Gammajakaumamalli; $\hat{m}$=62.226, $\hat{v}^2$=14.744, $\hat{g}$=0.057 $\hat{r}_n$=1.024, $\hat{h}$=1.138 ja $\hat{\upsilon}$=1.311.

33.177	17.638	8.809	4.305	2.081	0.999
36.007	19.142	9.560	4.672	2.258	1.084
39.077	20.775	10.376	5.071	2.451	1.177
42.410	22.547	11.261	5.503	2.660	1.277
46.027	24.469	12.221	5.973	2.887	1.386
49.952	26.556	13.263	6.482	3.133	1.504

Wilson-Hilferty-malli; $\hat{m}$=63.012, $\hat{v}^2$=14.771, $\hat{g}$=0.034 $\hat{r}_n$=1.025, $\hat{h}$=1.158 ja $\hat{\upsilon}$=1.344.

33.092	17.656	8.777	4.256	2.038	0.969
35.946	19.179	9.534	4.623	2.214	1.052
39.046	20.833	10.356	5.022	2.405	1.143
42.414	22.630	11.249	5.455	2.612	1.242
46.072	24.581	12.219	5.926	2.838	1.349
50.045	26.701	13.273	6.437	3.083	1.465

Varianssiestimaatit:

Normaalijakaumamalli; $\hat{m}$=62.900, $\hat{v}^2$=15.000, $\hat{r}_n$=1.025, $\hat{h}$=1.169 ja $\hat{\upsilon}$=1.365.

2.813	2.136	1.562	1.128	0.809	0.578
3.019	2.293	1.677	1.211	0.868	0.620
3.240	2.461	1.800	1.299	0.932	0.665
3.478	2.641	1.932	1.395	1.000	0.714
3.733	2.835	2.073	1.497	1.073	0.766
4.006	3.043	2.225	1.606	1.152	0.823

Gammajakaumamalli; $\hat{m}$=62.226, $\hat{v}^2$=14.744, $\hat{g}$=0.057 $\hat{r}_n$=1.024, $\hat{h}$=1.138 ja $\hat{\upsilon}$=1.311.

2.782	2.108	1.548	1.125	0.813	0.585
2.983	2.261	1.660	1.206	0.871	0.628
3.200	2.425	1.781	1.294	0.935	0.673
3.432	2.601	1.910	1.388	1.003	0.722
3.681	2.789	2.049	1.488	1.075	0.774
3.948	2.992	2.197	1.596	1.153	0.831

Wilson-Hilferty-malli; $\hat{m}$=63.012, $\hat{v}^2$=14.771, $\hat{g}$=0.034 $\hat{r}_n$=1.025, $\hat{h}$=1.158 ja $\hat{\upsilon}$=1.344.

2.785	2.114	1.549	1.121	0.806	0.578
2.989	2.269	1.662	1.203	0.865	0.620
3.207	2.434	1.784	1.291	0.928	0.665
3.441	2.612	1.914	1.385	0.996	0.714
3.693	2.803	2.054	1.486	1.069	0.766
3.962	3.008	2.204	1.595	1.147	0.822

2. $t = 10$

Generoitu matriisi $X_{10\times10}$:

3.449	11.474	15.085	16.146	11.273	6.566	4.131	1.974	0.579	1.494
4.460	10.941	21.833	19.856	15.625	8.929	4.554	3.180	1.574	0.965
4.591	15.652	19.903	16.988	17.130	10.325	4.851	3.125	1.576	0.782
4.638	14.093	21.090	23.128	16.974	10.939	5.788	2.725	1.534	1.207
6.785	14.433	20.602	22.964	22.235	12.114	3.422	1.908	1.046	1.816
7.070	15.454	21.917	24.308	20.224	11.293	8.980	4.887	3.937	3.356
4.490	20.507	28.719	18.722	25.101	13.924	6.262	4.353	2.651	2.503
8.383	16.261	28.388	34.767	29.119	14.644	4.876	5.685	2.374	1.888
7.058	20.824	37.405	30.451	25.822	19.388	8.092	2.585	3.514	0.987
8.914	27.613	33.228	37.919	29.523	16.866	10.561	3.427	1.287	0.676

Odotusarvojen estimaatit:

Normaalijakaumamalli; $\hat{m}$=62.579, $\hat{v}^2$=13.833, $\hat{r}_n$=1.018, $\hat{\lambda}$=2.859.

3.589	11.080	17.103	17.601	13.585	8.388	4.316	1.904	0.735	0.252
3.872	11.955	18.455	18.992	14.658	9.051	4.657	2.054	0.793	0.272
4.178	12.900	19.913	20.493	15.817	9.766	5.025	2.216	0.855	0.293
4.509	13.919	21.487	22.112	17.067	10.538	5.422	2.391	0.923	0.317
4.865	15.019	23.185	23.859	18.415	11.371	5.851	2.580	0.996	0.342
5.249	16.206	25.017	25.745	19.871	12.269	6.313	2.784	1.075	0.369
5.664	17.487	26.994	27.780	21.441	13.239	6.812	3.004	1.159	0.398
6.112	18.869	29.127	29.975	23.135	14.285	7.350	3.242	1.251	0.429
6.595	20.360	31.429	32.344	24.964	15.414	7.931	3.498	1.350	0.463
7.116	21.969	33.913	34.900	26.936	16.632	8.558	3.774	1.457	0.500

Gammajakaumamalli; $\hat{m}$=62.672, $\hat{v}^2$=11.290, $\hat{g}$=0.065, $\hat{r}_n$=1.016, $\hat{\lambda}$=2.836.

3.676	11.260	17.244	17.606	13.481	8.258	4.216	1.845	0.706	0.240
3.961	12.133	18.580	18.970	14.526	8.898	4.542	1.987	0.761	0.259
4.268	13.073	20.020	20.440	15.651	9.587	4.894	2.141	0.820	0.279
4.599	14.085	21.571	22.023	16.864	10.330	5.273	2.307	0.883	0.301
4.955	15.177	23.242	23.730	18.170	11.131	5.682	2.486	0.952	0.324
5.339	16.353	25.043	25.568	19.578	11.993	6.122	2.679	1.026	0.349
5.753	17.620	26.983	27.549	21.095	12.922	6.597	2.886	1.105	0.376
6.198	18.985	29.074	29.684	22.729	13.923	7.108	3.110	1.191	0.405
6.679	20.456	31.327	31.983	24.490	15.002	7.658	3.351	1.283	0.437
7.196	22.041	33.754	34.461	26.388	16.165	8.252	3.611	1.382	0.470

Wilson-Hilferty-malli; $\hat{m}$=62.514, $\hat{v}^2$=11.358, $\hat{g}$=0.062, $\hat{r}_n$=1.018, $\hat{\lambda}$=2.859.

3.585	11.068	17.085	17.583	13.571	8.379	4.312	1.902	0.734	0.252
3.868	11.943	18.436	18.972	14.643	9.042	4.652	2.052	0.792	0.272
4.174	12.887	19.892	20.471	15.800	9.756	5.020	2.214	0.854	0.293
4.504	13.905	21.465	22.089	17.049	10.527	5.417	2.389	0.922	0.316
4.860	15.004	23.161	23.835	18.396	11.359	5.845	2.578	0.995	0.341
5.244	16.190	24.991	25.718	19.850	12.257	6.307	2.781	1.073	0.368
5.658	17.469	26.966	27.751	21.419	13.225	6.805	3.001	1.158	0.397
6.106	18.850	29.097	29.944	23.111	14.270	7.343	3.238	1.250	0.429
6.588	20.339	31.397	32.310	24.938	15.398	7.923	3.494	1.349	0.463
7.109	21.947	33.878	34.864	26.909	16.615	8.549	3.771	1.455	0.499

Varianssiestimaatit:

Normaalijakaumamalli; $\hat{m}$=62.579, $\hat{v}^2$=13.833, $\hat{r}_n$=1.018, $\hat{\lambda}$=2.859.

0.891	1.626	2.100	2.214	2.021	1.651	1.230	0.849	0.548	0.334
0.953	1.739	2.246	2.368	2.162	1.765	1.316	0.908	0.586	0.357
1.019	1.860	2.402	2.532	2.312	1.888	1.407	0.971	0.627	0.382
1.089	1.989	2.569	2.708	2.472	2.019	1.505	1.039	0.671	0.408
1.165	2.128	2.747	2.896	2.644	2.159	1.610	1.111	0.717	0.437
1.246	2.275	2.938	3.097	2.828	2.309	1.721	1.188	0.767	0.467
1.333	2.434	3.142	3.313	3.024	2.470	1.841	1.271	0.820	0.499
1.425	2.603	3.360	3.543	3.234	2.641	1.969	1.359	0.877	0.534
1.524	2.783	3.594	3.789	3.459	2.825	2.106	1.453	0.938	0.571
1.630	2.977	3.843	4.052	3.699	3.021	2.252	1.554	1.003	0.611

Gammajakaumamalli; $\hat{m}$=62.672, $\hat{v}^2$=11.290, $\hat{g}$=0.065, $\hat{r}_n$=1.016, $\hat{\lambda}$=2.836.

0.814	1.480	1.904	1.999	1.818	1.478	1.098	0.755	0.485	0.294
0.870	1.582	2.034	2.136	1.943	1.580	1.173	0.806	0.519	0.314
0.929	1.690	2.174	2.283	2.076	1.689	1.254	0.862	0.554	0.336
0.993	1.807	2.323	2.440	2.219	1.805	1.340	0.921	0.592	0.359
1.062	1.931	2.483	2.607	2.371	1.929	1.432	0.984	0.633	0.384
1.135	2.063	2.654	2.787	2.534	2.061	1.530	1.052	0.676	0.410
1.212	2.205	2.836	2.978	2.708	2.203	1.636	1.124	0.723	0.438
1.296	2.357	3.031	3.183	2.894	2.354	1.748	1.202	0.773	0.468
1.385	2.519	3.239	3.401	3.093	2.516	1.868	1.284	0.826	0.501
1.480	2.692	3.462	3.635	3.306	2.689	1.996	1.372	0.882	0.535

Wilson-Hilferty-malli; $\hat{m}$=62.514, $\hat{v}^2$=11.358, $\hat{g}$=0.062, $\hat{r}_n$=1.018, $\hat{\lambda}$=2.859.

0.807	1.474	1.903	2.006	1.832	1.496	1.115	0.769	0.497	0.302
0.863	1.576	2.035	2.145	1.959	1.600	1.192	0.823	0.531	0.323
0.923	1.686	2.176	2.294	2.095	1.711	1.275	0.880	0.568	0.346
0.987	1.803	2.328	2.454	2.240	1.830	1.364	0.941	0.608	0.370
1.056	1.928	2.489	2.624	2.396	1.957	1.459	1.007	0.650	0.395
1.129	2.062	2.662	2.807	2.562	2.093	1.560	1.077	0.695	0.423
1.208	2.205	2.847	3.002	2.740	2.238	1.668	1.151	0.743	0.452
1.291	2.358	3.045	3.210	2.931	2.393	1.784	1.231	0.795	0.484
1.381	2.522	3.256	3.433	3.134	2.560	1.908	1.317	0.850	0.517
1.477	2.697	3.483	3.672	3.352	2.737	2.041	1.408	0.909	0.553