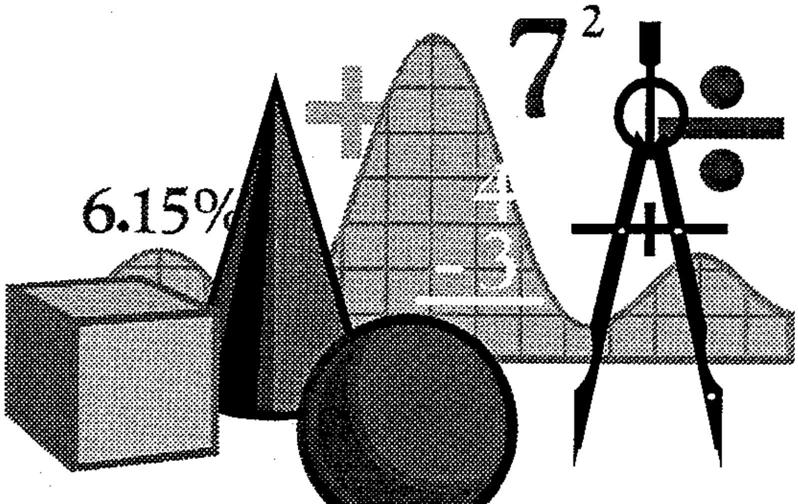
SHV-TUTKIELMA: KORVAUSVASTUUN ESTIMOINTIMENETELMÄT

KAI NIEMI, 1997



ABSTRACT

Niemi, K. (1997): Methods developed for estimating outstanding claims.

In practice, one of the most important tasks of actuary is to estimate outstanding claims i.e. how much insurance company has to pay in future claims which has already happened but are not yet fully or at all reported to company. In this study there is first introduced the general problematic involved with estimating these claims: the reasons that such claims are arisen, run-off pattern and the concept of run-off error. Then there is presented five different estimating methods developed for the purpose and at the end these methods are compared with simulation models and also one practical example is also presented.

1. Johdanto

VYL 10 luvun 2§ mukaisesti vakuutusyhtiön vakuutussopimuksista aiheutuva vastuu kirjataan vastuuvelaksi, jonka muodostavat vakuutusmaksuvastuu ja korvausvastuu. Näistä jälkimmäisen kyseisen laki määrittelee vastaavan jo sattuneiden vakuutustapahtumien johdosta suoritettavia, maksamatta olevia korvaus- ja muita määriä sekä runsasvahinkoisten vuosien varalta vastuuopillisesti laskettavaa tasoitusmäärää. Tässä tutkielmassa onkin tarkoitus lähemmin tarkastella kyseistä vastuuvelan osaa.

Tässä yhteydessä on tarkoitus esitellä erilaisia laskentamenetelmiä varsinaisen korvausvastuun (tasoitusmäärää ei huomioitu) määräämiseksi. Jatkossa puhuttaessa korvausvastuusta tarkoitetaan juuri tätä varsinaista korvausvastuuta. Korvausvastuun laskemismenettely riippuu merkittävästi vakuutuslajista ja sen luonteesta. Henkivakuutuksen yhteydessä korvausvastuun laskenta on usein huomattavasti yksinkertaisempaa kuin vahinkovakuutuksen yhteydessä. Perinteinen henkivakuutus ja erityisesti eläkevakuutus on säästömuoto, jossa vakuutusmaksuja maksetaan usein lähes kymmeniä vuosia ja maksetaan sitten kerralla ulos (säästöhenkivakuutus), jaksottain (eläkevakuutus) tai ainoastaan kuolemantapauksessa (riskihenkivakuutus). Säästö- ja riskihenkivakuutusmuodoissa korvausvastuu koostuu yleisesti ottaen jo tunnettujen maksamattomien ja vielä tuntemattomien kuolintapauskorvausten yhteismäärästä. Eläkevakuutuksen yhteydessä korvausvastuu määräytyy puolestaan hieman eri tavalla. Siinä korvaustapahtuma tapahtuu yleensä eläkkeensaajan saavuttaessa eläkeiän ja kyseisen korvaustapahtuman korvausmeno on periaatteessa tulevien maksettavien eläkkeiden pääoma-arvo ja tämä summa varataan korvausvastuuseen. Kyseisen pääoma-arvon laskennassa otetaan huomioon sekä korko että kuolevuus. Tuntemattomien varauksella ei puhtaassa eläkevakuutuksessa ole merkitystä. Kokonaisuudessaan siis henkivakuutuksessa korvausvastuun satunnaisuus riippuu pääasiassa vakuutettujen kuolevuudesta, jonka voidaan sanoa olevan suurta kantaa tarkastellessa erittäin stabiilia ja näin ollen aktuaarin on helppo hallita kyseistä satunnaisuutta. Lisäksi voisi mainita, että huomattavasti suuremman riskin henkivakuutuksessa muodostaa vastuiden sijoitustuotot eli se, että vastuille saatu sijoitustuotto ylittää käytetyn laskuperustekoron pitkällä aika välillä.

Vahinkovakuutuksessa korvausvastuun estimointi on yleisesti ottaen hankalempaa (riippuen tietenkin vakuutuslajista), koska on olemassa monta eri tekijää, jotka aiheuttavat satunnaisuutta vahinkoprosessiin. Kyseinen satunnaisuus yleensä suurenee sen mukaan mitä pitkähäntäisempää harjoitettu liike on eli kuinka kauan voi vahingon selviäminen kestää. Yleisesti ottaen siis henkivakuutuksessa korvausvastuut määrätään pääasiassa vakuutuskohtaisesti ja vahinkovakuutuksessa käytetään enemmän kollektiivisia menetelmiä. Tässä yhteydessä esiteltävät korvausvastuun estimointimenetelmät soveltuvatkin parhaiten juuri pitkähäntäisen liikkeen yhteydessä käytettäviksi, joissa vaaditaan kollektiivisia menetelmiä.

Tutkielma koostuu kuudesta eri luvusta. Luvussa 2 esitellään korvausvastuun ongelmaa yleisesti ja esitellään korvausvastuun estimointimenetelmien rakennemalli eli run-off-malli. Luvussa 3 tullaan esittelemään 5 eri korvausvastuun estimointimenetelmää ja käydään läpi niiden periaatteet. Luvussa 4 on esitetty simulointimalleilla saadut tulokset ja luvussa 5 on esitetty käytännön esimerkki vastuuvakuutuksen korvausvastuun estimoinnista.

Tämä tutkielma perustuu pääasiassa tekijän Pro-Gradu-tutkielmaan. Tässä yhteydessä kyseistä aikaisempaa työtä on muokattu paremmin sopivammaksi SHV-tutkielman raameihin: käsitteitä on lisätty ja korjattu ja kieliasua hieman muokattu asiaankuuluvammaksi. Lisäksi alkuperäisestä työstä on jätetty ns. esitietoja esittämättä johtuen yliopistossa tehdyn tutkielman yleisluontoisuudesta - lukijakunnalle vakuutusmatemaattiset perusteet eivät välttämättä ole tuttuja. Täysin uutta on puolestaan käytännöstä poimittu esimerkki (luku 5). Simulointimallit ovat pitkälti muokattu Pentikäisen ja Rantalan julkaisujen pohjalta.

2. Yleistä korvausvastuusta

Korvausvastuun käsite on muodostunut täysin käytännön realiteettien pakosta. Yleensä, kun vahinko on sattunut, niin kestää jonkin aikaa ennenkuin vakuutusyhtiö saa edes tiedon asiasta. Lisäksi vakuutusyhtiön saatua tiedon vahingosta kestää jälleen oman aikansa ennenkuin vakuutusyhtiö on käsitellyt vahingon yksityiskohdat ja on valmis maksamaan korvauksen. Tämä aikaviive riippuu monesta asiasta ja sitä aiheuttaa mm.

- korvauskäsittelyprosessi,
- korvauksen luonne eli korvaus on sellainen, että sitä ei korvata kerralla. Tällainen on esimerkiksi eläke,
- asian riitautuminen, jolloin joudutaan asiaa käsittelemään oikeudessa.

Joissakin äärimmäisissä olosuhteissa tämä aikaviive saattaa olla jopa useita vuosikymmeniä. Vakuutusyhtiön pitää kuitenkin tilinpäätöshetkellä antaa sijoittajille, omistajille ja myös viranomaisille kuva yhtiön taloudellisesta tilanteesta. Jotta yhtiö voisi antaa todellisen kuvan taloudellisesta tilanteestaan, pitää jollakin tavalla arvioida näitä maksamattomia korvauksia ja asettaa tätä arviota vastaava reservi menojäämänä yhtiön taseeseen. Tämä reservi vastaa johdannossa esitetyn lain pykäliä ja sitä kutsutaan siis **korvausvastuuksi** (*reserve for outstanding claims*).

Näitä tulevaisuudessa maksettavia korvauksia (*outstanding claims*) on perusolemukseltaan kahdenlaisia; on niin kutsuttuja **tunnettuja (avoimia) korvauksia** (*open claims*) ja niin kutsuttuja **IBNR** (*Incurred But Not Reported*) eli **tuntemattomia korvauksia**. Näiden kahden korvaustyypin pääasiallinen ero on kuten niiden nimetkin osoittavat, että tunnetuiksi korvauksiksi luokitellaan ne korvaukset, jotka on ilmoitettu vakuutusyhtiölle, mutta joita ei ole vielä (kokonaan) maksettu. Näiden korvauksien katsotaan olevan usein sellaisia, että niille voidaan määrätä reservi suhteellisen tarkasti yksinkertaisesti tarkastelemalla kyseisiä vahinkoja tapauskohtaisesti. Joissakin tilanteissa tämä ei ole kuitenkaan mielekästä kuten esimerkiksi silloin, kun tunnettuja korvauksia on suuri määrä. Tällöin näihin voidaan käyttää samoja menetelmiä kuin IBNR–korvausten laskemiseen. Kirjallisuudessa tunnettuja korvauksia kutsutaan myös **IBNER** (*Incurred But Not Enough Reported*)-korvauksiksi.

IBNR–korvauksiksi katsotaan puolestaan ne tulevaisuudessa maksettavat korvaukset, jotka eivät ole tunnettuja. Lisäksi IBNR–korvauksille laskettuun reserviin voidaan katsoa kuuluvan myös lisäreservi tunnettuihin korvauksiin mahdollisesti liittyvästä epävarmuudesta ja lisäreservi niille vahingoille, jotka on jo maksettu mutta tulee mahdollisesti tulevaisuudessa uudelleen käsiteltäviksi. Myöhemmin tullaan esittelemään joitakin korvausvastuun estimointimenetelmiä ja näitä menetelmiä tullaan myös keskenään vertailemaan simulointimalleilla. Tällöin korvausvastuulla tullaan tarkoittamaan lähinnä pelkästään IBNR–korvauksia.

Tämän tutkielman pääideana on tutkia juuri korvausvastuun estimointia ja tämän korvausvastuun estimaatin ja todellisten maksettujen korvausten erotuksesta johtuvan niin kutsutun **kehitysvirheen** (*run-off error*) vaihtelua ja sen suuruutta. Korvausvastuun oikea estimointi on siis erittäin oleellista vakuutusyhtiön kannalta ei ainoastaan siinä mielessä, että saadaan oikea kuva vakuutusyhtiön tuloksesta tiettynä tilivuonna, vaan myös sen takia, että systemaattinen korvausvastuun aliarvioiminen johtaa liian alhaisiin vakuutusmaksuihin. Toisaalta sen yliarvioiminenkaan ei ole hyväksi, koska tällöin vakuutusyhtiön kilpailukyky heikkenee liian korkeiden vakuutusmaksujen takia.

Suomessa on yleisesti ottaen omaksuttu käytäntö estimoida korvausvastuuta harhattomasti eli sen odotusarvon mukaisesti. Tällöin korvausvastuun estimaatin ja todellisten korvausten välisen erotuksen katsotaan heikentävän vakuutusyhtiön tulosta kyseisenä vuotena, mikäli erotus on ollut negatiivinen tai parantavan vakuutusyhtiön tulosta kyseisenä vuotena, mikäli estimaatti on ollut todellisia kustannuksia suurempi. Joskus harhattomuusoletuksesta tietoisesti luovutaan. Tällöin on yleensä ideana se, että korvausvastuu on eräänlaista velkaa vakuutusyhtiölle vakuutuksenottajilta ja näin ollen vakuutusyhtiön kannattaa sijoittaa kyseinen rahamäärä takaisin rahamarkkinoille ja odottaa tällöin saavansa sijoitustuottoa sijoituksistaan. Kyse on siis korvausvastuun diskonttauksesta eli korvausvastuun sen hetkisen pääoma-arvon määräämisestä. Luonnollisestikaan diskonttauksella ei ole suurta merkitystä silloin, jos on oletettavissa tiettynä vuonna tapahtuneiden vahinkojen suhteellisen nopea selviäminen (2–3 vuotta). Kuitenkin mikäli kaikkien vahinkojen selviäminen ja niiden maksaminen vakuutusyhtiön puolesta voi viedä jopa vuosikymmeniä, niin tällöin diskonttauksella on huomattava merkitys. Käytännössä on olemassa kaksi eri mielipidettä siitä, että pitääkö tai oikeastaan voidaanko korvausvastuuta laskettaessa ottaa huomioon diskonttaus vai ei. Korvausvastuuta diskontatessahan sitoudutaan johonkin tiettyyn tulevaisuuden sijoitustoiminnasta saatavaan tuottoon.

Sosiaali- ja terveysministeriön Määräys- ja ohjekokoelmassa kotimaisille vakuutusyhtiöille asetetaan erinäisiä ehtoja korvausvastuun määräämiselle ja erityisesti sen diskonttauksessa käytetylle korkokannalle. Määräykset on hieman erilaiset riippuen onko kyseessä henki- vai vahinkovakuutusyhtiö. Henkivakuutuksen korvausvastuun diskonttauksessa käytettävä korkokanta riippuu ensinnäkin vakuutuksen luonteesta (unit-linked, with-profit, non-profit) ja toiseksi vastuuta kattavan omaisuuden luonteesta. Yleisesti ottaen voisi sanoa, että yli 4.5% korkokannan käyttö ei ole sallittua ylijäämän jakoon osallisissa vakuutuksissa. Vahinkovakuutuksessa elinkorkojen diskonttauksessa pätee samat säännöt kuin henkivakuutuksessa. Muiden kuin elinkorkovastuiden diskonttaukselle on annettu tiukat ehdot. Näistä vahinkovakuutusyhtiöiden on ilmoitettava ministeriölle tilinpäätöksen liitetietoina korvausvastuun laskemisessa käytetystä korkoutuksesta seuraavat tiedot:

- Vahinkoryhmät, joihin korkoutusta on käytetty, korvausten keskimääräiset selviämisajat ja käytetty korkokanta.

- Korvausvastuun ja jälleenvakuuttajien osuuden osalta vastuun bruttomäärä ennen korkoutusta, korkoutuksen määrä ja vastuun nettomäärä.
- Samat tiedot myös edelliseltä tilikaudelta.

Määräys- ja ohjekokoelmassa sanotaan erityisesti, että diskonttausta ei saa käyttää mikäli vahinkoryhmän vahinkojen keskimääräinen selviämisaika on alle neljä vuotta. Lisäksi käytettävä korkokanta pitää valita vastuuta kattavan omaisuuden turvaavasti määrättyä sijoitustuottoa vastaten kuitenkin siten, että korkokanta ei saa olla yli 5%.

Aivan viimeaikoina on ollut paljon puhetta laskuperustekorkojen eli diskonttauksessa käytettyjen korkokantojen alentamisesta vallalla olevan matalan korkotason vuoksi, jonka voidaan olettaa pysyvän myös tulevaisuudessa alhaisena johtuen Suomen EMU-kytkennästä ja Keski-Euroopassa kauan vallinneista matalista korkotasoista. Käytännössä käytettävää diskonttauskorkoa ei herkästi mennä laskemaan, koska esimerkiksi yhden prosentin laskemisella on merkittäviä muutoksia vastuiden absoluuttiseen määrään erityisesti pitkähäntäisissä vakuutuslajeissa (7-12%) ja kyseinen vastuun lisäys pitää kattaa yhtiön omilla varoilla.

Kannattaa myös pitää mielessä, että käytännössä korvausvastuun estimaatti on alttiina huomattavissa määrin myös muille virheille, joita tässä tutkielmassa ei tulla käsittelemään lainkaan. Kehitetty matemaattinen malli on aina vain idealisaatio todellisesta ongelmasta ja näin ollen kyseinen malli sisältää virhettä jo itsestäänkin. Toisaalta myös tämän approksimoivan mallin parametrit on sovitettu tilastollisesti kerättyyn tilastoaineistoon, jolloin syntyy jälleen virhettä. Parametrien estimoinnissa onkin tapana käyttää jonkinlaista varovaisuusperiaatetta eli ettei "todennäköisyys korvausvastuun osoittautumisesta estimoinnin vuoksi liian pieneksi" ole liian suuri.

Jatkossa tullaan olettamaan, että jokaista vahinkoa vastaa yksi korvausmaksu eli vahinko korvataan kerralla kokonaan. Tämä ei välttämättä vastaa aivan todellisuutta, mutta eri estimointimenetelmät ja näiden vertailumenetelmät soveltuvat suoraan myös tapauksiin, joissa on mahdollista tapahtua useampia korvausmaksukertoja koskien yhtä vahinkoa. Tällöin tietenkin täytyy huomioida erilainen tilastoaineiston käsittely, parametrien estimointi, jne..

2.1. Rakennemalli (Run-off–malli)

Korvausvastuun estimoinnissa lähdetään yleensä liikkeelle siitä, että vakuutusyhtiöllä jo tiedossa olevat korvaukset luokitellaan eri luokkiin sen mukaan, miltä vuodelta ne ovat alkuperäisin eli siis minä vuonna ne ovat tapahtuneet. Yhtä tällaista luokkaa kutsutaan **kohortiksi** (*cohort*). Vastaavasti nämä eri kohortit jaetaan vielä lisäksi luokkiin sen mukaan, milloin kyseiset vahingot ovat ilmoitettu vakuutusyhtiölle eli kuinka monta vuotta varsinaisen vahingon tapahtumisesta on kulunut. Näitä kohorttien eri luokkia kutsutaan puolestaan **soluiksi**. Tätä lähestymistapaa kutsutaan **vahinkovuosi analyysiksi** (*year of accident analysis*). Näin ollen siis vaaditaan, että jokaisesta vahingosta on määrättävissä sattumis- ja maksamishetki. Tämä vaatimuksen toteutuminen voi maallikosta tuntua itsestään selvältä, mutta käytännössä on kuitenkin olemassa sellaisia vakuutuslajeja, joille vahingon tapahtumishetken määrääminen ei ole yksiselitteistä.

Olkoon $X(t;\ s,\ s)$ niiden vahingonkorvauksien summa, jotka ovat sattuneet vuonna t ja maksettu vuonna $t + s$. Parametri s ilmaisee kohortin **kehitysajanhetken** (*development time*). Vastaavasti voidaan määritellä

(2.1) $$X(t;s_1,s_2) = \sum_{i=s_1}^{s_2} X(t;i,i)$$

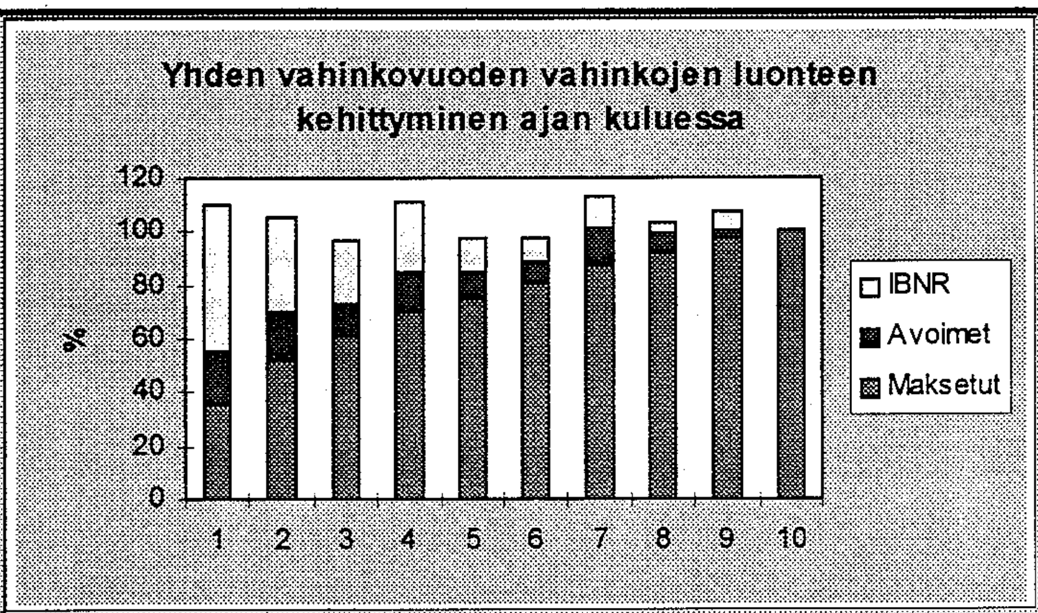
tarkoittamaan *kohortin t aikavälillä* $(t + s_1,\ t + s_2)$ *maksettujen* korvausten summaa,

(2.2) $$X(t;0,\infty) = \sum_{i=0}^{\infty} X(t;i,i)$$

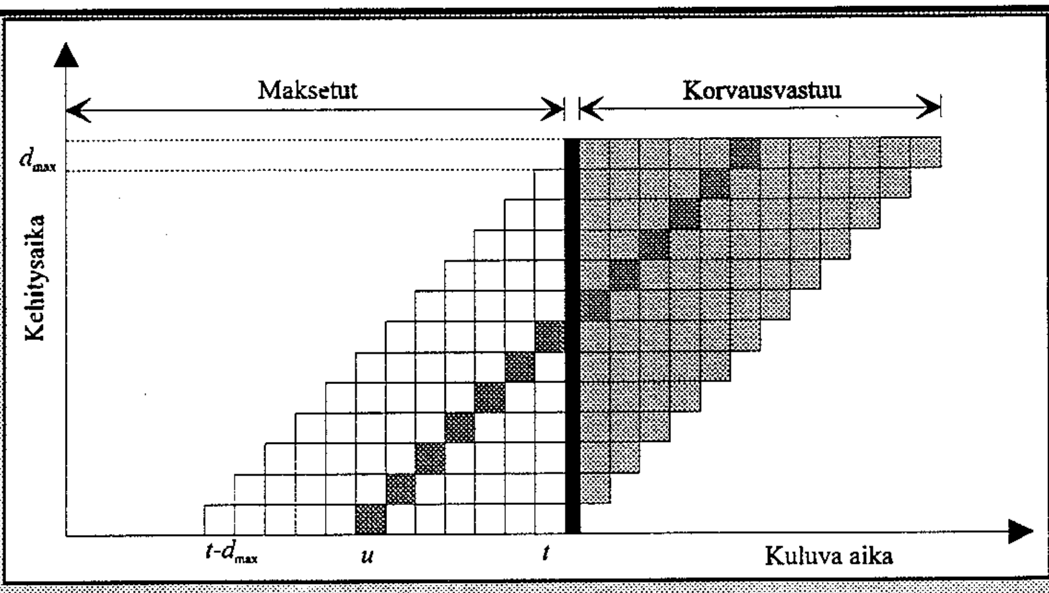
vuonna *t tapahtuneiden* vahinkojen kokonaissummaa ja

(2.3) $$X_p(t) = \sum_{s=0}^{\infty} X(t-s;s,s)$$

vuonna *t maksettujen* korvauksien kokonaismäärää. Käytännön sovellusten kannalta on yleensä syytä olettaa, että on olemassa jokin maksimiaika $d_{max} + 1$, jonka aikana kaikki yhtenä vuonna sattuneet vahingot ehditään ilmoittamaan, käsittelemään ja korvaamaan. Kuvassa 2.1 on graafisesti kuvattu yhden kohortin rakennetta ja sen kehittymistä ajan kuluessa. Kuvan pylväiden tumman harmaan osan eli maksettujen korvausten osuuden kehittymistä verrattuna lopulliseen korvausmenoon kutsutaan myös kyseisen kohortin vahinkojen **selviämisjakautumaksi**. Kyseisessä kuvassa olevien pylväiden vaihteleva korkeus johtuu juuri korvausvastuuseen liittyvästä epävarmuudesta.



Kuva 2.1. Yhden kohortin rakenne ja sen kehittyminen ajan kuluessa. Pylväiden epäsäännöllinen kehittyminen johtuu korvausvastuuseen liittyvästä epävarmuudesta.



Kuva 2.2. Yllä on esitelty vuoden t kohorttien kehittymiskaaviota. Tumman harmaalla väritetty diagonaali kuvaa yhden kohortin kehityskaarta (ks. kuva 2.1). Vuonna t kyseisen kohortin vahingoista on maksettu solujen (u, u), $(u+1, u+1)$, ..., (t, t) korvaukset $X_p(u,u), X_p(u+1,u+1), \ldots, X_p(t,t)$. Vastaavasti tilivuoden t korvausvastuu määräytyy harmaalla väritettyjen solujen summasta.

Vastaavasti kuvassa 2.2 on havainnollistettu korvausvastuun muodostumista (tummemmalla varjostettu alue). Kyseinen kolmion muotoinen alue pitää siis sisällään ne vahingot, joita ei ole vielä ilmoitettu vakuutusyhtiölle, mutta jotka ovat jo kuitenkin tapahtuneet. Näin ollen korvausvastuu $C(t)$ muodostuu juuri kyseisen kolmion muodostavien solujen estimoiduista vahinkomenoista eli

(2.4a) $$C(t) = \text{estimaatti summalle} \sum_{s=0}^{d_{max}-1} X(t-s; s+1, d_{max}).$$

Vastaavasti määritellään myös kohortin t korvausvastuu vuoden $t+s$ lopussa eli

(2.4b) $$C(t, s) = \text{estimaatti termille } X(t; s+1, d_{max}).$$

Kuten edellä jo mainittiin, niin yleensä vaaditaan, että ennuste 2.4b on harhaton eli

(2.5) $$E[C(t,s)] = E[X(t; s+1, d_{max})].$$

Vaihtoehto yllä olevalla ehdolle olisi vaatia ainoastaan kokonaiskorvausvastuun harhattomuutta eli

(2.6) $$E[C(t)] = E\left[\sum_{s=0}^{d_{max}-1} X(t-s; s+1, d_{max})\right].$$

Tämä onkin usein parempi vaihtoehto, koska tällöin korvausvastuun estimaatin variaatio on yleensä vähäisempää . Harhattomuusoletuksen ideana on taustalla oletus, ettei korvausvastuun estimaattiin sisällytetä mitään varmuusmarginaaleja mahdollisille odotusarvosta ylöspäin poikkeaville eli erittäin vahinkorikkaille korvausvuosille. Kyseisen marginaalin oletetaankin yleensä olevan sisällytetty jollakin tavalla vakuutusyhtiön tasoitusmäärään. Nyt kannattaa kuitenkin kiinnittää huomiota siihen tosiasiaan, että jos diskonttausta ei suoriteta (ks. seuraava kappale), niin tällöin periaatteessa korvausvastuuseen sisältyy oma epämääräinen marginaali, joka syntyy rahoille saatavasta sijoitustuotosta.

Mikäli korvausvastuulle halutaan suorittaa diskonttaus, niin silloin korvausvastuun $C(t)$ määritelmä muuttuu hieman eli

(2.7) $$C(t) = \text{estimaatti summalle} \sum_{s=1}^{d_{max}} X_p(t+s) \cdot v(t,s)$$

ja vastaavasti myös kohortin t korvausvastuun määritelmä muuttuu hieman. Edellä $v(t, s)$ on diskonttaustekijä ja sen määritelmä on tässä yhteydessä

(2.8) $$v(t,s) = \left(\frac{1}{1+j(t+s)}\right)^{\frac{1}{2}} \cdot \prod_{i=1}^{s-1} \frac{1}{1+j(t+i)},$$

missä puolestaan $j(t)$ on vuoden t sijoituskorko. Käytännön sovelluksissa pitää yleensä yllä esitettyä diskonttausmallia muokata tilanteeseen sopivaksi.

Kuvan 2.2 vaaleista soluista muodostunutta kolmiota eli sitä osaa kohorteista, jotka ovat jo tiedossa, kutsutaan **kehityskolmioksi** (*run-off triangle*). Tarkemmin ottaen kyseisen kolmion muodostavat solut voivat sisältää mitä tahansa satunnaismuuttujia, jotka liittyvät jotenkin vahinkovuoden $t - s$ jo havaittuun aineistoon. Tällaisia ovat esimerkiksi normeeratut kokonaisvahinkomenot, joista on poistettu inflaatio ja vakuutusyhtiön liikevolyymin vaikutus ja kumulatiiviset tai solukohtaiset vahinkomenot. Tarkemmin kehityskolmiosta selitetään luvussa 3.

2.2. *Kehitysvirhe (Run-off–virhe)*

Edellä määriteltiin kehitysvirheen muodostuvan korvausvastuun estimaatin ja todellisten maksettujen korvausten välisestä erotuksesta. Kyseistä virhettä voidaan tarkastella kahdesta eri näkökulmasta. Toinen on niin kutsuttu **break-up–näkökulma** ja toinen puolestaan niin kutsuttu **going-concern–näkökulma**.

2.2.1. Break-up

Ideana tässä tavassa on saada selville, mikä on kokonaisvirhe eli juuri kehitysvirhe vuoden t korvausvastuun estimaatissa $C(t)$. Tämä saadaan selville jättämällä huomioon ottamatta enää uusia vahinkovuosia ja antamalla keskeneräisten kohorttien kehittyä loppuun eli tilaan, jossa tapahtuneet vahingot ovat kaikki tiedossa ja maksettuina siten, ettei niihin liity enää minkäänlaista satunnaisuutta. Kehitysvirhe $R(t)$ on siis

$$\mathrm{R}(t) = \mathrm{C}(t) - \sum_{s=0}^{d_{\max}-1} \mathrm{X}(t-s; s+1, d_{\max}), \tag{2.9}$$

missä summaus otetaan siis kuvan 2.2 tummemmalla väritetystä alueesta. Jos korvausvastuu halutaan diskontata, täytyy yllä olevaa lauseketta hieman muuttaa. Tällöin on olemassa kaksi vaihtoehtoa kehitysvirheen määräämiseksi. Toinen yksinkertainen tapa on kertoa lausekkeen 2.9 oikean puolen summaus tekijä diskonttaustekijällä eli

$$\mathrm{R}(t) = \mathrm{C}(t) - \sum_{s=0}^{d_{\max}-1} \sum_{u=s+1}^{d_{\max}} \mathrm{X}(t-s; u, u) \cdot v(t, u-s). \tag{2.10a}$$

Vaihtoehtoinen tapa on matkia selviytymisprosessia ajan kuluessa eli maksaa korvaukset eräänlaisesta rahastosta, jonka vuosittain jäljellä olevasta määrästä saadaan sijoitustuloa. Tätä menetelmää voidaan approksimoida asettamalla ensin $C_r(t) = C(t)$ ja iteroimalla tästä eteenpäin kaavaa

$$\begin{aligned} C_r(t+s) &= C_r(t+s-1) \cdot (1 + \mathrm{j}(t+s)) \\ &\quad - (1 + \mathrm{j}(t+s))^{\frac{1}{2}} \cdot \sum_{d=s-1}^{d_{\max}-1} X(t+s-1-d; d+1, d+1) \end{aligned} \tag{2.10b}$$

niin kauan kuin kehittymisvuosia on jäljellä. Näiden yllä esitetyn kahden menetelmän ero on, että kaavalla (2.10a) saatu kehitysvirhe on tavallaan kaavan (2.10b) kehitysvirhe diskontattuna korvausvastuun määräämisvuoteen.

Käytännössä kehitysvirheen $R(t)$ arvoa ei tiedetä tarkasti vasta kuin useiden vuosien mahdollisesti jopa vuosikymmenien kuluttua, kun kaikki kohortit ovat täysin korvattuja. Näin ollen sitä voidaan käyttää ainoastaan *a posteriori* tarkistuksena käytetylle korvausvastuun estimointimenetelmälle. Simulointimalleja käytettäessä yllä esitelty tapa on sovelias, koska tarvittavat suureet ovat helposti tallennettavissa tietokoneen muistiin ja suureet ovat lopuksi helppo laskea.

2.2.2. Going-concern–näkökulma

Tässä tavassa on tarkoitus tutkia kehitysvirheen vaikutusta vuoden t **korvauskuluun** (*incurred claims, aggregate loss*) $X(t)$. Vuoden t korvauskulun määritellään olevan

(2.11) $$X(t) = X_p(t) + C(t) - C(t-1)$$

eli korvauskulu vuonna t muodostuu vuonna t maksetuista korvauksista sekä korvausvastuun estimaatin muutoksesta. Kun aikaisempiin kohortteihin liittyvä vahinko korvataan vuonna t, voidaan vapauttaa sille varattu reservi, joka on siis sisällytetty estimaattiin $C(t-1)$. Mikäli kyseinen reservi oli täsmälleen yhtäsuuri kuin maksettu korvaus, niin niiden vaikutukset kumoavat toisensa ja niistä ei muodostu lisäystä kehitysvirheeseen. Koska varattu reservi oli kuitenkin vain estimaatti maksettavalle korvaukselle, niin näin ei juuri tapahdu.

Korvauskulun määritelmä muuttuu jälleen hieman, jos korvausvastuu halutaan diskontata. Tällöin kyseistä korvauskulua voidaan approksimoida lausekkeella

(2.12) $$X(t) = X_p(t) + C(t) - C(t-1) - j(t)\cdot\frac{C(t-1)+C'(t)}{2}.$$

Edellä $C'(t)$ on se osa vuoden $t-1$ korvausvastuusta, joka sisältyy myös vuoden t korvausvastuuseen.

2.3. *Eri menetelmien vertailusta*

Kehitysvirhe $R(t)$ ja sen vaikutus korvauskuluun $X(t)$ riippuu luonnollisesti sekä korvausprosessista että tavasta, jolla korvausvastuuta estimoidaan. Eri menetelmiä tullaan luvussa 5 testaamaan kummastakin edellä esitetystä näkökulmasta katsoen käyttäen simulointimalleja. Tällöin tullaan käyttämään estimointimenetelmien epävarmuuden mittoina sekä kehitysvirheen $\boldsymbol{R}(t)$ ja korvauskulun $X(t)$ keskihajontoja σ_R ja σ_X. Tarkemmin ottaen lähinnä ollaan kiinnostuneita näiden tekijöihin ja riskimaksutulon suhteen keskihajonnasta ja kehitysvirheen tapauksessa myös kehitysvirheen ja todellisten vahinkomenojen suhteen keskihajonnasta. Näin ollen ollaan siis kiinnostuneita suureista

$$\sigma\left(\frac{R}{P}\right), \sigma\left(\frac{X}{P}\right) \text{ ja } \sigma\left(\frac{R}{X_{tod}}\right).$$

Edellä merkinnälle X_{tod} tarkoitetaan todellisia vahinkomenoja eli niiden vahinkomenojen todellisia kustannuksia, joita korvausvastuun on tarkoitus estimoida.

Joillekin korvausvastuun estimointimenetelmille nämä keskihajonnat voidaan laskea myös analyyttisesti ja saada näin ollen luottamusvälejä korvausvastuun estimaateille, mutta tällöin joudutaan usein tekemään korvausprosessiin vaikuttavista taustatekijöistä melko rajoittavia oletuksia. Yksi esimerkiksi tällaisista rajoittavista oletuksista on, että inflaatio on vakio koko tarkasteltavan aikavälin. Itsestään on kuitenkin selvää, että se voi vaihdella melko rajustikin. Tästä syystä tässä tutkielmassa onkin valittu lähestymistavaksi simulointimallit eri estimointimenetelmien eroavuuksien havaitsemiseksi. Pääsääntöinen syy tähän on, että tällöin voidaan tarvittavien oletuksien määrä supistaa melko vähäiseksi ja näin ollen voidaan korvausvastuuprosessia mallintaa tavalla, joka on lähempänä todellisuutta.

Seuraavassa on kerrattu viitteessä Pentikäinen & Rantala (1992) luetellut vakuutusyhtiön johdon kannalta tärkeimmät ominaisuudet hyvälle estimointimenetelmälle:

1. Todennäköisyys korvausvastuun estimaatin osoittautumisesta liian vähäiseksi pitäisi olla pieni eli $P[\boldsymbol{R}(t) + L(t) < 0] \le \varepsilon$. Edellä $L(t)$ on varmuusmarginaali.
2. Varmuusmarginaalin $L(t)$ pitäisi olla mahdollisimman pieni.
3. Korvauskulun heilahtelun pitäisi olla mahdollisimman pieni.

Tässä tutkielmassa on vertailussa keskitytty lähinnä ehtoon kolme eli eri menetelmien antamien estimaattien vaihteluihin.

3. Korvausvastuun estimointimenetelmiä

Vakuutusmatemaattisessa kirjallisuudessa on esitelty lähes lukematon määrä erilaisia korvausvastuun estimointimenetelmiä. Erityisesti 1970-luvulta lähtien asiaa on käsitelty varsin laajassa mittakaavassa alan kirjallisuudessa. Ensimmäiset esitetyt kollektiiviset menetelmät ovatkin peräisin juuri 60-luvun ja 70-luvun vaihteesta ja ovat rakenteeltaan melko yksinkertaisia. Niiden vaatima tilastoaineisto on yleisesti ottaen niin triviaalia, että se on käsillä jokaisessa vakuutusyhtiössä. Sittemmin menetelmien rakenne on huomattavasti monimutkaistunut ja ideat, joihin ne perustuvat, eivät enää välttämättä ole niin helposti ymmärrettävissä. Kun ensimmäisissä korvausvastuun estimointimenetelmissä estimaatti saatiin periaatteessa kertomalla vahinkovuonna kerätty vahinkomaksutulo jollakin sopivalla kertoimella, niin nykyaikaisimpien estimointimallien sovittamisessa käytetään esimerkiksi Kalman-suotimia. Hyvä puoli näissä nykyisissä menetelmissä on se, että niillä saadaan myös estimaatit korvausvastuun estimaatin keskihajonnalle ja näin ollen voidaan määrätä luottamusvälejä korvausvastuun estimaatille. Toisaalta näiden kehittyneempien estimointimenetelmien vaatima tilastoaineisto on jo useimmiten sen verran yksityiskohtaista, etteivät ne enää sovellu kaikkiin tilanteisiin. Tämä on usein ongelmana esimerkiksi jälleenvakuutuksen yhteydessä. Nykyään myös eri maiden hallituksien säätelemät määräykset sallivat näiden kehittyneiden korvausvastuun estimointimenetelmien käytön, kun esimerkiksi joskus 1970-luvulla oli yleisesti hyväksyttyjä ainoastaan käsin tapauskohtaisesti tehdyt korvausvastuun estimaatit. Olemassa olevasta kirjallisuudesta voi mainita viitteet Eeghen (1981) ja Taylor (1986), joissa on annettu kummassakin jonkinlainen läpileikkaus eri estimointimenetelmistä.

Tässä yhteydessä on valittu esiteltäviksi ja testattaviksi estimointimenetelmiksi seuraavat menetelmät:

1. Peräkkäisten askelmien (*Chain-Ladder*) menetelmä,
2. Bornhutterin Fergusonin menetelmä,
3. Mackin kredibiliteettimenetelmä,
4. Empiirinen Bayes-menetelmä.

Syy yllä olevien valintaan on lähinnä ollut se, että ne edustavat jokainen hieman toisistaan poikkeavaa tapaa estimoida korvausvastuuta ja antavat näin ollen jonkinlaisen yleiskuvan olemassa olevista menetelmistä. Seuraavissa pykälissä on tarkoitus antaa lyhyt kuvaus kustakin menetelmästä ja kiinnostuneita varten annetaan viitteitä perusteellisempiin lähdeteoksiin.

Lähes kaikissa seuraavissa pykälissä esitetyissä korvausvastuun estimointimenetelmissä käytetään hyväksi luvussa 2 esitettyä rakennemallia. Nämä estimointimenetelmät olettavat kehityskolmion kuitenkin muodostuneen hieman eri satunnaismuuttujien arvoista. Kunkin menetelmän yhteydessä tullaan erikseen selventämään, mitä tietoa tarvittavan kehityskolmion pitää sisältää. Yleisesti ottaen nämä menetelmät kuitenkin olettavat, että kyseessä on kuvan 3.1 mukainen kehityskolmio, missä satunnaismuuttujat X_{ij} kuvaavat jotain informaatiota vahinkomenoprosessista vahinkovuoden i ja kehittymisvuoden j suhteen. Osa menetelmistä suoriutuu myös tilanteesta, jossa osa soluista puuttuu jostakin syystä, mutta tarkempi kuvaus menetelmien vaatimasta tilastoaineistosta ja sen muodosta tullaan antamaan kuvattaessa kyseistä menetelmää. Jatkossa kuitenkin puhuttaessa kehityskolmiosta oletetaan ellei erikseen muuta määritellä, että käsillä tilastoaineisto todellakin muodostaa kolmion eli kuvassa 3.1 $t = s$.

Vahinkovuosi i	Kehittymisvuosi j						
	1	2	3	.	.	.	s
1	X_{11}	X_{12}	X_{13}	.	.	.	X_{1s}
2	X_{21}	X_{22}				.	
3	X_{31}				.		
.	.						
.	.						
.	.	.					
t	X_{t1}						

Kuva 3.1. Kehityskolmio hieman toisenlaisessa muodossa ilmaistuna. Havaintoarvot X_{ij} ilmaisevat jotain vahinkovuoteen i ja kehittymisvuoteen j liittyvää informaatiota, jonka avulla korvausvastuun estimaatti on tarkoitus johtaa.

Edeltä kannattaa huomioida, että määrättäessä vuoden t lopussa korvausvastuuta, niin kyseisen vuoden t aikana maksettujen vahinkojen määrät oletetaan jo olevan tiedossa. Käytännössä tämä ei kuitenkaan ole mahdollista, mutta kyseinen ongelma on enemmänkin filosofinen ja sen kiertäminen tuskin tuottaa suuria vaikeuksia.

3.1. Peräkkäisten askelmien menetelmä (Chain-Ladder)

Tämän menetelmän alkuperä ei ole aivan selvä, mutta se on kuitenkin kehitetty joskus 1970-luvun alkupuoliskolla tai mahdollisesti jopa jo 1960-luvun lopulla ja myöskään sen alkuperäisestä kehittäjästä ei ole tarkkaa tietoa. Mahdollista onkin, että se on kehitetty lähes samanaikaisesti useammalla eri taholla johtuen juuri sen yksinkertaisesta ideasta. Tämä menetelmä on kuitenkin ollut yksi ensimmäisistä kollektiivisista menetelmistä, joita käytettiin korvausvastuun estimointiin. Sitä voidaankin pitää eräänlaisena korvausvastuun estimointimenetelmien esi-isänä ja monet jälkeenpäin kehitetyt menetelmät ovatkin omalla tavallaan sen kehittyneempiä versioita, kuten tullaan myös tässä yhteydessä jatkossa havaitsemaan.

Peräkkäisten askelmien menetelmä ja kaikki sen suorat variaatiot perustuvat periaatteessa oletukseen, että kaikkien kohorttien kehittyminen on samankaltaista vuodesta toiseen eli että vahinkojen selviämisjakautuma ei riipu vahinkovuodesta. Tämä oletus on itse asiassa lähes tulkoon myös kaikkien muidenkin eri korvausvastuun estimointimenetelmien taustalla. Varsinainen menetelmän yksilöivä oletus on nyt, että kohorttien kehittymisvuosittaiset kumulatiiviset vahinkomenot $X(t;0,s)$ ovat keskenään verrannollisia tiettyjen kasvukertoimien suhteen. Tällä tarkoitetaan sitä, että

$$E\left[\frac{X(t;0,s+1)}{X(t;0,s)}\right]=E\left[\frac{X(u;0,s+1)}{X(u;0,s)}\right]=a(s)$$

kaikille parametrien s, t ja u arvoille, joille yllä olevat lausekkeet on määritelty. Tarkoituksena on siis jo tiedossa olevasta tilastoaineistosta laskea estimaatit näille kehittymisvuosien välisille kasvukertoimille ja käyttää niitä sitten hyväksi estimoitaessa korvausvastuuta.

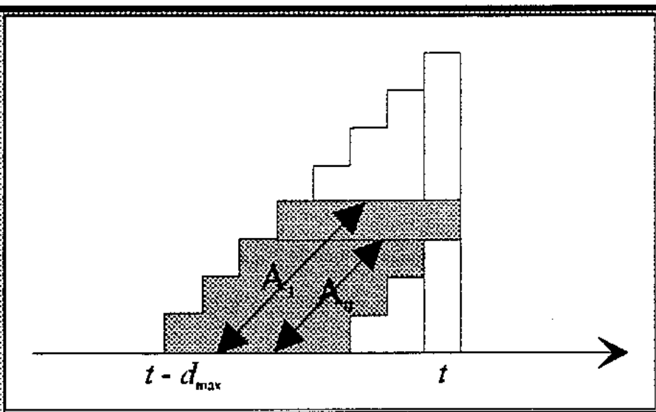
Nyt on siis käsiteltävänä kehityskolmio (ks. kuva 3.1), jonka muodostavat kohorttien hetkellä t tiedossa olevat kumulatiiviset korvaussummat

$$X_{ij}=X(i;0,j-1),\ i+j<t+1,$$

ja tarkoituksena on estimoida termejä X_{ij}, joille $i+j\geq t+1$. Kohorttien kasvukertoimiksi eli kohorttien välisiksi kasvusuhteiksi kehittymisvuosittain saadaan tiedossa olevan kehityskolmion avulla

$$a(d)=\frac{A_1(d)}{A_0(d)}=\frac{\sum_{i=d}^{d_{max}}X(t-i;0,d)}{\sum_{i=d}^{d_{max}}X(t-i;0,d-1)}. \tag{3.1}$$

Yllä olevan kaavan summauslausekkeita on selvennetty kuvassa 3.2.



Kuva 3.2. Kasvukertoimien $a(s)$ laskemisessa käytettyjen summalausekkeiden graafinen selvennys.

Varsinainen korvausvastuun estimaatti hetkellä t saadaan estimoimalla ensin jokaiseen kohorttiin liittyvää korvausvastuuta ja sitten summaamalla nämä eri kohorttien estimaatit yhteen. Nyt kohortin $t-s$ korvausvastuun estimaatiksi saadaan

$$
\begin{aligned}
C(t-s,s) &= \left[\prod_{\tau=s}^{d_{max}-1} a(\tau)\right]\cdot X(t-s;0,s) - X(t-s;0,s) \\
&= \left[\prod_{\tau=s}^{d_{max}-1} a(\tau) - 1\right]\cdot X(t-s;0,s).
\end{aligned}
\tag{3.2}
$$

Yllä oleva kaava on saatu soveltamalla iteratiivisesti kehittymisindeksejä $a(s)$ kehittymisvuosien loppuun asti eli

$$
\begin{aligned}
X(t-s;0,s+1) &= X(t-s;0,s)\cdot a(s), \\
X(t-s;0,s+2) &= X(t-s;0,s+1)\cdot a(s+1) \\
&= X(t-s;0,s)\cdot a(s)\cdot a(s+1).
\end{aligned}
$$

Näin ollen vuoden t korvausvastuun estimaatiksi saadaan

$$
C(t) = \sum_{s=0}^{d_{max}-1} C(t-s,s). \tag{3.3}
$$

Estimoitaessa korvausvastuuta vuosittain kannattaa huomioida, että kasvukertoimet $a(s)$ pitää laskea joka vuosi uudelleen.

Mikäli korvausvastuu halutaan diskontata, muuttuu kaava (3.2) hieman. Nyt korvausvastuu täytyy erikseen ilmaista kunkin kohortin jokaiselle tuntemattomalle solulle vastaavalla tavalla kuin yllä. Näin ollen

$$
\begin{aligned}
X(t-s;s+1,s+1) &= (a(s)-1)\cdot X(t-s;0,s)\cdot v(t,1) \\
X(t-s;s+2,s+2) &= (a(s+1)-1)\cdot a(s)\cdot X(t-s;0,s)\cdot v(t,2)
\end{aligned}
$$

ja vastaavasti jatkamalla saadaan

(3.4) $$C(t-s,s)=X(t-s;0,s)\cdot\sum_{i=1}^{d_{max}-s}\left[\prod_{k=0}^{i-2}a(s+k)\right]\cdot\left(a(s+i-1)-1\right)\cdot v(t,i).$$

Varsinainen korvausvastuun estimaatti saadaan sitten summaamalla nämä eri kohorttien antamat estimaatit yhteen kuten edelläkin. Yllä termi $v(t, s)$ on diskonttaustekijä (ks. kaava (2.8)).

Yllä esitetyllä niin kutsutulla perinteisellä peräkkäisten askelmien menetelmällä on muutamia heikkouksia. Sen perusoletushan on, että selviämisjakautuma pysyy ajan kuluessa suhteellisen stabiilina. Näin ollen sen ja oikeastaan minkään muunkaan estimointimenetelmä käyttö ei ole perusteltua, jos jokin sisäinen tai ulkoinen tekijä aiheuttaa muutoksen korvausten kehityskäyttäytymiseen. Tällainen on esimerkiksi

- muutos vakuutusyhtiön hallinnollisessa tavassa käsitellä korvauksia,
- muutos oikeuskäytännössä,
- poikkeuksellisen suuri inflaatio.

Peräkkäisten askelmien menetelmän suurin ongelma onkin juuri yllä lueteltu inflaatio. Se aiheuttaa ongelmia lähinnä kahdesta syystä. Menetelmä ottaa inflaation implisiittisesti huomioon, joten sen suuruusluokka ei ole tiedossa. Lisäksi menetelmä olettaa tulevaisuuden inflaation olevan vakiokorkoista, jonka suuruus puolestaan oletetaan olevan menneisyyden inflaatioindeksien keskiarvo. Tämä on selvästikin melko rajoittunut oletus. Tämä voidaan kuitenkin välttää sovittamalla tilastoaineisto siten, että kaikki eri havainnot koskevat samaa rahanarvoa. Tämä on mahdollista tehdä, mikäli inflaatio on tiedossa jostakin muusta lähteestä kuin kehityskolmiosta.

Tämä menetelmä vaatii myös, että vähintään yksi kohortti on kehittynyt loppuun asti eli ettei ainakaan yhdeltä vahinkovuodelta ole enää vahingonkorvauksia maksamatta. Tämä vaatimus ei ole kuitenkaan mitenkään liian rajoittava, sillä jos on odotettavissa, että varhaisimmalta vahinkovuodelta tulee vielä korvauskanteita, niin tällöin voidaan selviämisjakautumaa ekstrapoloida jollakin sopivalla tavalla. Käytännössä jonkinlaiseen ekstrapolointiin joudutaankin turvautumaan melko usein, koska estimoitavat vakuutuslajit ovat yleisesti ottaen etupainoitteisia ja suhteellisen pitkähäntäisiä.

Vielä yksi peräkkäisten askelmien menetelmälle ominainen ongelma on sen herkkyys kehityskolmion koillisnurkalle. Tämä johtuu siitä, että näille viimeisille kehittymisvuosille on kertynyt melko vähän havaintoarvoja. Toisaalta kyseinen ilmiö on melko yleinen eri estimointimenetelmien joukossa, mutta erityisesti peräkkäisten askelmien menetelmälle se korostuu johtuen juuri menetelmän kasvukertoimien multiplikatiivisuusoletuksesta. Simulointimallissa tämän vaikutusta testattiin ottamalla testattaviin menetelmiin mukaan tämän menetelmän pieni modifikaatio, jossa kehityskolmioon otettiin mukaan kolme valmiiksi kehittynyttä kohorttia muiden kohorttien lisäksi. Ongelma ei ole välttämättä liian suuri riippuen selviämisjakauman muodosta: mikäli suurin osa korvauksista selviää muutaman ensimmäisen vuoden aikana, niin tällöin selviämisjakauman hännän paino on melko pieni (kertoimet $a(s)$ ovat lähellä 1, kun s on suuri).

Vastaavanlaisen ongelman aiheuttaa myös lounaisnurkkaus. Mikäli parin ensimmäisen selviämisvuoden korvausmenot vaihtelevat suuresti vuosittain, multiplikatiivisuus-ominaisuudesta seuraa, että myös uusimpien vahinkovuosien lopullisten korvausmenojen estimaatit heilahtelevat vastaavassa suhteessa. Äärimmäisenä tilanteena kyseiselle menetelmälle voidaan pitää tilannetta, jossa uusimman vahinkovuoden ensimmäisenä vuonna ei ole lainkaan maksettu korvauksia, mistä seuraa että kyseisen vahinkovuoden korvausvastuun estimaatti on nolla.

Toisaalta menetelmän hyviin puoliin voidaan lukea se, että menetelmän taustalla oleva idea on yksinkertainen ja näin ollen aktuaarin on helppo muokata menetelmää, jos jokin taustatekijä rikkoo menetelmän oletuksia. Myös vakuutusyhtiön markkinavolyymin muuttuminen ei aiheuta ongelmia.

Peräkkäisten askelmien menetelmälle on olemassa lukuisia erilaisia variaatioita. Näistä voi katsoa tarkemman kuvauksen esimerkiksi viitteestä Eeghen (1981).

3.2. Bornhutterin Fergusonin menetelmä

Tämä menetelmä on oikeastaan enemmänkin jonkinlainen periaate kuin varsinainen menetelmä ja on alkuperäisin 1970-luvun alkupuolelta. Van Eeghen (ks. Eeghen (1981) s. 25) luettelee kyseisen menetelmän kuuluvan niin kutsuttuihin yksinkertaisiin korvausvastuun estimointimenetelmiin, joille on ominaista

- niiden yksinkertaisuus. Ne ovat enemmänkin kirjanpidollisia toimenpiteitä kuin matemaattisia algoritmeja,
- yleinen aktuaarin oman harkinnan hyväksikäyttö,
- käytettävän tilastoaineiston yleisyys.

Varsinainen idea on estimoida korvausvastuuta jokaiselle vahinkovuodelle erikseen käyttämällä hyväksi kyseisenä vahinkovuotena kerätyn riskimaksutulon määrää ja korvausten keskimääräisiä selviämisjakaumia. Vuonna $t - s$ sattuneiden vahinkojen korvausvastuun estimaatti vuonna t määritellään olevan

(3.5*a*) $$C(t-s,s) = P(t-s)\cdot c(t,s).$$

Yllä $P(t-s)$ on vuoden $t-s$ **riskimaksutulo** (*unloaded net premium*) ja parametri $c(t, s)$ on vahinkovuodesta t ja kehittymisajasta s riippuva estimaatti vielä maksamattomina olevien korvausten ja vahinkovuoden t kokonaisvahinkomenon suhteelle. Toinen vaihtoehtoinen tapa ilmaista korvausvastuun estimaatti on

(3.5*b*) $$C(t-s,s) = E\left[X(t-s;0,d_{\max})\right]\cdot c(t,s).$$

Yllä olevilla tavoilla ei ole suurta eroa, sillä tietyn vuoden riskimaksutulon pitäisi olla samaa kertaluokkaa kuin kyseisen vuoden arvioitu kokonaisvahinkomeno, mutta kirjallisuudessa kaavasta (3.5*b*) käytetään yleisesti nimitystä Bornhutterin Fergusonin menetelmä ja kaava (3.5*a*) tunnetaan vastaavasti **riskimaksutuloon perustuvana menetelmänä** (*premium-based method*).

Nyt kannattaa erityisesti huomioida, ettei eri vahinkovuosilta kerättyjä tilastoaineistoja suoranaisesti käytetä millään tavalla hyväksi. Ainoastaan kiinnitetään huomiota tietyn vahinkovuoden aikana kerättyyn riskimaksutuloon tai estimoituun kokonaisvahinkomenoon, ja aikaan, joka on kulunut kyseisestä vahinkovuodesta. Tilastoaineisto voi kuitenkin välillisesti vaikuttaa estimointiin riippuen parametrien $c(t, s)$ määrittelystä.

Tässä yhteydessä on tarkoitus suorittaa simulointitestejä kaavaa 3.5*a* vastaavalle menetelmälle. Tällöin täytyy jollakin tavalla määritellä vuoden $t - s$ riskimaksutulo $P(t - s)$. Kyseisen suure on nyt määritelty lähes vastaavasti kuin viitteessä Pentikäinen & Rantala (1992) eli

(3.6)
$$P(t-s) = \frac{\sum_{u=0}^{T} X_p^*(t-s-u)}{T},$$

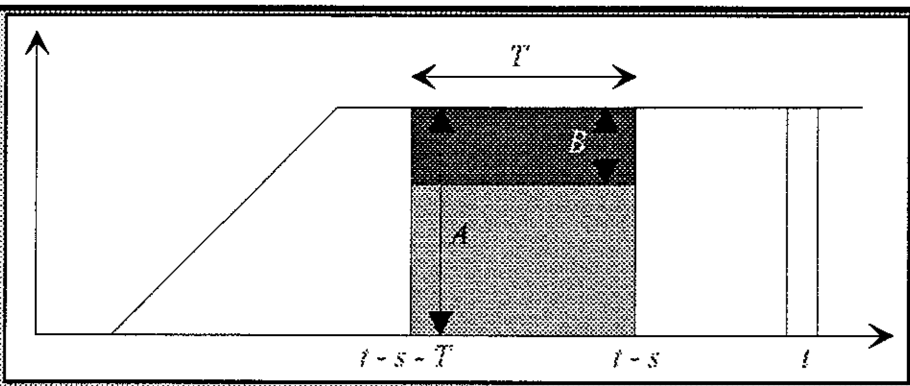
missä termit $X_p^*(t-s-u)$ tarkoittavat vuonna $t - s - u$ *maksettujen* korvausten kokonaismääriä, jotka on normeerattu vuoden t tasolle eli muutettu vastaamaan vuoden t rahanarvoa. Parametri T on summaukseen otettavien vuosien lukumäärä.

Vastaavasti myös kertoimet $c(t, s)$ määritellään tässä yhteydessä simuloidun vahinkohistorian avulla eli

(3.7)
$$c(t,s) = \frac{\sum_B X^*}{\sum_A X^*} = \frac{\sum_{u=0}^{T-1} \sum_{v=s}^{d_{max}} X^*(t-s-u-v;v,v)}{\sum_{u=0}^{T-1} X_p^*(t-s-u)}.$$

Käytännössä kaavoja (3.6) ja (3.7) vastaavia menetelmiä käytettäessä normeerattujen vahinkomenojen määräämisestä olla tarkkana vallitsevan inflaation takia ja erityisesti vahinkovuoden sisäisen korvausinflaation takia. Vahinkovuoden sisäisellä korvausinflaatiolla tarkoitetaan siis sitä, kuinka maksettavien korvausten määrät muuttuvat selviämisvuoden suhteen (sillä ei siis tarkoiteta sitä, että mahdollisesti yksittäisten korvausten jakaumat riippuisivat selviämisvuodesta). Lisäksi haluttaessa todellisia korvausvastuun estimaatteja pitää saatuihin normeerattuihin korvausmenon estimaatteihin lisätä tulevaisuuden inflaatiokehitys jollakin tavalla. Se voidaan yksinkertaisimmillaan ottaa huomioon joko määrittelemällä se menneisyyden inflaatiokehityksen mukaan tai sitten olettamalla se vakioksi. Tässä tutkielmassa simuloinnin yhteydessä on päädytty edelliseen.

Toinen vaihtoehto kyseisten parametrien määrittelyksi olisi olettaa niiden olevan kiinnitetty etukäteen. Kuvassa 3.3 on graafisesti esitetty, miten kaavojen 3.2 ja 3.2 summaukset lasketaan. Lisäksi nyt kannattaa huomioida, että mikäli selviämisjakaumaa vastaavia parametreja $c(t, s)$ ei kiinnitetä etukäteen eli että ne estimoidaan vuosittain uudelleen, niin voi olla että näiden parametrien summa on erisuuri kuin 1 eli ne eivät edustakaan varsinaista perinteistä todennäköisyysjakautumaa.



Kuva 3.3. Yllä on graafisesti selvennetty eri vuosien riskimaksutulon ja parametrien $c(t, s)$ laskemista. Vuoden $t-s$ riskimaksutulo saadaan jakamalla alueen A sisältämät normeeratut korvausmenot välin pituudella T. Vastaavasti parametrit $c(t, s)$ on saatu jakamalla alueen B normeeratut korvaukset alueen A normeeratuilla korvauksilla.

Mikäli nyt halutaan ottaa diskonttaus huomioon, niin asiat muuttuvat jälleen hieman. Tässä yhteydessä menetelmän laskemaa riskimaksutuloa ei ole määritelty diskontattuna, vaan diskonttaus on huomioitu vasta laskettaessa eri kohorttien korvausvastuun estimaatteja. Tavallisessa tapauksessa estimaatti saatiin kertomalla riskimaksutulo kertoimella $c(t, s)$. Nyt on laskettu erikseen kunkin kohortin eri solujen estimaatit ja diskontattu nämä sopivasti eli

$$C(t,s) = P(t-s) \cdot \sum_{i=s}^{d_{max}-1} \big(c(t,i) - c(t,i+1)\big) \cdot v(t,i-s+1).$$

Yleisesti kuitenkin riskimaksutulo edustaa jonkinlaista vahinkovuoden korvausmenon pääoma-arvoa jollakin tavalla laskettuna. Tällöin luonnollisesti pitää selviämisparametrien määrittelyssä ottaa pääoma-arvon laskemismenettely huomioon.

Riskimaksutuloon perustuvan menetelmän hyviin puoliin voidaan katsoa omalta osaltaan sen yksinkertaisuus. Sen vaatima tilastoaineisto on niin yleisluonteinen, että menetelmää voidaan soveltaa lähes aina. Erityisesti jälleenvakuutuksen yhteydessä tämä menetelmä tai sen modifikaatiot ovat erittäin suosittuja, koska jälleenvakuuttajan tilastoaineisto on yleensä melko rajoittunutta. Esimerkiksi joissakin kotimaisen ja ulkomaisen yhtiön välisissä sopimuksessa mainitaan, että tuntemattomia korvauksia varten varataan 50% vuoden riskimaksutulosta. Menetelmä sopii myös tilanteeseen, jossa uuden vakuutuslajin tilastokertymä on niin vähäinen, ettei muita menetelmiä voi edes ajatella käytettäviksi.

Tämän menetelmän ongelma onkin oikeastaan näennäinen yksinkertaisuus eli menetelmä luonnollisesti edellyttää, että kerätty riskimaksutulo on asianmukaisesti määrätty.

3.3. Mackin kredibiliteettimenetelmä

Tämä menetelmä on jo hieman uudempaa vuosikertaa kuin aikaisemmat ja lukeutuu korvausvastuun estimointimenetelmien joukossa niin kutsuttuihin kredibiliteettiteoriaan perustuviin menetelmiin. Menetelmän on alunperin kehittänyt De Vylder (ks. De Vylder (1982)) ja tässä yhteydessä esitetyn parannusehdotuksen kyseiselle mallille on julkaissut Mack (ks. Mack (1990)). Tämä menetelmä on jo sen verran edellä esitettyjä menetelmiä kehittyneempi, että nyt korvausvastuun estimaatille voidaan antaa myös jonkinlaiset luottamusvälit, koska sille on analyyttisesti laskettavissa keskihajonnat.

Oletetaan nyt, että kehityskolmion muodostavat normeeratut korvausmenot X_{ij} eli kyseinen kolmio sisältää luvut $X(i; j, j)$, jotka ovat kaikki muutettu vastaamaan samaa rahanarvoa ja jotka ovat vielä lisäksi jaettu vakuutusyhtiön markkinavolyymia kuvaavilla suhdeluvuilla p_i. Korvausmenojen muuttaminen vastaamaan samaa rahanarvoa tarkoittaa ensisijaisesti inflaation huomioimista ja vaatii näin ollen jonkin sopivan vuosittaisen indeksikertoimen. Toisaalta korvausmenojen jakaminen jollakin volyymia kuvaavalla tekijällä vastaa eri vuosien havaintoarvojen tekemistä suoraan vertailukelpoisiksi. Tällainen volyymikerroin voi olla esimerkiksi johdettu yhtiön vahinkomaksutulosta tai ensimmäisen vuoden aikana ilmoitettujen vahinkojen lukumäärästä. Toisaalta kyseistä volyymikerrointa määrättäessä pitää tässä yhteydessä ottaa huomioon myös mahdollinen vahinkosuhteen muutos, mikäli käytetään suoraan vuotuisia kerättyjä vakuutusmaksutuloja volyymikertoimina.

Menetelmän varsinainen idea on oletus, että normeeratut korvausmenot ovat vahinkovuosittain riippuvaisia tietyn struktuurimuuttujan Θ_i. tuntemattomasta realisaatiosta θ_i. Kyseisen riippuvuuden muodostavat oletukset, että

1. vahinkovuodet eli parit $(\Theta_1,(X_{11},\dots,X_{1d})),\dots,(\Theta_t,(X_{t1},\dots,X_{td}))$ ovat toisistaan riippumattomia,
2. struktuurimuuttujat $\Theta_1,\dots,\Theta_t$ ovat samoin jakautuneita,
3. $\mathrm{Cov}(X_{js},X_{ju}|\Theta_j)=0$ kaikille $u \neq s$,
4. on olemassa positiiviset luvut ja mitallinen funktio β siten, että
$$E[X_{js}|\Theta_j]=y_s\beta(\Theta_j),$$
5. on olemassa mitallinen funktio σ^2 siten, että

$$\operatorname{Var}\left(X_{js} \middle| \Theta_j\right) = \frac{y_s \sigma^2(\Theta_j)}{p_j}.$$

Nyt muistutukseksi ja selvennykseksi lukijalle johtuen eri menetelmien erilaisista merkinnöistä, niin nyt on tarkoitus estimoida tekijöitä $\mathrm{E}\left[X_{ij} \middle| \Theta_j\right]$ niille satunnaismuuttujille X_{ij}, jotka eivät ole vielä tiedossa. Näiden saamiseksi täytyy johtaa tilastoaineistosta eli kehityskolmiosta estimaatit parametrivektorille $\hat{y} = (y_1, y_2, \ldots, y_d)$ ja struktuurifunktioista $\beta(\Theta_i)$ ja $\sigma^2(\Theta_i)$ vaadittaville suureille.

Oletuksista 1–5 seuraa, että korvausvastuun estimaatti voidaan laskea käyttäen hyväksi kredibiliteettiteorian tunnettua Bühlmannin Straubin mallia. Kyseinen malli voidaan yksinkertaisimmillaan ilmaista kaavalla (kredibiliteettimaksukaava)

$$\tilde{X}_n = \frac{n}{n+\kappa}\overline{X}_n + \frac{\kappa}{n+\kappa}\mu$$

missä $\kappa = \dfrac{E[Var[X|\Theta]]}{Var[E[X|\Theta]]}$. Luonnollisesti tässä yhteydessä kyseistä mallia joudutaan hieman muokkaamaan ottamalla huomioon korvausten kehittymismalli eli run-off-kolmio. Käytännössä tämä kredibiliteettimalli tarkoittaa tässä yhteydessä, että korvausvastuun soluttaiset estimaatit saadaan sopivasti painottamalla vahinkovuoden sisäistä kehittymistä sekä kaikkien eri vahinkovuosien keskimääräistä kehittymistä.

Kyseiseksi estimaatiksi saadaan erikseen jokaiselle tuntemattomalle solulle

(3.8) $$\hat{X}_{ij} = y_j b \cdot \left(z_i \cdot Z_i + (1 - z_i)\right).$$

Yllä

$$Z_i = \frac{\sum_{j \in T_i} X_{ij}}{\sum_{j \in T_i} y_j b}, \quad z_i = \frac{p_i \cdot \sum_{j \in T_i} y_j b}{p_i \cdot \sum_{j \in T_i} y_j b + \frac{c}{a}}, \quad T_i = \left\{ j \middle| X_{ij} \text{ on tunnettu} \right\}$$

$$b = \mathrm{E}\left[\beta(\Theta_j)\right], \quad c = \frac{1}{b}\mathrm{E}\left[\sigma^2(\Theta_j)\right] \text{ ja } a = \frac{1}{b^2}\operatorname{Var}\left[\beta(\Theta_j)\right].$$

Nyt termi $y_j b$ vastaa siis periaatteessa kredibiliteettimaksun kaavassa olevaa kollektiivista termiä μ ja termi $y_j b \cdot Z_i$ vahinkovuoden sisäisen informaation termiä $\overline{X}_n$.

Yllä esitetyt kehityskertoimet $y_1, y_2, \ldots, y_d$ ja struktuurifunktiot on mallia johdettaessa oletettu tunnetuksi, mutta kuten yllä jo osittain mainittiin, niin niistä tarvittava informaatio pitää johtaa tilastoaineistosta. Tarvittavien muuttujien estimaateiksi saadaan

(3.9*a*)
$$\hat{y}_j b = \frac{\sum_{i \in K_j} p_i X_{ij}}{\sum_{i \in K_j} p_i}, \quad K_j = \{i \mid X_{ij} \text{ on tunnettu}\},$$

$$\hat{c} = \frac{1}{m} \sum_{i=1}^{t} p_i \cdot \sum_{j \in I_i} y_j b \cdot \left(\frac{X_{ij}}{y_j b} - Z_i \right)^2$$

ja

(3.9*b*)
$$\hat{a} = \frac{1}{t} \cdot \sum_{i=1}^{t} z_i \cdot (Z_i - 1)^2 .$$

Edellä parametri m on käytetyn kehityskolmion havaintojen lukumäärä vähennettynä vahinkovuosien lukumäärällä t. Edeltä kannattaa huomioida parametrin a estimaatin $\hat{a}$ määrittely tai oikeastaan estimaatin $\hat{a}$ ja tekijän z_j välinen riippuvuus. Kyseisille suureille pitääkin laskea estimaatit iteratiivisesti lähtien esimerkiksi ehdosta $z_j = 1$. Toinen vaihtoehto parametrin a estimaatiksi on

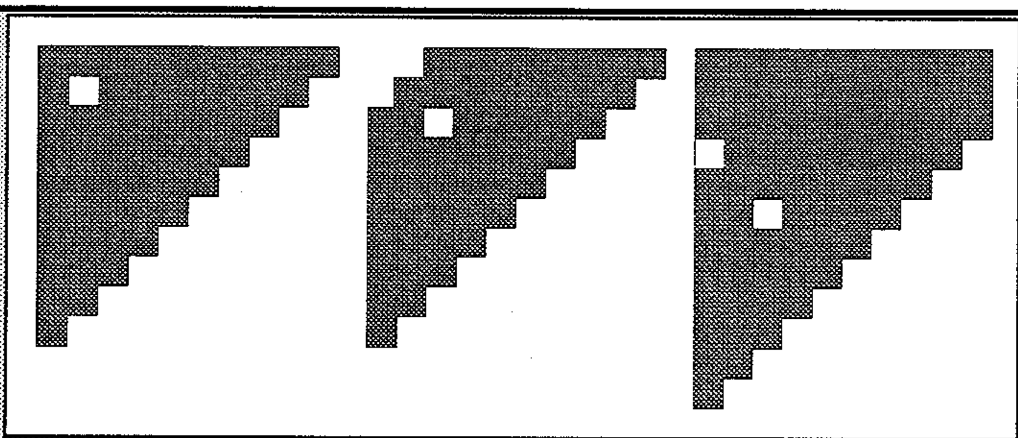
(3.9*c*)
$$\hat{a} = \frac{\sum_{i=1}^{t} p_i \cdot \sum_{j \in I_i} y_j b \cdot (Z_i - 1)^2 - t \cdot \hat{c}}{\sum_{i=1}^{t} p_i \cdot \sum_{j \in I_i} y_j b} .$$

Tämä onkin parempi vaihtoehto, koska tällöin parametrin a ja parametrien z_i määrittelyt eivät riipu toisistaan ja näin ollen niiden välillä ei tarvitse suorittaa iterointia, joka ainakin simulointimallien yhteydessä on suuri etu. Käytännössä laskettaessa korvausvastuun estimaatteja kannattaa olla myös tarkkana, ettei kaavan (3.9c) antama estimaatti ole negatiivinen. Mikäli näin käy, niin tällöin näissä tietyissä tilanteissa voidaan käyttää iteratiivista menetelmää.

Kuten edellä osittain jo mainittiinkin, varsinainen korvausvastuun estimaatti saadaan laskemalla kaavalla (3.8) solukohtaiset korvausvastuun estimaatit erikseen jokaiselle tuntemattomalle solulle. Nämä solukohtaiset estimaatit täytyy kuitenkin ensin muuttaa vastaamaan sopivaa rahanarvoa eli kertoa markkinavolyymia kuvaavilla luvuilla p_i ja sopivilla inflaatiokertoimilla. Lisäksi täytyy ottaa vielä huomioon tulevaisuuden inflaatio, mikäli korvaukset maksetaan maksamisvuoden rahana. Tämä vaatii jonkinlaisen inflaatioennusteen tekemistä tulevaisuuden varalle ja luonnollisesti tällöin syntyy jälleen virhettä. Kyseisen inflaatioennusteen tekeminen onkin menetelmän suurin ongelma, joka korostuu mitä kauemmin vahinkojen selviäminen kestää. Tulevaisuuden inflaation on tässä yhteydessä määritelty olevan vakiokorkoista ja kyseinen korko on laskettu kolmen edellisen vuoden painotettuna keskiarvona. Toisaalta tämä inflaatioennuste ei ole kovin suuri ongelma, jos menetelmää verrataan muihin menetelmiin, sillä esimerkiksi peräkkäisten askelmien menetelmä ottaa sen implisiittisesti huomioon ja näin ollen sen suuruudesta ei ole mitään tietoa. Nyt siis tulevaisuuden inflaation voidaan ottaa huomioon eksplisiittisesti ja näin ollen aktuaari voi käyttää omaa subjektiivista harkintaansa sen määräämisessä.

Menetelmän oletuksista voi mainita sen verran, että ainakin oletukset 1,2 ja 3 ovat soveliaita, mikäli vakuutusyhtiön korvauskäsittelyprosessi ei liiemmin tarkasteltavien vuosien aikana muutu. Vahinkomenoprosessin multiplikatiivisuusoletus eli oletus 4 on käytössä myös monissa muissa estimointimenetelmissä; esimerkiksi edellä esitetyssä peräkkäisten askelmien menetelmässä on vastaava oletus. Ideana kyseisessä oletuksessa on, että vahinkojen selviämisjakautuma on sama jokaisena vahinkovuotena, mutta varsinaisten korvausmenojen taso voi vaihdella.

Kuten monelle muullekin korvausvastuun estimointimenetelmälle, niin myös nyt kehityskolmion koillisnurkka aiheuttaa ongelmia. Nyt kyseinen ongelma ei ole niin suuri kuin mitä se oli peräkkäisten askelmien menetelmässä johtuen hieman erilaisesta tavasta estimoida korvausvastuuta. Tässä yhteydessä tilastoaineiston ei täydy kuitenkaan muodostaa mitään kolmiota, vaan havaintoja voi puuttua sieltä täältä (ks. kuva 3.3). Toisaalta voidaan käyttää hyväksi myös niitä vanhempia vahinkovuosia, jotka ovat jo kehittyneet loppuun asti ja tällöin koillisnurkan ongelma osittain poistuu. Tässä kuitenkin se ongelma taas, että jopa vuosikymmeniä vanhat tilastoaineistot tuskin enää toteuttavat mallin oletuksia. Luultavammin vakuutusyhtiön korvausprosessi, yhteiskunnan oikeuskäytännöt ja monet muut vahinkomenoprosessiin vaikuttavat tekijät ovatkin muuttuneet jo useampaan otteeseen, mikäli edes kyseistä vakuutuslajia on ollut olemassa niin kauan.



Kuva 3.4. Yllä on graafisesti kuvattu erilaisia kehityskolmion variaatioita, joihin yllä esitettyä menetelmää voidaan suoraan soveltaa.

Tarkemman kuvauksen edellä esitetystä menetelmästä ja yleisesti eri kredibiliteettiteoriaan liittyvistä menetelmistä voi katsoa viitteistä Mack (1990), De Vylder (1982), Hadidi (1985) ja Goovaerts *et al.* (1990).

3.4. Empiirinen Bayes-menetelmä

Seuraavaksi on tarkoitus esitellä viitteessä Verrall (1990) esitetty empiirinen Bayes-estimointimenetelmä. Kyseisessä julkaisussa esitetään myös tavallinen Bayes-estimointimenetelmä, joka eroaa tässä esitetystä siinä suhteessa, että se vaatii käyttäjältä priorijakauman tietyille parametreille. Nyt on kuitenkin valittu empiirinen menetelmä, koska jatkossa on tarkoitus tutkia simulointimalleilla eri menetelmien antamia tuloksia. Tämä lähinnä siksi, että empiirisessä menetelmässä kaikki tarvittavat priorijakautumat johdetaan tilastoaineistosta, eikä niitä näin ollen tarvitse erikseen määrätä. Käytännössä kuitenkin tavallinen Bayes-menetelmä on ehkä parempi valinta, koska tällöin aktuaari käyttää omaa subjektiivista näkemystään hyväksi määrittelemällä vaadittavat priorijakautumat sopivasti.

Nyt esitettävä menetelmä on omalla tavallaan melko suora modifikaatio peräkkäisten askelmien menetelmälle. Ainoa olennainen ero on oikeastaan vain teoria, jota käytetään korvausvastuun estimaattien johtamiseksi kehityskolmiosta. Menetelmän lähtökohtana on ensin muuttaa peräkkäisten algoritmien menetelmän pohjalla oleva multiplikatiivinen malli

(3.10) $$E[X(t;0,s)] = a(s)\cdot X(t;0,s-1)$$

lineaariseksi malliksi

(3.11) $$E[\ln(X(t;s,s))] = \mu + \alpha_t + \beta_s$$

olettaen, että lisäykset $X(t;s,s)$ ovat positiivisia ja log-normaalisti jakautuneita. Nyt malli (3.10) on selvästikin ekvivalentti mallin

$$E[X(t;s,s)] = a^*(s)\cdot X(t;0,d_{\max})$$

kanssa. Ottamalla tästä luonnolliset logaritmit kummaltakin puolelta saadaan

$$\begin{aligned}\ln(E[X(t;s,s)]) &= \ln(a^*(s)\cdot X(t;0,d_{\max}))\\ &= \ln(a^*(s)) + \ln(X(t;0,d_{\max}))\\ &= \ln(X(1;0,d_{\max})) + \ln(a^*(0))\\ &\quad + \ln(X(t;0,d_{\max})) - \ln(X(1;0,d_{\max}))\\ &\quad + \ln(a^*(s)) - \ln(a^*(0)),\end{aligned}$$

joka on juuri malli (3.11) merkitsemällä

(3.12) $$\begin{aligned}\mu &= \ln(X(1;0,d_{\max})) + \ln(a^*(0))\\ \alpha_t &= \ln(X(t;0,d_{\max})) - \ln(X(1;0,d_{\max}))\\ \beta_s &= \ln(a^*(s)) - \ln(a^*(0))\end{aligned}$$

Edeltä kannattaa huomioida, että rajoituksen $\sum_{s=0}^{d_{max}} a^*(s) = 1$ takia on yllä päädytty formulointiin, jossa uusille parametreille muodostuu rajoitus $\alpha_1 = \beta_1 = 0$. Useat kirjallisuudessa esitellyt korvausvastuun estimointimenetelmät perustuvatkin juuri tähän peräkkäisten askelmien menetelmän linearisointiin ja eroavatkin toisistaan lähinnä sen suhteen, mitä matemaattisen tilastotieteen teorian osa-aluetta on haluttu soveltaa eli miten parametrit α_i ja β_j on tarkoitus estimoida tilastoaineistosta.

Oletetaan nyt siis, että kohorttien solukohtaiset vahinkomenot $X(i; j, j)$ ovat positiivisia ja lognormaalisti jakautuneita. Tällöin on luonnollista ottaa niistä logaritmit ja merkitään

(3.13) $$Y_{ij} = \ln(X(i; j, j)).$$

Näin ollen varsinaiseksi lineaariseksi malliksi saadaan

(3.14) $$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij},$$

missä μ, α_i ja β_j ovat määritelty yhtälön (3.12) tavalla ja virhetermien ε_{ij} oletetaan olevan keskenään riippumattomia ja normaalisti jakautuneita parametrein 0 ja σ^2.

Kyseinen johdettu lineaarinen malli on lähes kaksisuuntainen ANOVA-malli ainoastaan sillä erolla, että osa arvoista on vielä tuntemattomia. Kuitenkin tälle lineaariselle mallille voidaan käyttää lineaarisen tilastollisten mallien teoriaa ja näin ollen tälle on löydettävissä hieman paremmin käyttäytyvät parametrien estimaatit kuin alkuperäiselle peräkkäisten askelmien menetelmälle. Lisäksi on mahdollista myös johtaa korvausvastuun estimaatille luottamusvälejä, koska sen keskihajonnat on laskettavissa tilastoaineistosta. Tarkemman kuvauksen peräkkäisten askelmien menetelmän multiplikatiivisen mallin johtamisesta lineaariseksi voi katsoa esimerkiksi viitteistä Kremer (1982) ja Verrall (1991a).

Nyt on tarkoitus käyttää hyväksi hierarkisten bayesilaisten lineaaristen mallien teoriaa ja tarkemmin ottaen niin kutsuttujen kaksivaihe (*two stage*) ja kolmivaihe (*three stage*) bayes-mallien kombinaatiota. Kaksivaihe bayes-malli on tässä yhteydessä tavallinen bayes-menetelmä, jossa siis parametreille määrätään jokin priorijakautuma. Kolmivaihemalli on puolestaan menetelmä, jossa parametrien priorijakautuman parametreille eli niin kutsutuille hyperparametreille määritellään myös priorijakautuma. Koska kaikki informaatio on tarkoitus johtaa tilastoaineistosta, niin hyperparametreille tullaan käyttämään epämääräistä priorijakautumaa. Tällä tarkoitetaan sitä, etteivät niiden varianssit ole äärellisinä olemassa.

Nyt olennainen oletus tässä menetelmässä on, että riviparametrit α_2, α_3, ..., α_t eli vahinkovuodesta riippuvat tekijät ovat riippumattomia ja samoin jakautuneita. Tämä tarkoittaa itseasiassa sitä, että käytettävän tilastoaineiston pitää sisältää normeerattuja korvausmenoja eli vakuutusyhtiön volyymilla ja inflaatiolla sovitettuja lukuarvoja. Varsinaisena ideana käyttää riviparametreille epämääräistä priorijakautumaa eli riviparametrit ovat jonkinlaisia hyperparametreja. Toisaalta myös muille parametreille eli sarakeparametreille β_2, β_3, ..., β_t ja keskiarvolle μ on tarkoitus käyttää epämääräistä priorijakautumaa, mutta tämä oletus tehdään mallia johdettaessa hieman eri vaiheessa.

Lineaarinen malli 3.14 on matriisimuodossa kirjoitettuna

(3.15) $$\hat{y} = \mathbf{X}\hat{\varphi} + \hat{e}$$

ja tässä tapauksessa hieman toisella tavalla ilmaistuna

$$\hat{y}|\hat{\varphi} \sim N(\mathrm{X}\hat{\varphi}, \mathrm{S}),$$

koska kehityskolmion muodostaneiden kokonaisvahinkomenojen $X(t; s, s)$ oletettiin olevan lognormaalisti jakautuneita. Edellä

- $\hat{y} = (y_{11}, y_{12}, y_{21}, y_{13}, y_{22}, y_{31}, \ldots, y_{t1}, y_{t-1,2}, \ldots, y_{1t})^{\mathrm{T}}$ on havaintovektori. Havaitut arvot ovat siis järjestetty vektoriin maksamisjärjestyksessä. Muistin virkistämiseksi kerrattakoon, että vuoden u maksetut korvaukset koostuvat kaikkien vahinkovuosien $u - v$ kehittymisvuonna $v = 0, 1, \ldots, d_{\max}$ maksetuista vahingoista $X(u - v; v, v)$,
- **X** on rakennematriisi (*design matrix*),
- **S** on virhetermien kovarianssimatriisi. Koska jatkossa oletetaan virhetermien olevan toisistaan riippumattomia, niin $\mathbf{S} = \sigma^2\mathbf{I}$,
- $\hat{\varphi} = (\mu, \alpha_2, \ldots, \alpha_t, \beta_2, \ldots, \beta_t)^{\mathrm{T}}$ on tuntematon parametrivektori,
- $\hat{\varepsilon} = (\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{21}, \ldots, \varepsilon_{t1}, \varepsilon_{t-1,2}, \ldots, \varepsilon_{1t})^{\mathrm{T}}$ virhevektori.

Seuraavassa esitellään lyhyesti lemma, jonka tulosta tarvitaan estimoitaessa satunnaismuuttujan $\hat{\varphi}|\hat{y}$ posteriorijakautumaa ja sen keskiarvoa.

Lemma 1. Oletetaan, että

$$\hat{y}|\hat{\varphi} \sim N(\mathrm{X}\hat{\varphi}, \mathrm{S}),$$
$$\hat{\varphi}|\hat{\theta}_1 \sim N(\mathrm{A}_1\hat{\theta}_1, \mathrm{C}_1),$$
$$\hat{\theta}_1|\hat{\theta}_2 \sim N(\mathrm{A}_2\hat{\theta}_2, \mathrm{C}_2),$$

missä $\hat{y}, \hat{\varphi}$ ja **X** ovat kuten yllä, $\hat{\theta}_1, \hat{\theta}_2$ ovat toisen vaiheen ja kolmannen vaiheen priorijakautumien parametrivektoreita ja $\mathbf{C}_1$ ja $\mathbf{C}_2$ ovat niiden kovarianssimatriisit. Oletetaan lisäksi, että parametrivektorin $\hat{\theta}_2$ priorijakautuma on epämääräinen eli C_2^{-1} on nollamatriisi. Tällöin posteriorijakauma parametrivektorille $\hat{\varphi}$ on

$$\hat{\varphi}|\hat{y}, \hat{\theta}_2 \sim N(\mathrm{D}\hat{d}, \mathrm{D}),$$

missä

$$D^{-1} = X^T S^{-1} X + C_1^{-1} - C_1^{-1} A_1 \left[A_1^T C_1^{-1} A_1 \right]^{-1} A_1^T C_1^{-1}$$

ja

$$\hat{d} = X^T S^{-1} \hat{y}.$$

Lisäksi parametrivektorin $\hat{\varphi}$ Bayes-estimaatti $\hat{\varphi}_B$ on posteriorijakautuman keskiarvo eli

$$\hat{\varphi}_B = D \cdot \hat{d} = D \cdot X^T S^{-1} X \hat{\varphi}_M.$$

Edellä $\hat{\varphi}_M$ on parametrivektorin $\hat{\varphi}$ suurimman uskottavuuden (*maximum-likelihood*) estimaatti, joka puolestaan saadaan yhtälön

$$X^T S^{-1} X \cdot \hat{\varphi}_M = X^T S^{-1} \hat{y}$$

ratkaisusta.

♦

Sovellettaessa yllä olevaa lemmaa tähän yhteyteen saadaan parametrivektorin $\hat{\varphi}$ Bayes-estimaatiksi muutaman välivaiheen jälkeen

(3.16)
$$\hat{\varphi}_B = \left[\sigma^{-2} X^T X + C^{-1} - B \right]^{-1} \sigma^{-2} X^T X \hat{\varphi}_M,$$

missä

$$B = \left[\begin{array}{c|ccc|c} 0 & 0 & \cdots & 0 & 0_{(1 \times t-1)} \\ \hline 0 & \frac{\sigma_\alpha^{-2}}{t-1} & \cdots & \frac{\sigma_\alpha^{-2}}{t-1} & 0_{(1 \times t-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \frac{\sigma_\alpha^{-2}}{t-1} & \cdots & \frac{\sigma_\alpha^{-2}}{t-1} & 0_{(1 \times t-1)} \\ \hline 0_{(t-1 \times 1)} & 0_{(t-1 \times 1)} & \cdots & 0_{(t-1 \times 1)} & 0_{(t-1 \times t-1)} \end{array} \right]$$

ja merkinnällä $0_{(n \times m)}$ tarkoitetaan $(n \times m)$ nollamatriisia. Yllä on oletettu, että varianssit σ^2 ja σ_α^2 ovat tunnettuja. Näin ei asian laita kuitenkaan ole, joten ne pitää vielä estimoida kehityskolmiosta. Näiden tarkat estimaatit ovat erittäin hankalia ellei lähes mahdottomia johtaa, joten tässä yhteydessä niille käytetään approksimoivia ratkaisuja. Nyt varianssien σ^2 ja σ_α^2 estimaateiksi s^2 ja s_α^2 saadaan

(3.17)
$$s^2 = \frac{\nu\lambda + \left(\hat{y} - X\hat{\varphi}_B \right)^T \left(\hat{y} - X\hat{\varphi}_B \right)}{n + \nu + 2}$$

ja

(3.18)
$$s_\alpha^2 = \frac{\nu_\alpha \lambda_\alpha + \sum_{i=1}^{t} \left(\alpha_{B,i} - \alpha_{mean}\right)^2}{n + \nu_\alpha + 1},$$

missä

$$\alpha_{mean} = \frac{\sum_{i=1}^{t} \left(\alpha_i\right)_B}{t-1}$$

ja merkinnällä $(\alpha_i)_B$ tarkoitetaan riviparametrien α_i bayes-estimaatteja. Jatkossa tullaan käyttämään ainoastaan täydellisiä kehityskolmioita ja näin ollen estimoinnissa käytettyjen vahinkovuosien lukumäärä $t = d_{max} + 1$. Lisäksi kannattaa jälleen kiinnittää huomiota siihen, että näiden varianssien estimaatit sisältävät varsinaisten parametrien estimaatit. Tästä johtuen menetelmän varsinaisen ratkaisun löytämiseksi täytyy yhtälöiden (3.16), (3.17) ja (3.18) välillä suorittaa iterointia kunnes haluttu konvergenssitaso on saavutettu. Yllä olevissa kaavoissa on vielä selvittämättä parametrien ν ja λ määrittely. Nämä parametri on tarkoitus määritellä varianssien priorijakautumien perusteella ja näiden oletetaan olevan käänteisen khinneliösuurejakautuman mukaan jakautuneita eli

$$\frac{\nu\lambda}{\sigma^2} \sim \chi_\nu^2 \text{ ja } \frac{\nu_\alpha \lambda_\alpha}{\sigma_\alpha^2} \sim \chi_\nu^2 .$$

Parametri v voidaan määritellä yhtäsuureksi kuin nolla ja näin ollen parametrin λ arvoa ei tarvitse määrätä. Parametrin ν_α arvo pitää kuitenkin olla nollasta eroava ja tässä yhteydessä sille on käytetty arvoa 0.8. Vastaavasti parametrin λ_α arvo on simulointimalleissa määritelty ensin laskettujen suurimman uskottavuuden estimaattien otosvarianssin avulla eli

$$\lambda_\alpha = \sum_{i=2}^{t} \left(\alpha_i - \alpha_{mean}\right)^2 .$$

Edellä α_{mean} on riviparametrien suurimman uskottavuuden estimaattien keskiarvo. Tarkemman kuvauksen edellä lyhyesti esitettyjen asioiden johtamisesta ja todistamisesta voi katsoa viitteistä Verrall (1990), Lindley (1971), Lindley (1965) ja Lindley & Smith (1971).

Edellä esitetyt tulokset täytyy vielä muuttaa koskemaan todellisia vahinkomenoja. Tähän voidaan käyttää seuraavaa lemmaa.

Lemma 2. Oletetaan, että solukohtaiset vahinkomenot $X(i; j, j)$ ovat lognormaalisti jakautuneita parametrein θ ja σ. Oletetaan lisäksi myös, että parametrin θ posteriorijakautuma on normaalijakautuma parametreilla m ja τ. Tällöin

$$E\left[X(i; j, j) | \Im\right] = \exp\left(m + \frac{1}{2}\sigma^2 + \frac{1}{2}\tau^2\right)$$

ja

$$\mathrm{Var}\left[X(i;j,j)\middle|\Im\right]=e^{\left(2m+\sigma^2+\tau^2\right)}\left(e^{\sigma^2+\tau^2}-1\right).$$

♦

Sovellettaessa yllä olevaa lemmaa tähän yhteyteen saadaan vielä tuntemattomille solukohtaisille normeeratuille korvausmenoille estimaatit

$$X_{ij}=X(i;j,j)=\exp\left(\mu^*+\alpha_i^*+\beta_j^*+\frac{1}{2}s^2+\frac{1}{2}s_{ij}^s\right). \tag{3.19}$$

Edellä

$$s_{ij}^2=\mathrm{X}_{(i,j)}\mathrm{D}^{-1}\mathrm{X}_{(i,j)}^{\mathrm{T}}$$

missä puolestaan merkinnällä $\mathbf{X}_{(i,\ j)}$ tarkoitetaan rakennematriisin sitä riviä, joka kuvaa parametreja α_i ja β_j eli riviä, jolla kaikki muut alkiot ovat nollia paitsi alkiot 1, i ja $t+j$ ovat ykkösiä.

Vastaavasti kuin Mackin menetelmässäkin täytyy yllä olevat normeeratut vahinkomenot X_{ij} muuttaa vastaamaan todellista rahanarvoa. Nyt kyseinen menettely on ihan samanlaista kuin pykälässä 3.3 ja myös luonnollisesti samat inflaatioennusteen ongelmat pätevät nytkin.

Testattaessa simuloimalla eri estimointimenetelmiä on yhdeksi menetelmäksi otettu tässä empiirisessä bayes-menetelmässä oikeastaan alkuarvoina käytetyt suurimman uskottavuuden estimaatit. Korvausvastuun estimaatin arvo käyttäen suurimman uskottavuuden estimointia saadaan samalla tavalla kuin ylläkin ainoastaan sillä erolla, että kaavassa (3.19) olevat varianssit lasketaan kaavoilla

$$\begin{aligned} s^2&=\frac{\left(\hat{y}-\mathrm{X}\hat{\varphi}_M\right)^{\mathrm{T}}\left(\hat{y}-\mathrm{X}\hat{\varphi}_M\right)}{n-p}, \\ s_{ij}^2&=\mathrm{X}_{(i,j)}\left(\mathrm{X}^{\mathrm{T}}\mathrm{X}\right)^{-1}\mathrm{X}_{(i,j)}\cdot s^2. \end{aligned} \tag{3.20}$$

Yllä n ja p ovat $(n\times p)$ rakennematriisin $\mathbf{X}$ dimensiot. Nyt kannattaa kuitenkin kiinnittää huomiota siihen, että tässä yhteydessä suurimman uskottavuuden menetelmällä saadut estimaatit eivät ole harhattomia johtuen käytetystä tavasta muuttaa saadun parametrivektorin tulokset koskemaan normeerattuja korvausmenoja. Viitteessä Verrall (1991b) on esitetty tapa, jolla suurimman uskottavuuden menetelmälle saadaan harhattomat estimaatit.

4. Eri menetelmien vertailuja

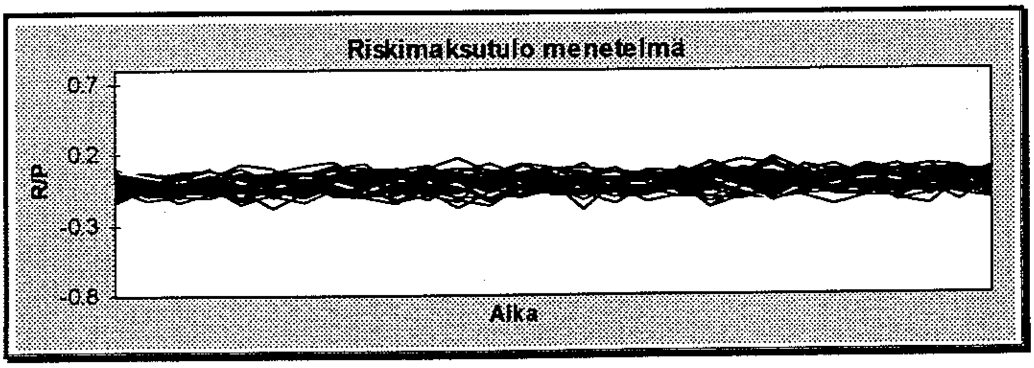
Seuraavissa pykälissä on esitelty saatuja tuloksia. Vertailuja on tehty tavallaan hierarkisesti aina lisäämällä uusia taustatekijöitä vahinkomenoprosessiin ja samalla tarkasteltu, kuinka ne vaikuttavat eri menetelmien antamiin tuloksiin. Taustatekijöihin on nyt laskettu inflaatio, vahinkojen lukumäärän vaihtelu ja diskonttaus. Vahinkojen lukumäärän vaihtelulla tarkoitetaan tässä yhteydessä trendin, pitkäjaksoisen heilahtelun ja lyhytjaksoisen heilahtelun huomioimista. Mikäli kyseistä tekijää ei ole huomioitu, niin vahinkojen lukumäärä on generoitu läpi koko realisaation Poisson–prosessina käyttäen samaa aloitusvuoden parametrin arvoa.

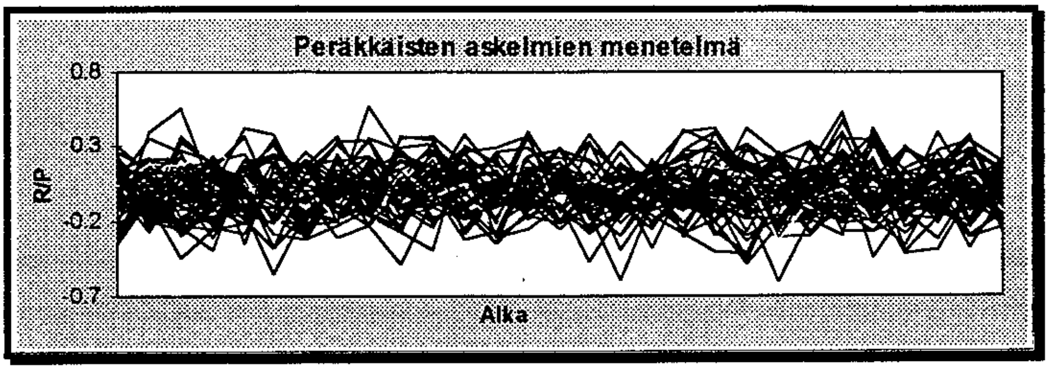
4.1. Ei taustatekijöitä

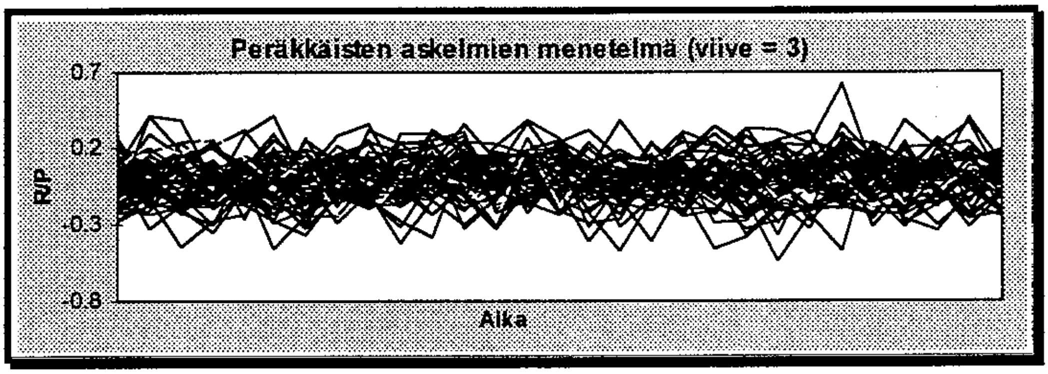
4.1.1. Ei diskonttausta

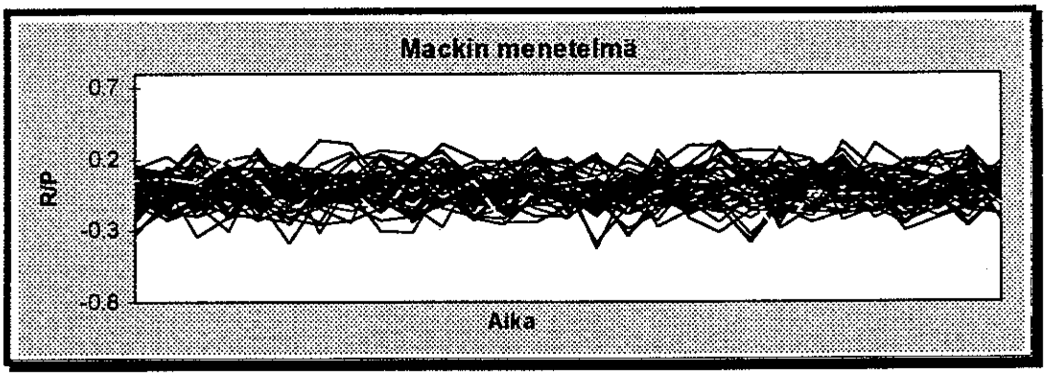
Ensiksi on tarkasteltu aktuaarin kannalta kaikkein ihanteellisinta tilannetta eli tilannetta, jossa inflaation ja vahinkojen vaihtelun vaikutusta ei ole otettu huomioon. Tämä tilanne ei luonnollisestikaan ole todellisuudessa realistinen, mutta tarkoitus onkin vain tutkia kuinka eri menetelmät toimivat tällöin.

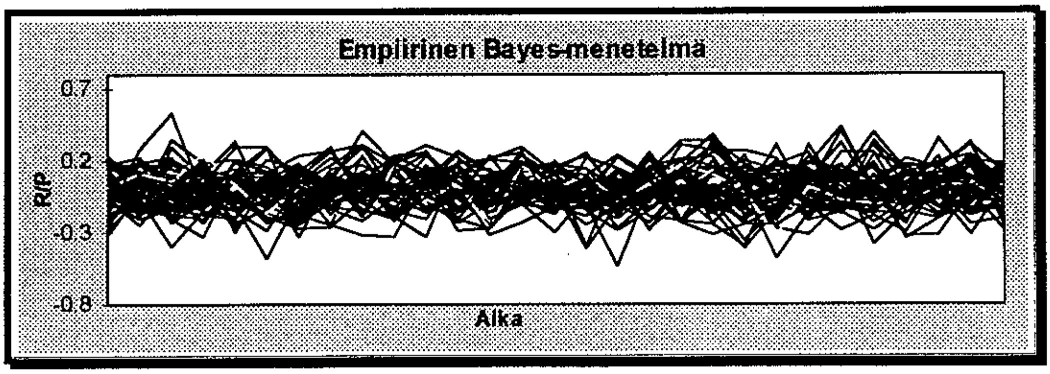
Seuraavassa on kuvattu graafisesti 50 ensimmäistä realisaatiota kunkin menetelmän antamasta kehitysvirheen ja riskimaksutulon suhteesta (ks. kuva 4.1). Tarkemmin ottaen muutenkin aina jatkossa annetut luvut eivät suoraan mittaa kehitysvirheen tapauksessa saatua kehitysvirhettä, vaan sen suhdetta vuosittaiseen riskimaksutuloon. Vastaavasti korvauskulun tapauksessa luvut kuvaavat korvauskulun ja riskimaksutulon suhdetta. Alla olevat realisaatiot on lähinnä annettu, jotta lukija saa jonkinlaisen kuvan menetelmien käyttäytymisestä ajan kuluessa.

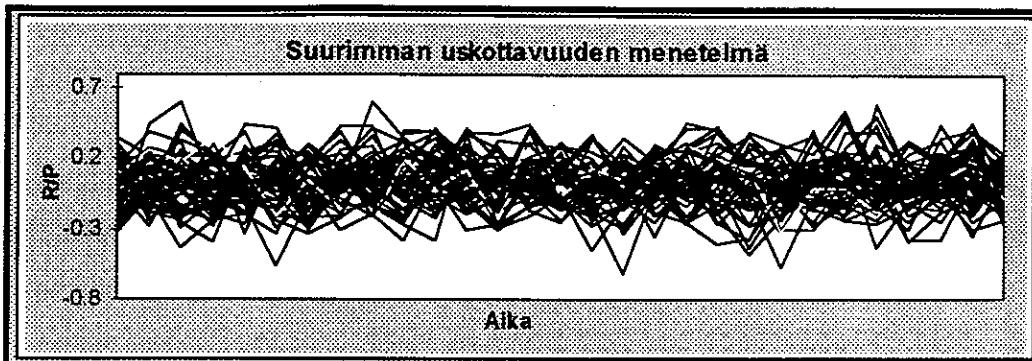



Peräkkäisten askelmien menetelmä
0.8
0.3
-0.2
-0.7
R/P
Aika


Peräkkäisten askelmien menetelmä (viive = 3)
0.7
0.2
-0.3
-0.8
R/P
Aika


Mackin menetelmä
0.7
0.2
-0.3
-0.8
R/P
Aika

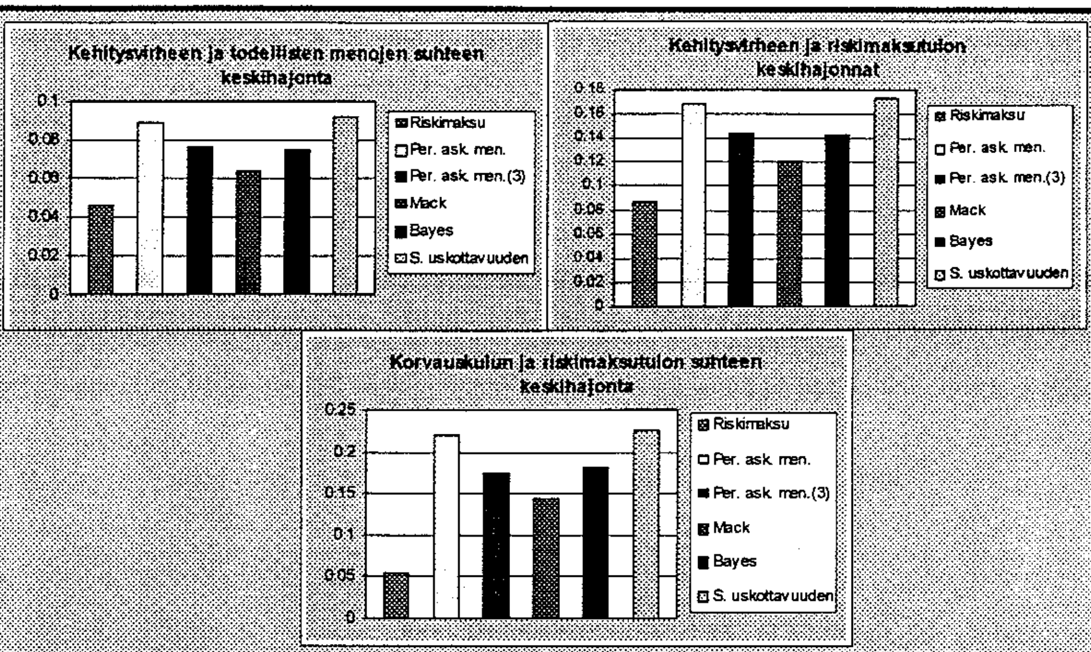

Empiirinen Bayes-menetelmä
0.7
0.2
-0.3
-0.8
R/P
Aika



Kuva 4.1. Jokaisen estimointimenetelmän 50 ensimmäistä realisaatiota kehitysvirheen R ja riskimaksutulon P suhteesta. Realisaation kesto on 30 vuotta. Vahinkomenoprosessi on generoitu ilman inflaation ja vahinkojen lukumäärän vaihtelua. Vahinkojen odotusarvo on vuosittain ollut 10000 ja yksittäisen vahingon koon jakautuman alimmat momentit olivat $a_1 = 0.006$ (miljoonaa markkaa), $a_2 = 0.001$ ja $a_3 = 0.0001$. Lisäksi käytetyn selviämisjakautuman pituus on ollut 7 vuotta eli $d_{max} = 6$.

Yllä olevista kuvista nähdään, että kehitysvirheen ja riskimaksutulon suhde on melko stabiili heilahdellen enemmän tai vähemmän säännöllisesti nollatason kummallakin puolella. Näin ollen voidaan simuloinnissa laskea mielenkiinnon kohteena olevat suureet jokaisena realisaation vuotena ja laskea näistä sitten tasapainotilaa kuvaavien keskihajontojen arvot (ks. Pentikäinen & Rantala (1992)).

Kuvassa 4.2 on graafisesti esitetty pylväsdiagrammin avulla kunkin menetelmän kehitysvirheen R ja riskimaksutulon P suhteen keskihajonnat ja samoin kehitysvirheen R ja todellisten vahinkomenojen suhteen keskihajonnat. Todellisilla vahinkomenoilla tarkoitetaan tässä yhteydessä niitä todellisia korvausmenoja, joiden estimaateista korvausvastuu muodostuu. Samassa kuvassa on myös kuvattu korvauskulun X ja riskimaksutulon suhteen keskihajonnat erikseen kullekin menetelmälle. Vahinkomenoprosessi on generoitu samoilla parametrien arvoilla kuin kuvassa 4.1 esitetyt realisaatiot. Alla olevissa kuvissa ja jatkossa aina muutenkin, kun on kyseessä eri estimointimenetelmistä saatujen tulosten esittäminen pylväsdiagrammien avulla, niin eri estimointimenetelmien järjestys vasemmalta oikealle on aina sama: riskimaksutuloon perustuva menetelmä, peräkkäisten askelmien menetelmä, sen modifikaatio, Mackin menetelmä, empiirinen Bayes-menetelmä ja suurimman uskottavuuden menetelmä.



Kuva 4.2. Eri menetelmien kehitysvirheiden keskihajontoja suhteutettuna hieman eri suureisiin. Ylhäällä vasemman puoleisessa kuvassa on kehitysvirheen ja todellisten vahinkomenojen suhteen keskihajonnat ja oikean puoleisessa kuvassa taas kehitysvirheen ja riskimaksutulon suhteen keskihajonnat. Alla on puolestaan korvauskulun ja riskimaksutulon suhteen keskihajonnat.

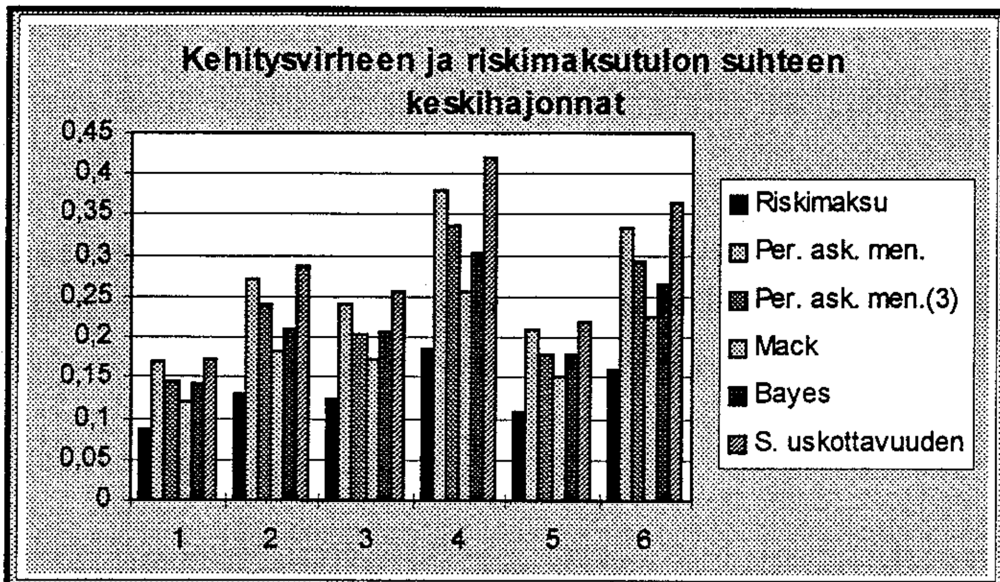
Yllä olevista pylväsdiagrammeista nähdään, että tässä ideaalisessa tapauksessa, kun inflaatiota ja vahinkojen lukumäärän vuosittaisia vaihteluja ei huomioida, menetelmistä pienimmän keskihajonnan korvausvastuun estimaatille antaa riskimaksutuloon perustuva menetelmä. Toiseksi pienimmän keskihajonnan antaa Mackin kredibiliteettimenetelmä. Yllä olevista kuvista nähdään myös, että ottamalla peräkkäisten askelmien menetelmään lisää täysin kehittyneitä kohortteja, niin tulokset paranevat noin 15–20 prosenttia. Tämä on seurausta juuri niistä seikoista, joita mainittiin esiteltäessä peräkkäisten askelmien menetelmän heikkouksia. Yllä kannattaa kiinnittää huomiota myös siihen tosiasiaan, että riskimaksutuloon perustuva menetelmä antaa melko pienen keskihajonnan korvauskulua tarkasteltaessa. Tämä johtuu kyseisen menetelmän eräänlaisesta vaimennusominaisuudesta eli se voi antaa jopa tasaisemman korvauskulun käyttäytymisen kuin tapaus, jossa koko kehitysvirhe on unohdettu eli oletus, että kaikki vahingot maksetaan heti niiden tapahtuessa, on jälleen otettu mukaan. Tämä kyllä puolestaan yleensä sitten näkyy suurempana kehitysvirheen arvona.

Nyt kun on simuloinnista ja erityisesti kyseisen ohjelman tekemisestä on kyse, niin hyvän kuvan ohjelman antamien tulosten oikeellisuudesta antavat osittain myös eri menetelmillä saadut kehitysvirheen ja riskimaksutulon suhteen ja korvauskulun ja riskimaksutulon suhteen keskiarvot. Taulukossa 4.1 onkin annettu jokaisen menetelmän keskiarvot kyseisille suhteille. Suurimman uskottavuuden hieman poikkeavat arvot johtuvat sen estimoinnin harhaisuudesta. Erityisesti suurimman uskottavuuden menetelmässä olisi pitänyt käyttää eri estimointitapaa, mikäli olisi haluttu harhattomat estimaatit. Tässä on kuitenkin osittain haluttu myös testata, kuinka paljon empiirisellä bayes-menetelmällä saadut tulokset poikkeavat sen "alkuarvoista" eli suurimman uskottavuuden menetelmän tuloksista.

Taulukko 4.1. Eri menetelmien antamat kehitysvirheen ja korvauskulun keskiarvot suhteutettuna kehitysvirheen tapauksessa sekä riskimaksutuloon (sarake 1) että todellisiin korvausmenoihin (sarake 2) ja korvauskulun tapauksessa riskimaksutuloon (sarake 3). Periaatteessa kyseisten keskiarvojen pitäisi kehitysvirheen tapauksessa olla nollia ja korvauskulun tapauksessa ykkösiä.

Menetelmä	**1**	**2**	**3**
Riskimaksuhdo	-0.0011	0.0012	1.0000
Per. askelmien	0.0017	0.0031	1.0000
Per. askelmien (3)	-0.0006	0.0019	0.9993
Mack	-0.0019	0.0010	0.9999
Empiirinen Bayes	0.0032	0.0038	0.9998
Suurimman usk.	0.0364	0.0214	0.9997

Yllä on tähän asti tarkasteltu ainoastaan yhtä numeerista esimerkkiä, jonka identifioi kuvan 4.1 kuvatekstissä annetut parametrien arvot. Tässä tapauksessa testattiin myös eri parametrien arvojen vaihteluiden vaikutusta. Kuvassa 4.3 on esitelty pylväsdiagrammina testattujen taustatekijöiden vaikutusta. Kyseisessä kuvassa on kaikki parittomat pylväskokoelmat generoitu käyttämällä 7 vuoden selviämisjakautumaa ja kaikki parilliset on generoitu 13 vuoden selviämisjakautumalla. Pylväskokoelman 1 parametrit ovat samat kuin edelläkin ja kokoelman 2 parametrit muuten samoja lukuunottamatta siis selviämisjakautuman kestoa. Pylväskokoelmat 3 ja 4 on generoitu muuten samoilla parametrien arvoilla paitsi, että vuosittaisten vahinkojen odotusarvo oli puolet pienempi eli 5000. Kokoelmat 5 ja 6 ovat taas muuten samoja kuin 1 ja 2, mutta yksittäisen vahingon koon jakautuman toinen momentti oli hieman suurempi eli 0.0015 antaen näin ollen vahinkojen koolle noin kaksi kertaa suuremman keskihajonnan. Kyseisestä kuvasta nähdään, että parametrien vaihtelut antavat tuloksia, jotka ovat intuitiivisesti melko luonnollisia. Jos yksittäisen vahingon koon keskihajonta on hieman suurempi, niin on luonnollista, että eri menetelmien antamien estimaattien keskihajonnat ovat myös hieman suurempia. Vastaavasti, jos vahinkoja sattuu suhteessa vähemmän, niin tällöin tilastoaineistosta on myös vaikeampi ennustaa tulevaisuutta johtuen juuri sen pienemmästä kattavuudesta. Mielenkiintoisinta kuvassa on kuitenkin eri menetelmien keskinäinen käyttäytyminen, joka on nyt lähes tulkoon täysin identtistä.

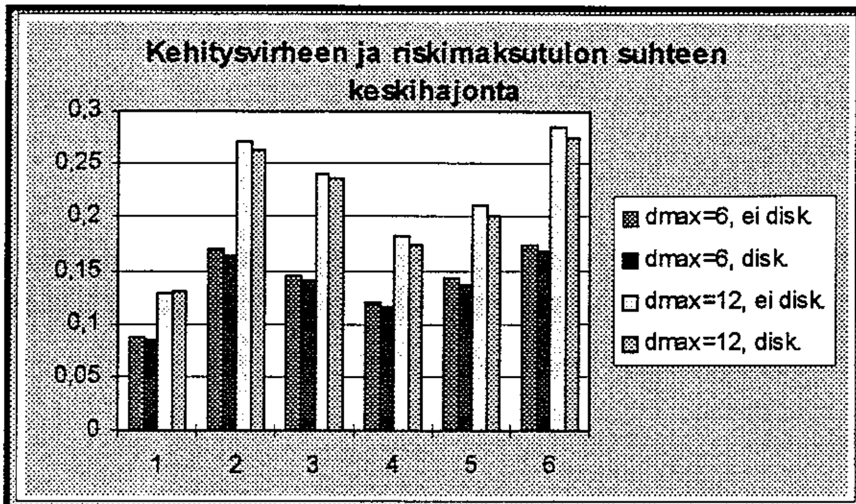


Kuva 4.3. Parametrien vaihteluiden vaikutus eri menetelmiin. Pylväät ovat jaettu ryhmiin niin, että kukin eri ryhmä kuvaa omaa parametrien arvojoukkoa. Kussakin ryhmässä on lisäksi kullekin menetelmälle oma pylväs, joiden järjestys vasemmalta oikealle seuraa kuvan oikealla olevaa kuvatekstiä.

4.1.2. Diskonttaus

Seuraavassa on tarkasteltu diskonttauksen vaikutusta menetelmien antamiin estimaatteihin tai lähinnä näiden keskihajontoihin. Tässä yhteydessä nyt ja jatkossa muutenkin aina kun tarkastellaan diskontattua korvausvastuuta, niin myös riskimaksumaksutulo on määritelty kyseinen diskonttaus huomioon ottaen. Tämä on tehty laskettaessa riskimaksutuloa kehityskolmiosta siten, että kyseisen termin määrittelyssä (ks. pykälä 4.2) kaikki kehittymisvuoden s termit on diskontattu diskonttaustekijän $v(t, s)$ mukaan. Nyt on diskonttaustekijänä käytetty vakioarvoa 0.05 eli diskonttaus on suoritettu viiden prosentin korkokehityksen mukaan. Kuvassa 4.4 on saatuja tuloksia verrattu edellisessä pykälässä olleeseen perusmalliin. Nyt siis käytetyt arvot ovat

- yksittäisen vahingon koon jakautuman origomomenteille: $a_1 = 0.006$, $a_2 = 0.001$ ja $a_3 = 0.0001$
- realisaatiolle: kesto 29 vuotta ja lukumäärä 275 lyhyemmällä selviämisjakautumalle ja 33 ja 185 pidemmälle selviämisjakautumalle (ovat jatkossa aina nämä),
- vahinkojen lukumäärälle vuosittain: 10000.



Kuva 4.4. Diskonttauksen vaikutus eri menetelmien antamiin kehitysvirheen ja riskimaksutulon suhteen keskihajontaan. Sarake 1 kuvaa riskimaksutulomenetelmällä saatuja tuloksia, sarake 2 peräkkäisten askelmien menetelmällä saatuja tuloksia, jne. vastaavassa järjestyksessä kuin aikaisemminkin.

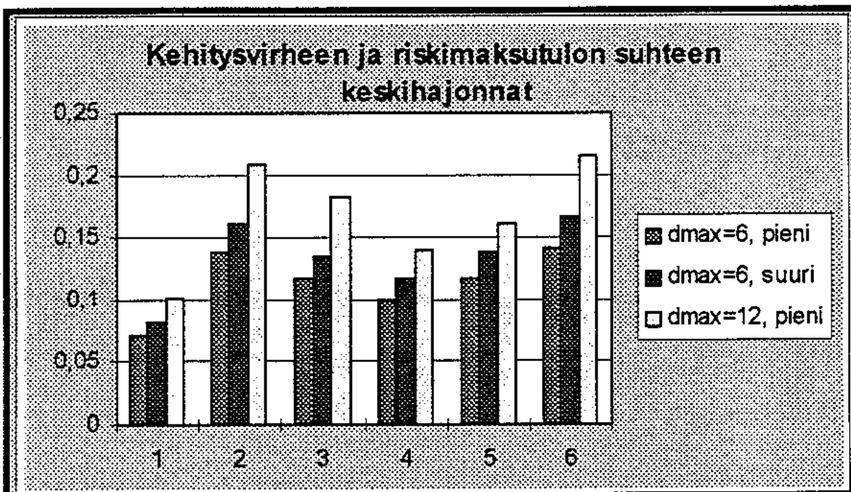
Nyt nähdään, että diskonttauksella ei ole suurta merkitystä saatuihin tuloksiin. Kannattaa kuitenkin kiinnittää huomiota selvään yleiseen trendiin eli siihen, että diskonttaus kuitenkin hieman pienentää saatuja keskihajontoja. Toisaalta tämä ei kuitenkaan ole suuri yllätys, koska eri kohorttien viimeiset solut eli ne vahingot, joiden selviäminen kestää kauimmin, ovat yleensä kaiken herkimpiä estimointivirheille. Tämä johtuu siitä, että kyseisille kehittymisvuosille on vähän tilastoaineistoa. Nyt kuitenkin näiden pääoma-arvo on huomattavasti pienempi ja näin ollen virheiden, joita kyseisissä soluissa ilmenee, vaikutus pienenee.

Sekä diskonttaamattomassa tapauksessa että diskontatussa tapauksessa on merkille pantavaa kehitysvirheen suhteellisen suuri vaihtelu verrattuna vuosittaisiin riskimaksuihin tai todellisiin vahinkomenoihin, vaikka tässä ideaalisessa tapauksessa on kaikki taustatekijät unohdettu. Todellisiin vahinkomenoihin verrattuna kehitysvirheen vaihtelun suuruus oli 5–9 prosenttia ja riskimaksutuloon verrattuna kehitysvirheen vaihtelu oli 9–19 prosentin luokkaa.

4.2. Vahinkojen lukumäärän vaihtelu

4.2.1. Ei diskonttausta

Nyt on vahinkomenoprosessia mallinnettu hieman todenmukaisemmin ja otettu vahinkojen lukumäärän vaihtelu mukaan. Eri menetelmiä on testattu kahdelle eri selviämisjakautuman kestolle: 7 vuoden ja 13 vuoden. Lisäksi lyhyemmälle selviämisjakautumalle on testattu kahta hieman erilaista vahinkojen lukumäärän heilahteluprosessia. Kummatkin ovat olleet toisen asteen autoregressiivisiä aikasarjoja ja parametrin arvot (ks. luku 3) ovat olleet toiselle $\alpha_1 = 0.438$, $\alpha_2 = -0.096$, $\sigma = 0.18$, $\gamma = 0.2$ ja multiplikatiivisen trendin kasvukerroin $r_g = 1.02$ sekä toiselle muuten samat paitsi $\sigma = 0.23$, $\gamma = 0.25$ ja multiplikatiivisen trendin kasvukerroin $r_g = 1.01$. Jälkimmäiseen on kuvassa 4.5 viitattu termillä suuri vaihtelu ja vastaavasti ensimmäiseen termillä pieni vaihtelu.



Kuva 4.5. Menetelmien antamat keskihajonnat kehitysvirheen ja riskimaksutulon suhteelle. Lyhyelle selviämisjakautumalle on testattu kahta erilaista vahinkojen lukumäärän heilahtelua – suurta ja pientä vaihtelua.

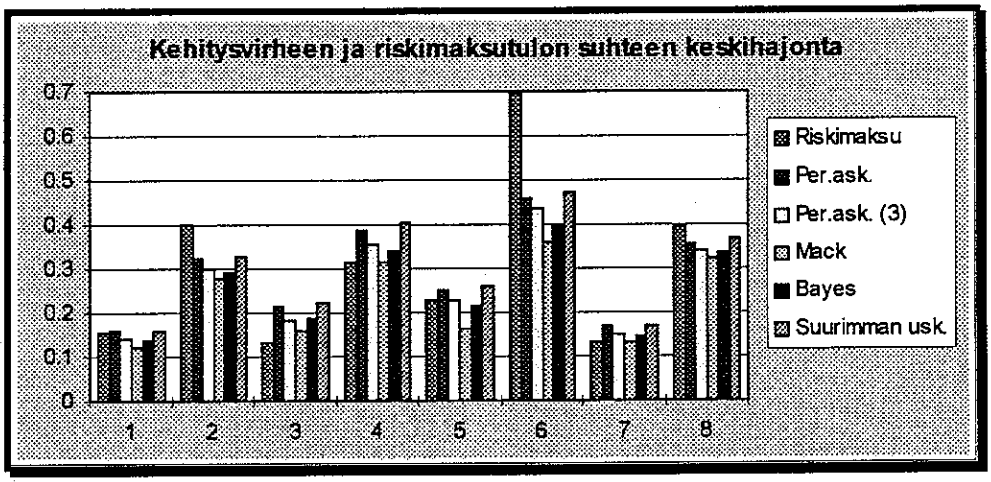
Yllä olevasta kuvasta nähtävät tulokset ovat jälleen melko luonnollisia, mitä tulee eri parametrien ja taustatekijöiden vaihteluun. Mitä suurempi vahinkojen lukumäärän vaihtelu sitä suurempi myös kehitysvirheen ja riskimaksutulon suhteen vaihtelu. Vastaavasti mitä kauemmin selviäminen kestää sitä suurempi kehitysvirheen vaihtelu on. Yllättävintä ehkä nyt kuitenkin on, että verrattaessa pienemmällä vahinkojen lukumäärän vaihtelulla saatuja tuloksia edellisessä pykälässä saatuihin tuloksiin, joissa siis vahinkojen vaihtelua ei ollut mallinnettu, niin jokaisella menetelmällä kehitysvirheen ja riskimaksutulon suhteen keskihajonta oli nyt hieman pienempi.

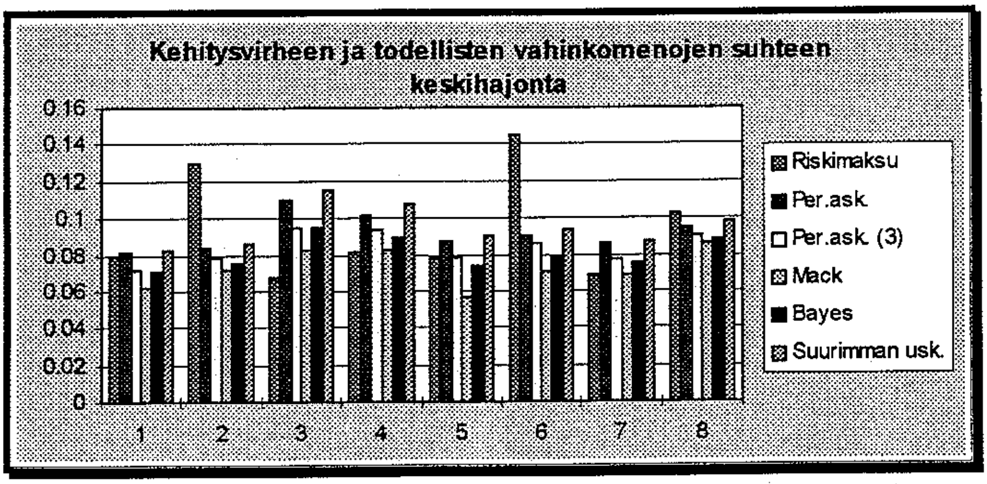
4.3. Vahinkojen lukumäärän vaihtelu ja inflaatio

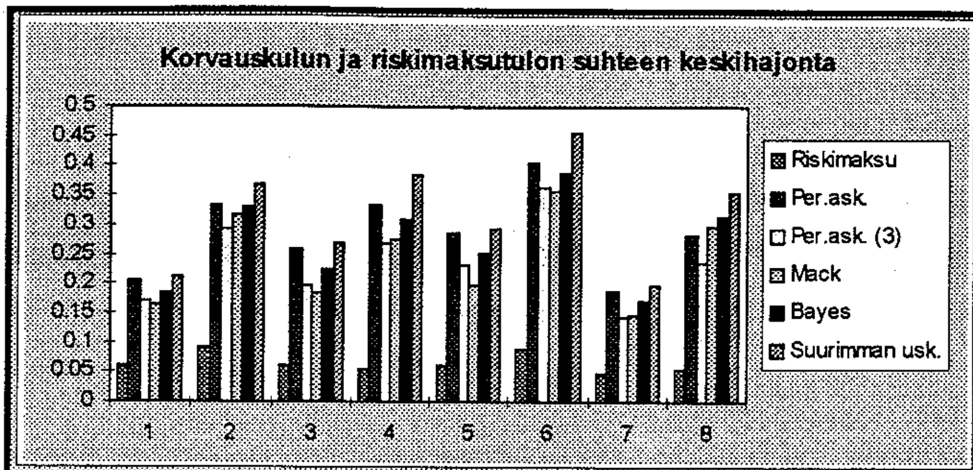
4.3.1. Ei diskonttausta

Lopuksi on tarkoitus ottaa myös inflaation vaikutus huomioon. Aikaisemmin ei varsinaisesti ole mainittu, mitä inflaatiolla tarkalleen ottaen tässä yhteydessä tarkoitetaan. Useastihan riskiteoriassa mallinnettaessa inflaatiota siihen luetaan mukaan myös niin kutsuttu sosiaalinen inflaatio. Tässä yhteydessä kyseisellä asialla ei ole merkitystä, joten sitä ei sen tarkemmin nyt määritellä.

Inflaatio on nyt tarkoitus mallintaa käyttäen hyväksi ensimmäisen asteen autoregressiivistä aikasarjaa (ks. Daykin et. al). Inflaatio on jatkossa generoitu käyttäen kyseiselle aikasarjalle seuraavia parametrien arvoja: kerroin a = 0.7, keskimääräinen inflaatio iav = 0.05 ja virhetermin keskihajonta σ = 0.015. Mikäli käytetään jotain muita arvoja, tästä mainitaan erikseen asian yhteydessä.







Kuva 4.6. Yllä on kolme eri kuvaa, joissa kahdessa ensimmäisessä on kuvattu kehitysvirhettä ja viimeisessä korvauskulua. Kussakin kuvassa on kahdeksan eri vahinkomenoprosessia, joille jokaiselle on annettu kunkin eri menetelmän antamat keskihajonnat.

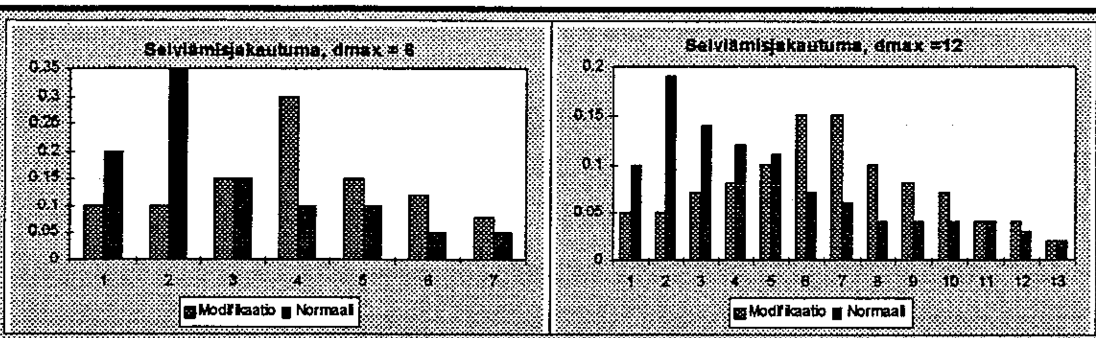
Yllä olevissa kuvissa on esitetty eri menetelmien antamat kehitysvirheen ja korvauskulun keskihajonnat kahdeksalle eri vahinkomenoprosessille. Seuraavassa on erikseen määritelty kukin prosessi:

1. Normaali inflaatio, pieni vahinkojen lukumäärän vaihtelu (ks. edellinen pykälä), aloitusvuoden vahinkojen lukumäärän taso 10000, normaali yksittäisen vahingon koon jakautuma ja selviämisjakautuman kesto on 7 vuotta.
2. Sama kuin yllä, mutta vahinkojen selviämisjakautuman kesto on 13 vuotta.
3. Sama kuin 1, mutta aloitusvuoden vahinkojen lukumäärän taso 5000.
4. Sama kuin yllä, mutta vahinkojen selviämisjakautuman kesto on 13 vuotta.
5. Sama kuin 1, mutta vahinkojen selviämisjakautuman muoto hieman erilainen (ks. kuva 4.7).
6. Sama kuin yllä, mutta modifioidun vahinkojen selviämisjakautuman kesto 13 vuotta.
7. Sama kuin 1, mutta inflaatio mallinnettu eri tavalla eli kerroin $a = 0.7$, keskimääräinen inflaatio $iav = 0.06$ ja virhetermin keskihajonta $\sigma = 0.02$.
8. Sama kuin yllä, mutta vahinkojen selviämisjakautuman kesto 13 vuotta.

Tarkasteltaessa yllä olevista kuvista kahta ensimmäistä eli kehitysvirheeseen liittyviä kuvia nähdään, että riskimaksutuloon perustuva menetelmä menettää aikaisemmissa pykälissä ilmenneen ehdottaman "paremmuutensa" muihin nähden. Tämä johtuu suureksi osaksi vuosittaisen riskimaksutulon määräämisen vaikeutumisesta eli inflaatio on tuonut vahinkomenoprosessiin lisää vaihtelua ja näin ollen myös uusien riskimaksujen määrääminen vanhojen avulla ei toimi enää niin hyvin kuin aikaisemmissa pykälissä. Muuten menetelmien "paremmuus" pysyy suurin piirtein samana eli erityisesti Mackin menetelmä näyttää suoriutuvan melko hyvin muihin verrattuna. Vastaavasti empiirinen Bayes-menetelmä näyttää kussakin tapauksessa olevan jonkin verran kyseistä menetelmää heikompi, mutta taas hieman paremmin käyttäytyvä kuin peräkkäisten askelmien menetelmän modifikaatio. Mitä eri vahinkomenoprosesseihin tulee, saadut tulokset ovat melko luonnollisia. Mielenkiintoista on kuitenkin ja ehkä yllättävääkin Mackin menetelmän reagointi modifioituun selviämisjakautumaan verrattaessa kehitysvirheen ja todellisten korvausmenojen suhdetta; keskihajonnat näyttävät olevan normaaliin jakautumaan verrattuna pienempiä, vaikka jakautuma onkin keskittynyt keskemmällä selviämistä.

Kyseisistä kuvista kannattaa huomioida kehitysvirheen ja riskimaksutulon suhteen ja kehitysvirheen ja todellisten vahinkomenojen suhteen poikkeavat pylväsdiagrammikuviot. Erityisesti ensimmäisen suhteen kyseessä ollessa selviämisjakautuman pituus aiheuttaa melko suuren vaihtelun verrattuna lyhyempään vahinkojen selviämiseen.

Korvauskulun ja riskimaksutulon suhteen keskihajonta on suurin piirtein samanlaista kuin kehitysvirheen ja riskimaksutulon suhteen keskihajonnan käyttäytyminen ainakin, jos selviämisjakautumien käyttäytymistä verrataan (ks. kuva 4.6). Muuten tulokset ovatkin lähes täysin erilaisia. Edellä jo mainittiinkin riskimaksutuloon perustuvan menetelmän vaimennusominaisuudesta ja nyt kyseinen ilmiö näkyy selvästi. Ennen kaikkea kyseisen ilmiön syy näkyy eli se, että korvauskulun pienempään vaihteluun päästään suuremman kehitysvirheen kustannuksella. Vastaavasti kuvasta 4.6 näkyy osittain myös toinen hyvin tunnettu ilmiö erityisesti peräkkäisten askelmien menetelmän suhteen. Kyseinen ilmiö on, että peräkkäisten askelmien menetelmä sopii paremmin tilanteisiin, joissa vaihtelu on suurta. Hieman yllättävää on myös muutamassa tapauksessa peräkkäisten askelmien menetelmän modifikaation "paremmuus" Mackin menetelmään nähden. Tämä näyttää erityisesti toteutuvan pidemmän selviämisjakautuman suhteen. Yllättävää on myös vaihtelevamman inflaation antamat pienemmät keskihajonnat.



Kuva 4.7. Vahinkojen lukumäärän selviämisjakautuma. Vasemman puoleisessa kuvassa on lyhyemmän selviämisjakautuman kuvaaja ja oikean puoleisessa pidemmän jakautuman kuvaaja. Modifioitua selviämisjakautumaa on yllä käytetty vahinkomenoprosesseissa 5 ja 6.

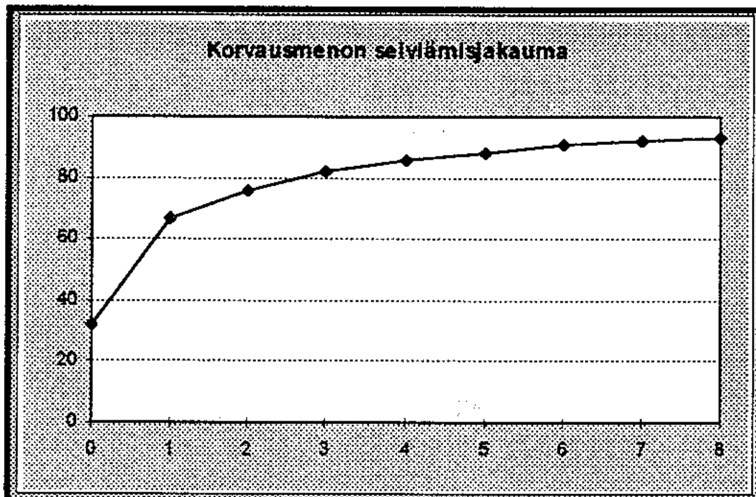
5. Käytännön esimerkki

Edellisessä luvussa testattiin luvussa 3 esiteltyjä menetelmiä simulaatiomallien avulla. Nyt tässä luvussa on tarkoitus testata kuinka eri menetelmät suoriutuvat todellisen tilastoaineiston kanssa. Tässä yhteydessä ei voida kyseisiä menetelmiä testata vastaavalla tavalla kuin simulaatiomallien yhteydessä, vaan tarkoituksena on vain esimerkin avulla näyttää kuinka menetelmät toimivat ja mitä informaatiota ne tarvitsevat ja mihin taustatekijöihin pitää kiinnittää erityistä huomiota kussakin menetelmässä.

Käytettävä tilastoaineisto koostuu vastuuvakuutuksen korvausten selviämisestä, joka perustuu yhtiöiden toimittamiin aineistoihin vahinkovuosilta 1978-1993. Aineisto on kerätty ja valmisteltu Suomen Vakuutusdata Oy:ssa ja se on pitkälti vastaava kuin aineisto, jota H.Lonka, J.Jacobsson ja V. Knuuti ovat käyttäneet omissa tutkimuksissaan tarkastellessaan vastuuvakuutuksen korvausten selviämistä. Käsiteltävän tilastoaineiston korvaukset on jaettavissa luokkiin vahingonkorvausten ominaisuuksien perusteella; esimerkiksi aineisto voidaan jakaa vakuutustyypin mukaan (tuotevastuuvakuutus, yritysten toiminnan vastuuvakuutus, jne.) ja korvauslajin mukaan (selvittelykulut, henkilövahingot ja omaisuuskorvaukset). Yleisesti ottaen käytettäessä jotakin tilastoaineistoa on yleensä pyrittävä jakamaan tilastoaineisto luokkiin siten, että kukin luokka olisi tarkasteltavilta ominaisuuksiltaan mahdollisimman homogeeninen kuitenkin siten, että estimoinnin stabiilisuuden kannalta kunkin osa-aineiston tulisi sisältää riittävästi havaintoja. Nyt on päädytty tarkastelemaan vastuuvakuutuksen vahinkoja (pl. oikeusturva) pitämällä kaikki eri vakuutustyypit samassa luokassa ja muodostamalla jako sen mukaan ovatko ne henkilövahinkoja vai omaisuuskorvauksia. Tällä tavalla on saatu suhteellisen homogeeniset luokat esimerkin tarkoituksen kannalta. Todellisuudessa eri vakuutustyyppejä olisi ehkä syytä ja tarvetta tarkastella erikseen, mutta tässä yhteydessä näin ei ole menetelty, koska nyt on tarkoitus vain esittää eri menetelmien etuja ja haittoja sovellettaessa näitä todelliseen ongelmaan.

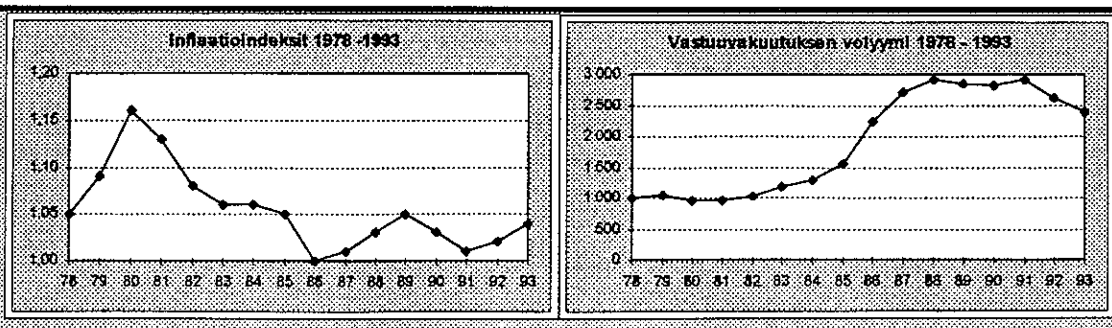
Nyt on tarkoitus siis laskea suorat estimaatit korvausvastuille. Korvausvastuun estimaatteja tullaan laskemaan ensin käsittelemättömälle aineistolle ja sitten asteittain aina prosessoidummalle aineistolle. Tällä tarkoitetaan, että ensin lasketaan kunkin menetelmän antamat estimaatit alkuperäiselle aineistolle ja tämän jälkeen otetaan huomioon asteittain inflaatio, volyymi ja muut mahdolliset tekijät. Varsinainen korvausvastuu estimoidaan olettaen, että koko korvausvastuu koostuu tuntemattomista vahingoista, koska tässä yhteydessä ei ole ollut käytössä vahinkovuosittaisia arvioita vahinkokohtaisesti arvioiduista korvauksista (tunnetuista maksamattomista vahingoista). Todellisuudessa tämä aiheuttaa melkolailla epästabiilisuutta estimaatteihin, koska todellisuudessa vahinkokohtaisten varausten osuuden voi olettaa olevan merkittävän ja näiden arvioiden tarkkuus melko hyvä verrattuna tuntemattomien arvioihin. Lisäksi vahingoista puhuttaessa on nyt syytä korostaa, että korvauksilla tarkoitetaan maksettuja korvauksia eli vahinkojen alkuperäiskoko ei aina ilmene käytetystä aineistosta.

Kyseinen vakuutuslaji on valittu tarkastelun pohjaksi, koska se on keskimääräisesti melko pitkähäntäistä. Lisäksi aineistoa aiemmin tutkineiden tutkimuksissa on tultu siihen tulokseen, että kyseisen vakuutuslajin selviäminen ei ole merkittävästi muuttunut vuosien varrella. Eri menetelmillehän oli yhteistä, että tilastoaineiston vahinkovuosina ei ole tapahtunut merkittävää muutosta vahinkojen selviämisessä. Kuvassa 5.1 on esitetty edellä mainituissa tutkimuksissa estimoitu korvausten selviämisjakautuma kyseinen vakuutuslajin osalta.



Kuva 5.1. Vastuuvakuutuksen selviämisjakautuma (lähde: Vastuuvakuutuksen korvausmenon selviäminen, V. Knuuti 1994).

Osa menetelmistä käyttää estimoinnissa hyväkseen taustatekijöiden vaihteluita eli mikäli inflaatio ja vakuutuslajin volyymin mitta on tiedossa tai estimoitavissa muusta aineistosta kuin käytettävästä korvausaineistosta, niin tällöin näiden menetelmien oletukset saadaan realistisemmaksi. Inflaation kuvaajaksi tässä yhteydessä valittu J. Jacobssonin tutkimusta mukaillen tilastokeskuksen julkaisema tukkuhintaindeksi, huomioimatta kuitenkaan indeksin laskuja. Volyymin mittariksi on puolestaan tässä yhteydessä valittu koko vastuuvakuutusluokan vuosittainen vakuutusmaksutulo, suodattamalla näistä kuitenkin inflaation vaikutus pois edellä mainitun inflaatiomitan mukaisesti. Kuvassa 5.2 on esitetty sekä inflaatioindeksit 1978-1993 että normeeratut volyymiluvut vastaavalta ajanjaksolta.



Kuva 5.2. Inflaatioindeksit ja volyymi 1978–1993. Volyymia tarkasteltaessa on vuoden 1978 volyymiksi asetettu 1000.

Eri menetelmistä riskimaksutuloon perustuva menetelmä on estimoinnin kannalta hieman muista poikkeava, koska se pohjautuu vuoden riskimaksutuloon ja näin ollen se ei ole suoraan sovellettavissa pelkän korvausaineiston pohjalta. Lisäksi, koska kyseinen menetelmä on enemmänkin periaate kuin jonkinlainen estimointimenetelmä, selviämisjakauma pitää määrätä käyttäen joko subjektiivista harkintaa tai estimoida käyttäen hyväksi esimerkiksi muita tässä tutkielmassa esitettyjä menetelmiä. Tässä yhteydessä kyseistä menetelmää ei ole otettu mukaan esimerkkeihin, koska vuotuinen riskimaksutulo olisi pitänyt johtaa jotenkin vakuutusmaksutulosta. Tämä olisi tietenkin voitu tehdä estimoimalla jonkinlaiset vuotuiset vahinkosuhteet. Koko prosessi olisi kuitenkin järkevien tuloksien kannalta ollut melko laaja ja näin ollen kyseinen menetelmä on jätetty testattavista menetelmistä pois. Simulointimalleilla kyseinen menetelmä oli helppo ottaa mukaan, koska yksinkertaisilla keinolla oli helposti johdettavissa jonkinlaiset arviot riskimaksutulolle ja selviämisjakautumalle.

5.1. Henkilövahinkojen korvausvastuun estimointi

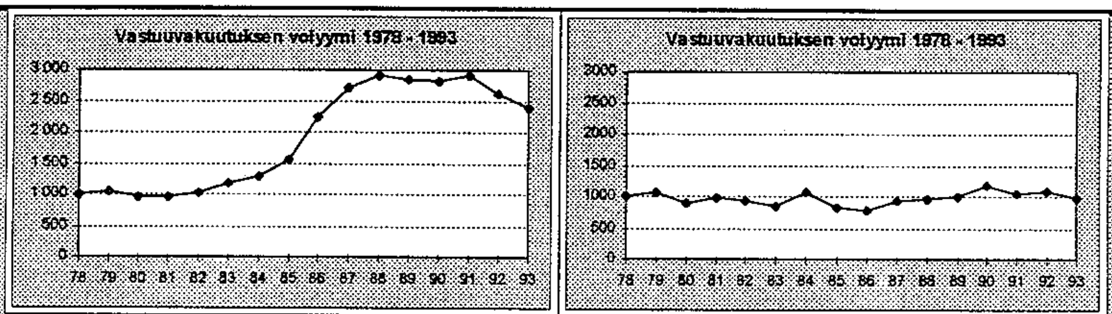
Ensin eri menetelmiä on testattu vastuuvakuutuksen henkilövahingoista kerätyn tilastoaineiston kanssa (liite A). Syynä tähän on, että intuitiivisesti ajatellen voisi kuvitella näiden käyttäytyvän stabiilimmin ja paremmin jonkin tietyn mallin mukaisesti kuin omaisuuskorvausten. Näin myös tässä tapauksessa onkin ja jälkimmäisellä tilastoaineistolla saatuihin estimaatteihin palataan myöhemmin.

Taulukossa 5.1 on esitetty kunkin menetelmän estimaatit neljässä eri tapauksessa eli kun mitään taustatekijöitä ei ole huomioitu ja kun taustatekijät (volyymi ja inflaatio) on lisätty kumpikin vuorollaan ja lopuksi kumpikin yhdessä. Diskonttausta ei ole nyt huomioitu ollenkaan, koska vaikkakin sillä on merkittävä vaikutus estimaattien tasoihin, niin varsinaisiin eri menetelmien käyttäytymiseen sillä ei ole vaikutusta.

	Chain-Ladder	**Mackin menetelmä**	**Empiirinen Bayes**	**Maximum likelihood**
Ei mitään	48 799	36 820	44 616	53 442
Volyymi	48 744	51 839	50 743	53 442
Inflaatio	45 216	39 891	43 623	50 991
Kummatki	44 414	57 688	53 653	50 991

Taulukko 5.1. Eri menetelmien antamat korvausvastuun estimaatit eri taustaoletuksilla.

Tarkasteltaessa Mackin menetelmää eri variaatioilla mieltä jää kaivamaan hieman outo heilahtelu. Syy kyseiseen heilahteluun on volyymin huomioon ottaminen. Tämä johtuu vastuuvakuutuksen vahinkosuhteen parantumisesta 1980-luvun alusta vuosikymmenen loppuun ja 1990–luvun alkuun mennessä. Mackin menetelmä tulkitsee tämän puolestaan siten, että alhainen korvausmeno on ollut vain satunnaista ja siis näin ollen keskimääräisesti korvausmenon pitäisi olla suurempaa. Vastaavasti mikäli volyymia ei oteta huomioon ilmiö kärjistyy päinvastaisesti. Vastaanlainen ilmiö oli huomattavissa myös Empiirisen bayes-menetelmän kanssa, joskaan ei yhtä suuressa mittakaavassa. Maximum-likelihood-menetelmään volyymilla ei ole merkitystä estimointitavasta johtuen. Chain-Ladder-menetelmällä volyymi puolestaan vaikuttaa jonkin verran, mutta sen yhteydessä muutos johtuu siitä, että vanhempien vahinkovuosien paino nousee ja luonnollisesti jonkinlaista pientä variaatiota on vahinkojen selviämisessäkin ollut. Volyymin aiheuttama heilahtelua estimaatteihin voidaan kiertää valitsemalla volyymimittari jollakin muulla tavalla. Vaihtoehtoisia valintoja voisi olla joko esimerkiksi ensimmäisen suoritusvuoden vahinkomeno tai vahinkojen lukumäärä. Valitsemalla volyymimittariksi ensimmäisenä vuonna sattuneiden vahinkojen lukumäärä saadaan melkoisesti aikaisemmasta poikkeava volyymimitta. Kuvassa 5.3 on esitetty uusi ja vanha volyymimitta rinnakkain

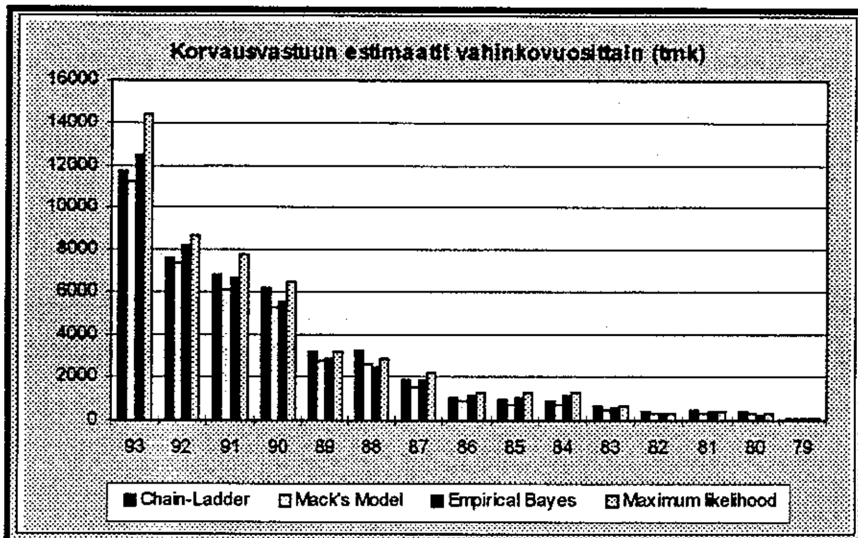


Kuva 5.3. Volyymimitat vakuutusmaksutulosta johdettuna (vasen puoli) ja ensimmäisen vuoden vahinkojen lukumäärästä johdettuna.

Laskettaessa korvausvastuun estimaatit uudelleen tällä uudella volyymimitalla saadut tulokset näyttävät käyttäytyvän huomattavasti paremmin kuin aiemmin (taulukko 5.2). Sovitetun tilastoaineiston tarkempi analysointi osoitti, että tämä uusi volyymitta onkin huomattavasti paremmin sopiva tähän yhteyteen, kuin suoraan vakuutusmaksutulosta johdettu volyymi. Syykin tähän on jo mainittu eli jälkimmäinen ei huomio ollenkaan vahinkosuhteen mahdollista vaihtelua.

	Chain-Ladder	**Mackin menetelmä**	**Empiirinen Bayes**	**Maximum likelihood**
Ei mitään	48 799	36 820	44 616	53 442
Volyymi	48 903	37 252	45 159	53 442
Inflaatio	45 216	39 891	43 623	50 991
Kummatki	45 425	40 611	44 855	50 991

Taulukko 5.2. Eri menetelmien antamat korvausvastuun estimaatit (tuhansina markkoina) eri taustaoletuksilla, kun volyymi on huomioitu uudella tavalla.



Kuva 5.4. Eri menetelmien antamat vahinkovuosittaiset korvausvastuun estimaatit, kun inflaatio ja volyymi (ensimmäisen vuoden vahingoista johdettu) on huomioitu. Luvut on esitetty tuhansina markkoina.

Kuvasta 5.4 nähdään, että mitään suurta vaikutusta ei eri estimointimenetelmillä ollut varsinaiseen selviämisjakautumaan. Lisäksi myöskään eri taustatekijöillä ei ollut suurtakaan vaikutusta, vaikka luonnollisesti inflaation vaikutuksen poistaminen painotti loppuosaa hieman vähemmän. Kaiken kaikkiaan henkilövahinkojen osalta ei eri menetelmiä pitänyt juurikaan sovittaa tilastoaineistolle sopivaksi tai oikeammin tilastoaineisto ei tässä yhteydessä vaatinut juuri minkäänlaisia etukäteismuunnoksia lukuunottamatta volyymimitan aiheuttamia ongelmia.

5.2. Omaisuuskorvausten korvausvastuun estimointi

Seuraavaksi eri menetelmiä testattiin vastuuvakuutuksen omaisuusvahingoista kerätyllä tilastoaineistolla (liite B). Tämän kanssa asiat eivät enää olletkaan niin yksinkertaista kuin edellisen aineiston kanssa. Ensimmäinen ongelma, johon muutamalla menetelmällä törmätään, on joinakin vahinkovuosina esiintyvät negatiiviset korvaussummat erityisesti myöhempinä suoritusvuosina. Esimerkiksi empiirinen Bayes- ja Maximum likelihood–menetelmät eivät näistä suoriudu johtuen vahinkojen log-normaali-oletuksesta. Muihin menetelmiin negatiiviset arvot soveltuvat siinä missä positiivisetkin.

Ensimmäinen mieleen tuleva keino kiertää negatiiviset korvaussummat on korvata ne jollakin pienellä positiivisella vakiolla (esimerkiksi 1), koska tällöin ainakin yksi korvausvastuun estimoinnissa huomioon otettava ehto eli turvaavuus toteutuu. Toinen vastaavanalainen vaihtoehto on lisätä kuhunkin soluun jokin vakio siten, että kaikki solut tulevat positiivisiksi Laskemalla estimaatit kummallakin tavalla ovat saadut tulokset ainakin empiirisen Bayes- ja Maximum likelihood-menetelmän osalta lähes järjettömiä. Tätä voisi kutsua esimerkiksi klassisesta sanonnasta “Garbage in, Garbage out”. Toisin sanoen, mitään estimointimenetelmiä ei ikinä kannata suin päin lähteä soveltamaan tuntemattomaan tilastoaineistoon. Syy näihin outoihin tuloksiin tässä yhteydessä selviää melko helposti ja se on poikkeuksellisen suuret korvaukset, joita on muutamina vuosina kohdattu. Pelkästään tilastoaineistoa tarkastelemalla näkee, että poikkeuksellisen suuria korvauksia on ollut enemmän kuin yksi. Tarkemman tarkastelun perusteella kyseiset korvaukset ovat peräisin pääosin tuotevastuuvakuutuksen vahingoista.

Tässä yhteydessä näiden poikkeuksellisen suurten korvausten aiheuttama ongelma on kierretty ottamalla näistä korvauksista huomioon vain osa. Tämä on tehty siten, että suoritusvuosina 5-7 silmään pistävät suuret solujen korvausmenot ovat leikattu 1.5 Mmk tasolle ja myöhemmät suoritusvuodet 0.75 Mmk tasolle korvausten lukumäärätiedoista kuitenkin tarkistaen, että kyseinen suuri korvausmeno ei ole johtunut suuresta vahinkomäärästä. Kyseinen melko karkea tapa on valittu, koska yksityiskohtaisia tietoja vakuutuskorvausten määristä ei ole.

Kyseisen sovituksen jälkeen lopullisiksi korvausvastuun estimaateiksi saatiin Chain-Ladder menetelmällä 121.2 Mmk ja Mackin menetelmällä 102.5 Mmk. Kaksi muuta menetelmää ei vielä tämän jälkeenkään oikein toimi stabiilisti. Edellä esitetyillä kahdella negatiiviset korvausmenot kierrettävällä tavalla laskettujen estimaattien tulokset eroavat toisistaan melkoisesti. Erityisesti tapa, jossa negatiiviset korvaukset korvataan pienellä vakiolla on erittäin herkkä sen suhteen, mikä kyseiselle vakiolle asetetaan arvoksi. Mikäli kyseiselle vakiolle asetetaan arvo 1 ovat saadut tulokset kertaluokkaa suuremmat kuin Chain-Ladder- ja Mackin menetelmällä saadut! Vastaavasti myös toinenkin vaihtoehtoinen tapa oli yhtä herkkä sille, mikä arvo kaikkiin soluihin lisättävälle arvolle annetaan. Tämä viittaa siihen, että kyseisten menetelmien log-normaalisuus oletus rikkoontuu. Yksi tapa olisi jollakin tavalla optimoida lisättävän vakion arvo siten, että modifioitun aineiston poikkeavuus log-normaalisuudesta olisi mahdollisimman pieni.

Tässä yhteydessä kyseistä aineiston sovitusta ei ole suoritettu ja näin ollen kuvassa 5.4 on esitetty vain Chain-Ladder ja Mackin menetelmällä saadut solukohtaiset estimaatit.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1978	8 045	5 740	1 093	651	583	1 666	438	454	282	-12	31	96	0	0	1	18	
1979	9 902	7 092	1 707	857	945	1 500	1 317	98	70	500	33	500	380	4	-133	17	17
																22	22
1980	11 673	10 862	2 838	1 021	1 117	791	760	260	123	-1	556	169	-7	-3	-65	19	-46
															-79	25	-54
1981	16 676	15 006	3 953	1 827	1 084	892	233	255	147	55	134	3	3	0	-79	23	-56
														0	-100	31	-69
1982	17 761	16 171	4 426	1 508	884	1 370	1 654	132	95	92	92	27	125	0	-85	25	65
													125	0	-106	33	52
1983	20 484	21 926	6 167	1 983	1 316	1 399	1 500	128	-17	44	-2	251	152	1	-103	30	331
												247	152	0	-129	40	310
1984	30 635	28 839	8 503	5 040	2 470	1 694	1 500	-46	518	321	292	349	212	1	-144	42	752
											308	353	218	1	-184	58	754
1985	28 982	33 783	6 021	12 354	4 499	731	379	342	390	319	330	395	240	1	-163	47	1 169
										280	341	391	241	1	-204	64	1 114
1986	34 567	29 001	8 071	2 972	1 064	911	648	104	396	308	318	381	232	1	-157	46	1 525
									319	255	311	356	220	1	-186	59	1 335
1987	44 913	46 653	8 908	10 900	940	1 806	146	507	578	450	465	558	339	1	-230	67	2 735
								506	475	379	463	531	327	1	-277	87	2 492
1988	41 013	30 883	5 820	4 756	3 760	1 469	1 576	407	464	361	373	447	272	1	-184	54	3 771
							1 642	394	369	295	360	413	255	1	-216	68	3 581
1989	33 040	35 563	8 618	11 571	4 601	2 276	1 684	435	496	386	399	478	290	1	-197	57	6 305
						2 424	1 767	424	398	318	388	444	274	1	-232	73	6 279
1990	46 966	97 416	19 102	3 986	5 342	4 067	3 008	777	886	689	713	854	519	2	-352	103	16 608
					5 378	4 547	3 315	795	746	596	727	834	514	2	-436	137	17 155
1991	36 758	34 796	9 573	6 520	2 994	2 279	1 686	435	497	386	399	479	291	1	-197	57	15 827
				5 474	2 811	2 376	1 733	415	390	312	380	436	269	1	-228	72	14 441
1992	31 603	42 185	9 251	6 600	3 031	2 307	1 707	441	503	391	404	485	294	1	-200	58	25 273
			8 647	5 551	2 850	2 410	1 757	421	395	316	386	442	272	1	-231	73	23 290
1993	29 372	35 193	8 180	5 836	2 680	2 040	1 509	390	444	346	358	428	260	1	-176	51	57 540
		29 252	6 856	4 401	2 260	1 911	1 393	334	313	250	306	350	216	1	-183	58	47 717
																Yhteensä (Mack)	131 816
																Yhteensä (Chain-Ladder)	118 419

Kuva 5.4. Solukohtaiset korvausvastuun estimaatit Chain-Ladder- (alemmat rivit) ja Mackin menetelmällä (ylemmät rivit). Lisäksi harmaalla tummennettu alue kuvaa sovitettua tilastoaineistoa. Estimaatit on laskettu huomioimalla inflaatio ja volyymi.

6. Yhteenveto

Tässä tutkielmassa on edellä esitetty erilaisia kollektiivisia korvausvastuun estimointimenetelmiä ja tutkittu eri korvausvastuun estimointimenetelmien "paremmuutta". Sana paremmuus on nyt asetettu lainausmerkkeihin, koska mitään kaiken kattavaa vertailuprotokollaa ei ole tässä yhteydessä käyty läpi ja koska edellä esitetyt tulokset on saatu pääosin käyttäen hyväksi simulointitekniikkaa. Näin ollen se, mitä on tehty, on muutaman erilaisen numeerisen vahinkomenoprosessiesimerkin antamien tulosten laskeminen. Toisaalta kyseiset esimerkit on kuitenkin pyritty valitsemaan siten, että ne antaisivat melko kattavan kuvan eri menetelmien käyttäytymisestä sekä suhteessa toisiin menetelmiin että suhteessa menetelmällä muissa esimerkeissä saatuihin tuloksiin.

Simulointimalleilla saatuja tuloksia katsomalla nähdään, että yleisesti ottaen parhaiten tässä yhteydessä toiminut menetelmä oli Mackin kredibiliteettiteoriaan perustunut menetelmä. Erityisen miellyttävä ominaisuus kyseisellä menetelmällä oli se, että tuntui toimivan hyvin sekä kehitysvirhettä että korvauskulua tarkasteltaessa. Myöskään se ei tuntunut olevan erityisen herkkä eri parametrien vaihteluille eikä myöskään liiemmin taustatekijöiden muutoksille – ainakaan, mikäli sen antamia tuloksia verrattiin muiden antamiin tuloksiin. Jos muita menetelmiä halutaan vielä asettaa paremmuusjärjestykseen simulointimalleilla saatujen tulosten perusteella, niin ehkä seuraavaksi tulisi empiirinen Bayes-menetelmä ja heti perässä peräkkäisten askelmien menetelmän modifikaatio. Kuten jo edellä mainittiin, niin vahinkoprosessin satunnaisen vaihtelun lisääminen paransi jälkimmäisen menetelmän tuloksia suhteessa muihin menetelmiin. Lisäksi myös peräkkäisten askelmien menetelmän modifikaation parantava vaikutus saatuihin tuloksiin verrattuna tavalliseen oli loppujen lopuksi melko selvä – noin 10–15 prosenttia. Nyt täytyy kuitenkin muistaa, että käytännössä on melko kyseenalaista voidaanko valmiiksi kehittyneitä kohortteja enää lisätä kehityskolmioon tai ainakaan tällöin ei enää voida olla varmoja menettelyn parantavasta vaikutuksesta. Myös mitä empiiriseen Bayes-menetelmään tulee, niin käytännössä järkevämpää olisi varmastikin tavallisen Bayes-menetelmän käyttö, sillä tällöin aktuaari voisi itse asettaa sopivat priorijakautumat ja näin ollen kehityskolmion merkitys pienentyisi.

Edellä ei ole vielä asetettu riskimaksutuloon perustuvaa menetelmää millekään sijalle eri menetelmien "paremmuus"-vertailussa. Tämä johtuu kyseisen menetelmän hieman erilaisesta luonteesta; kuten edellä mainittiin, niin menetelmä on enemmänkin jonkinlainen periaate kuin varsinainen menetelmä. Selvää kuitenkin on, että mikäli vuosittaiset riskimaksutulot kattavat hyvin tiettynä vuonna tapahtuneet vahingot, niin tietenkin myös korvausvastuun kehitysvirhe on melko pieni. Edellä olevassa pykälässä kuitenkin nähtiin ja muutenkin on itsestään selvää, että menetelmä on erittäin herkkä vuosittaisen riskimaksutulon riittävyydelle. Vuosittaisen korvauskulun vaihtelun kyseinen tapa pitää kylläkin pienenä, mutta valitettavasti tämä tapahtuu kehitysvirheen kustannuksella.

Edellä oleva "paremmuus"-järjestys on siis muodostettu simulointitekniikalla saaduista tuloksista. Kuten kuitenkin todellisuudesta poimitun esimerkin osoittamana nähdään, niin käytännössä näin ei välttämättä ole. Simulointimalleja tarkasteltaessa on pohjalla aina jokin matemaattinen malli ja vaikka malliin lisätäänkin kohinaa, niin se kuitenkin pohjautuu aina johonkin malliin. Käytännössä valittaessa käytettävää korvausvastuun estimointimenetelmää, niin luultavasti 95% aktuaareista valitsisi joko riskimaksutuloon perustuvan menetelmän tai Chain-Ladder-menetelmän tai näiden jonkinlaisen kombinaation. Tuskin kukaan valitsisi Empiiristä Bayes-menetelmää. Syy tähän on yksinkertaisesti niiden yksinkertaisuus. Ne ovat matemaattisesti yksinkertaisia ja näin ollen ei-matemaattisesti suuntautunutkin ymmärtää niiden periaatteen. Lisäksi vakuutuslajit kehittyvät koko ajan ja myös vakuutusehdot muuttuvat ja näin ollen myös korvauskäytäntö vaihtelee näiden mukana. Kuten luvusta 5 nähdään, niin vaikkakin kehittyneet menetelmät käyttävät pitkälle kehittyneitä teorioita hyväkseen, ne on kuitenkin jo niin monimutkaisia että niiden modifiointi kyseessä olevaan ongelmaan on melko monimutkaista. Lisäksi niiden antamien tuloksien analysointi vaatii aikaa, koska aina pitää kyetä vastaamaan, ovatko tulokset mielekkäitä, sellaisia, joita niiden intuitiivisesti olettaisi olevan, ja mikäli ne eivät ole, niin mikä on vialla?

Lopuksi voi mainita simulointimalleilla havaitun korvausvastuun estimointiin liittyvän virheen melko suuresta kokoluokasta; riippumatta millä tavalla virhettä verrataan, niin sen suhde sekä vuosittaiseen riskimaksutuloon että todellisiin korvausmenoihin on kuitenkin melko suuri. Riskimaksutuloon verrattuna virheen keskihajonta oli noin 10–30 prosenttia riskimaksutulosta ja todellisiin korvausmenoihin noin 7-15 prosenttia todellisista korvausmenoista. Myös vuosittaisen korvauskulunkin keskihajonta oli suhteellisen suuri – 5–35 prosenttia vuosittaisen riskimaksutulon suuruudesta. Käytännössä virheet ovat luultavimmin vielä suurempia mahdollisesti huomattavastikin suurempia.

Kirjallisuusluettelo

Beard R.E., Pentikäinen T., Pesonen E. (1984): *Risk Theory (3rd Edition)*. Chapman & Hall, London.

Daykin C.D., Pentikäinen T., Pesonen M. (1994): *Practical Risk Theory for Actuaries*. Chapman & Hall, London.

De Vylder F. (1982): Estimation of IBNR claims by credibility theory. *Insurance: Mathematics and Economics* Vol. 1, No 1, 35 - 40.

Goovaerts M.J., Kaas R., van Heerwaarden A.E., Bauwelinckx T. (1990): *Effective Actuarial Methods*. Elsevier, Amsterdam.

Hadidi N. (1985): A note on De Vylder's method of estimation of IBNR claims. *Insurance: Mathematics and Economics* Vol. 4, 263 - 266.

Jacobson J. (1992): *Vakuutusyhtiön korvausvastuusta*, Suomen Aktuaariyhdistys, Working Papers 37.

Kremer E. (1982): IBNR-claims and the two-way model of ANOVA. *Scandinavian Actuarial Journal*, No. 1, 47 - 55.

Mack T. (1990): Improved estimation of IBNR claims by credibility theory. *Insurance: Mathematics and Economics* Vol. 9, 51 - 57.

Lindley D.V., Smith A. (1972): Bayes estimates for the linear model. *Journal of Royal Statistic Society* Vol. 34, No. 1, 1–41.

Lindley D.V. (1971): The estimation of many parameters. *Foundations of statistical inference: A symposium*, ed. Godambe V.P., Sprott D.A., 435–455. Holt, Rinehart and Winston of Canada.

Lindley D.V. (1965): *Introduction to Probability & Statistics Part 2: Inference*. Cambridge University Press, London.

Pentikäinen T., Rantala J. (1992): A simulation procedure for comparing different claims reserving methods. *ASTIN Bulletin* Vol. 22, No. 2, 191-216.

Pentikäinen T., Rantala J.: Run-off risk as a part of claims fluctuation. *ASTIN Bulletin* Vol.16, No.2, 113-145.

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

Taylor G.C. (1986): *Claims Reserving in Non-life Insurance*. North-Holland, Amsterdam.

van Eeghen J. (1981): *Loss Reserving Methods. Surveys of Actuarial Studies 1*. Nationale-Nederlanden N.V., Rotterdam.

Verrall R.J. (1990): Bayes and empirical bayes estimation for the Chain Ladder model. *ASTIN Bulletin* Vol. 20, No. 2, 217 - 239.

Verrall R.J. (1991a): Chain Ladder and maximum likelihood. *Journal of the Institute of Actuaries* Vol. 118, No. 3, 489 - 499.

Verrall R.J. (1991b): On the estimation of reserves from loglinear models. *Insurance: Mathematics and Economics* Vol 10, 75–80.

Liite A.

Vastuuvakuutuksen (ei oikeusturva) henkilövahingojen selviäminen 1978-1993.

		Suoritusvuosi															
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	1978	1 765 813	1 476 891	493 882	359 156	363 443	514 437	231 311	147 295	115 948	150 774	127 136	123 337	97 174	120 276	323 641	111 306
	1979	1 465 569	1 727 270	580 628	1 110 089	715 755	206 393	42 322	104 367	65 334	111 641	56 918	27 905	27 889	33 217	76 384	
V	1980	1 334 738	2 577 519	1 014 848	1 300 223	910 782	147 610	569 554	75 151	121 663	26 627	77 218	84 083	46 137	13 531		
a	1981	1 998 234	2 024 656	1 445 107	1 150 328	944 694	625 977	112 617	190 061	110 448	87 247	108 498	32 915	34 305			
h	1982	1 747 029	1 946 027	638 712	878 291	241 543	288 489	155 701	88 681	84 184	58 059	239 294	18 812				
I	1983	1 774 921	3 438 718	1 129 425	837 596	1 038 425	573 114	135 304	353 406	145 447	48 272	176 723					
n	1984	2 560 609	2 418 191	1 347 410	905 106	1 274 658	269 052	1 032 540	759 233	621 671	187 224						
k	1985	2 075 068	2 794 589	1 420 640	1 575 627	827 847	363 268	760 547	292 500	338 760							
o	1986	2 053 671	2 639 207	1 383 097	936 092	521 641	307 375	659 652	477 599								
v	1987	2 816 566	3 872 758	1 827 784	1 074 519	710 122	786 985	839 796									
u	1988	3 580 227	5 450 722	2 740 418	1 365 742	1 192 535	291 898										
o	1989	3 292 960	4 126 095	1 685 539	1 341 842	927 713											
s	1990	4 877 233	5 278 792	2 359 220	2 246 571												
i	1991	3 738 708	3 900 590	2 718 477													
	1992	3 521 375	3 532 662														
	1993	3 254 740															

Liite B

Vastuuvakuutuksen (ei oikeusturva) omaisuusvahingojen selviäminen 1978-1993.

Suoritusvuosi

Vahinkovuosi	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1978	8 045 357	5 740 226	1 092 550	651 205	582 654	1 666 272	437 551	454 157	282 084	-11 802	30 552	96 116	0	0	1 179	18 248
1979	9 901 774	7 092 241	1 706 524	857 313	945 409	5 186 894	1 317 192	97 517	70 145	10 793 761	33 340	3 977 102	380 047	3 678	-133 108	
1980	11 673 448	10 862 033	2 838 018	1 021 381	1 116 888	790 948	759 610	260 396	122 692	-918	556 470	169 139	-7 269	-2 733		
1981	16 676 112	15 006 316	3 953 166	1 826 656	1 083 679	891 923	232 610	255 293	147 471	54 688	133 754	2 997	3 325			
1982	17 761 428	16 170 813	4 425 840	1 507 749	883 920	1 369 834	1 654 356	132 081	95 313	91 577	91 958	26 833				
1983	20 483 687	21 926 295	6 167 214	1 983 431	1 316 032	1 398 830	8 338 770	127 799	-17 267	44 268	-2 114					
1984	30 634 880	28 838 593	8 503 410	5 039 878	2 470 093	1 693 726	3 263 813	-45 545	517 747	321 209						
1985	28 981 958	33 783 498	6 021 309	12 353 534	4 498 955	731 144	378 578	341 846	389 964							
1986	34 567 207	29 001 436	8 070 510	2 972 281	1 063 980	911 289	647 730	104 098								
1987	44 912 838	46 652 958	8 907 756	10 899 553	940 055	1 806 209	145 890									
1988	41 012 939	30 882 533	5 820 347	4 756 260	3 760 242	1 469 475										
1989	33 039 712	35 563 270	8 618 141	11 570 562	4 601 180											
1990	46 965 513	97 416 105	19 102 263	3 985 963												
1991	36 757 934	34 795 735	9 573 225													
1992	31 603 377	42 184 502														
1993	29 371 888															