

## 5. Yksilöllinen tariffointi

Tariffitekiäjien avulla vakuutusmaksut pystytään mitoit-  
tamaan oikeudenmukaisesti ainakin jollain tarkkuudella.  
Tariffiluokan kohteet asetetaan tällöin samanlaisiksi:  
riskimaksu on sama kaikilla luokkaan kuuluvilla koh-  
teilla. Selvää kuitenkin on, että tariffitekiäjät eivät  
kuvaa riskiä täsmällisesti.

Tarkastellaan esimerkkinä autovakuutusta. Luonnolliselta  
luntau, että esimerkiksi ajoneuvon 'koti-paikka' vaikut-  
taa riskin suuruuteen. Tämä on helposti saatavilla  
oleva tieto ja siten varteenotettava tariffitekijähdokas.  
Auton omistajan ajotapa vaikuttaa ilmeisesti myös  
kolariin riskiin. Tätä on kuitenkin vaikea mitata, eikä  
se siksi sovi tariffitekijäksi. Tietoa ajotavasta saa-  
daan kuitenkin epäsuorasti seuraamalla auton omistajan  
vahinkohistoriaa. Yksilöllisen tariffoinnin ideana on-  
kin juuri ohjata vakuutusmaksua nojautuen kohteeseen  
omaan historiaan.

Ilmeisesti tariffin oikeudenmukaisuutta pystyt-  
tään parantamaan yksilöllisten menetelmien avulla.  
Toisaalta vakuutusmaksu ei saisi vaihdella tur-  
han paljon vuodesta toiseen. Tämä rajoittaa vahinko-  
historian käyttämällisyyttä tariffoinnissa.

Tarhusto tariffeilla on käytännössä aina oma  
merkityksensä, vaikka sovellettaisiinkin yksilöllisiä  
menetelmiä. Tyypillisesti uusi vakuutettava kohde  
hinnoitellaan nojautuen tarhusto tariffiin. Laajemmin,  
usein maksu määräytyy kompromissina tarhusto maksusta  
ja omasta vahinkohistoriasta.

Vahinkohistorias voidaan hyödyntää monin tavoin tariffoinnissa. Seuraavassa tarkastellaan kolmea lähestymistapaa:

- krediittily - teoria
- eksponentiaalinen tasointus
- bonusjärjestelmät.

## 5.1. Credibility-teoriaa

Tarkastellaan vakuutuskaikaa tai taikittimokkaa, joka muodostuu samankaltaisista muuttajista kuitenkin samanlaisista kohteista. Credibility-teorian lähtökohdaksi määritellään riskikollektiivin käsite:

1) Kukaan yksittäinen kohde kuvataan muuttajilla

$$v, X_1, \dots, X_n, \text{ missä}$$

$v$  = riskiparametri, ei havaittu,

$X_1, \dots, X_n$  = havaitut kohteen vahinkomäärät vuosina  $1, \dots, n$ .

Ehdolla  $v = v$ , vahinkomäärät  $X_1, \dots, X_n$  ovat riipp.

2) Riskikollektiivi

$\Theta$  = mahdollisten  $v$ :n arvojen joukko,

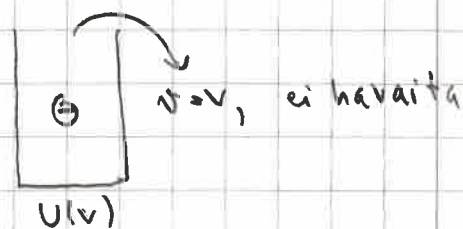
$U$  = takennefunktio,  $v$ :n jakauma  $\Theta$ :ssa,

$F(v) = v$ :n arvoa  $v$  vastaava vahinkomäärän kertymäfunktio,  $v \in \Theta$ .

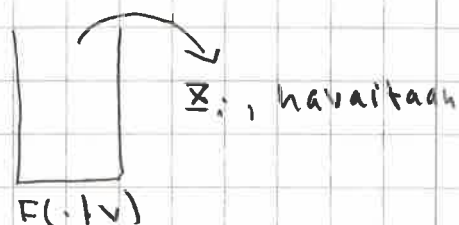
Riskiparametri voisi olla esimerkiksi kunkin kohteen vahinkomäärän odotusarvo eli riskimaksu. Mallissa tähän suhtaudutaan satunnaismuuttujana.

Samaava muuhavainnollistus antaa induktiivisen kuvan mallista. Kullekin kohteelle

- 1) alussa perimitaan  $v = v$  jakaumasta  $U$



- 2) Kunakin vuonna  $i$  perimitään  $X_i$  jakaumasta  $F(\cdot|v)$



Yllä  $F(x|v)$  on  $X_i$ -in ehdollinen kertymäfunktio,

$$F(x|v) = P(X_i \leq x | v = v), \quad x \in \mathbb{R}$$

Kiinteälle kohteelle pätee

$$P(X_1 \leq x_1, \dots, X_n \leq x_n | v = v) = F(x_1|v) \cdots F(x_n|v).$$

Rakennefunktio  $U$  kuvaa kollektiivista heterogeenisuutta. Voidaan ajatella, että  $dU(v)$  ilmaisee, millä painolla riskiparametrin arvo  $v$  esiintyy vakuumuokannassa. Umpimähtään kannasta valitulle rakennetulle pätee

$$\begin{aligned} P(X_1 \leq x) &= \int_0^1 F(x|v) dU(v) \\ &= F(x). \end{aligned}$$



Merkitämme ehdollista odotusarvoa ja varianssia symbolilla  $\mu(v)$  ja  $\sigma^2(v)$ ,

$$\mu(v) = \mathbb{E}(X_i | U=v) = \int_{\mathbb{R}} x dF(x|v),$$

$$\begin{aligned} \sigma^2(v) &= \mathbb{E}((X_i - \mu(v))^2 | U=v) \\ &= \int_{\mathbb{R}} (x - \mu(v))^2 dF(x|v). \end{aligned}$$

Umpimätkään valitulle kohteelle merkitään

$$\begin{aligned} \mu &= \mathbb{E}(X_i) = \mathbb{E}(\mathbb{E}(X_i | U)) \\ &= \mathbb{E}(\mu(U)) = \int_{\Theta} \mu(v) dU(v), \end{aligned}$$

$$\begin{aligned} \sigma^2 &= \text{Var}(X_i) = \mathbb{E}(\sigma^2(U)) + \text{Var}(\mu(U)) \\ &= \int_{\Theta} \sigma^2(v) dU(v) + \int_{\Theta} (\mu(v) - \mu)^2 dU(v). \end{aligned}$$

Yhteisesti

$$\mu = \int_{\mathbb{R}} x dF(x), \quad \sigma^2 = \int_{\mathbb{R}} (x - \mu)^2 dF(x).$$

5.6.1.

### 5.1.1. Yhden vakunteen crediitilidyy-maattu

On luonnollista tarkastella esimerkiksi taajuu-  
luokkaa ja mallintaa vakunteen heterogeenisyyttä  
riisiparametrien avulla. Tarkastellaan tässä kappa-  
leessa asiaa kuitenkin vain yhden vakunteen  
näkökulmasta.

Tavoitteena on minimoida keskinäisriippuvuus (5.2).  
Ainoastaan vakunteen oma vahinkohistoria siis on  
käytettävissä.

Pääongelmana on estimoida kohteen hiiliparametri, kun vuotuiset vahinkomäärät  $X_1, \dots, X_n$  on havaittu, tai pikemminkin tähän liittyvä vertaaminen. Luonnollinen estimaattori on

$$R_0 = E(\mu(v) \mid X_1, \dots, X_n).$$

Tätä kutsutaan usein Bayes-estimaattoriksi tai Bayes-maksuksi. Tunnetusti  $R_0$  minimoi keskineliöpoikkeaman

$$\begin{aligned} & E\left([\mu(v) - f(X_1, \dots, X_n)]^2\right) \\ &= \int_{\Theta \times \mathbb{R}^n} [\mu(v) - f(x_1, \dots, x_n)]^2 dF(x_1|v) \dots dF(x_n|v) dU(v) \end{aligned}$$

yli kaikkien mitallisten kuvausten  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ . Yksinkertaisuuden vuoksi valitaan keskineliöpoikkeaman mielessä paras lineaarinen  $f = f_L$  (aikeastaan affiini),

$$f_L(x_1, \dots, x_n) = a + b_1 x_1 + \dots + b_n x_n.$$

Parametrit  $a, b_1, \dots, b_n$  määrittävät siten, että keskineliöpoikkeama

$$(5.2) \quad E\left([\mu(v) - f_L(X_1, \dots, X_n)]^2\right)$$

minimoidaan. Optimaalisilla parametrien arvoilla määrittyvä maksu kutsutaan credibility-maksuksi.

Lause 5.1. Olkoon  $z^2 \in (0, \infty)$ . Silloin (5.2) minimoituu, kun

$$a = (1-z)\mu,$$

$$b_1 = \dots = b_n = z/n,$$

missä

$$z = \frac{ns}{ns+m},$$

$$m = \mathbb{E}(z^2(\nu)),$$

$$s = \text{Var}(\mu(\nu)).$$

Todistus. Tarkastellaan aluksi joitakin malliin liittyviä momenteja. Olkoon  $i \neq j$ . Silloin

$$\begin{aligned} (5.4) \quad \mathbb{E}(Z_i Z_j) &= \mathbb{E}(\mathbb{E}(Z_i Z_j | \nu)) \\ &= \mathbb{E}(\mathbb{E}(Z_i | \nu) \mathbb{E}(Z_j | \nu)) = \mathbb{E}(\mu(\nu)^2) \end{aligned}$$

ja

$$\begin{aligned} (5.5) \quad \text{Cov}(Z_i, Z_j) &= \mathbb{E}(Z_i Z_j) - \mathbb{E}(Z_i) \mathbb{E}(Z_j) \\ &= \mathbb{E}(\mu(\nu)^2) - \mu^2 = \text{Var}(\mu(\nu)) = s. \end{aligned}$$

Samaan tapaan nähdään, että

$$\begin{aligned} (5.6) \quad \mathbb{E}(Z_i^2) &= \text{Var}(Z_i) + \mu^2 \\ &= m + s + \mu^2. \end{aligned}$$



Olkoon annesi  $m > 0$ . Merkitään

$$g(a, b_1, \dots, b_n) = E([ \mu(\omega) - (a + b_1 \bar{X}_1 + \dots + b_n \bar{X}_n) ]^2).$$

Tavoitteena on siis minimoida  $g$  yli joukon

$$(a, b_1, \dots, b_n) \in \mathbb{R}^{n+1}.$$

Olkoon  $\mu > 0$  summi. Silloin

$$\begin{aligned} g(\mu a, \mu b_1, \dots, \mu b_n) &= \mu^2 E([ \frac{\mu(\omega)}{\mu} - (a + b_1 \bar{X}_1 + \dots + b_n \bar{X}_n) ]^2) \\ &= \mu^2 E([ a + b_1 \bar{X}_1 + \dots + b_n \bar{X}_n ]^2) + o(\mu) \mu^2, \quad \mu \rightarrow \infty. \end{aligned}$$

Todetaan, että  $g(\mu a, \mu b_1, \dots, \mu b_n) \rightarrow \infty$  tasaisesti

joukossa  $\|(a, b_1, \dots, b_n)\| = 1$ . Nimitetään

$E([ a + b_1 \bar{X}_1 + \dots + b_n \bar{X}_n ]^2)$  saavuttaa miniminsä yksikkö-

ympyrällä ja minimiarvo on aidosti positiivinen. Siis

$g$  saavuttaa minimiarvonsa eräessä pisteessä. Tässä

totentuvat välttämättömät optimalityshdot,

$$\begin{cases} \frac{\partial}{\partial a} g = -2 E(\mu(\omega) - a - b_1 \bar{X}_1 - \dots - b_n \bar{X}_n) = 0, \\ \frac{\partial}{\partial b_i} g = -2 E(\bar{X}_i (\mu(\omega) - (a + b_1 \bar{X}_1 + \dots + b_n \bar{X}_n))) = 0, \end{cases}$$

$$i = 1, \dots, n.$$

Tarkastellaan osittaisderivaattoja  $b$ -muuttujien suhteen. Saadaan

$$\mathbb{E}(Z_i; \mu(\omega)) = \mathbb{E}(\mathbb{E}(Z_i; \mu(\omega) | \omega)) = \mathbb{E}(\mu(\omega)^2).$$

Optimaalisuusehto ja kaavat (5.4) - (5.6) antavat

$$\begin{aligned} \mathbb{E}(\mu(\omega)^2) - a\mu - \mathbb{E}(\mu(\omega)^2) \sum_{j=1}^n b_j - (m+s+\mu^2) b_i \\ = \mathbb{E}(\mu(\omega)^2) - a\mu - \mathbb{E}(\mu(\omega)^2) \sum_{j=1}^n b_j - m b_i = 0. \end{aligned}$$

Vähentämällä saatuja yhtälöitä pareittain toisistaan nähdään, että optimipisteessä

$$b_1 = \dots = b_n = b.$$

Optimointitehtävä on siis saatu muotoon: määritä  $a$  ja  $b$  siten, että

$$\begin{cases} a + b\mu n = \mu, \\ a\mu + b(n\mathbb{E}(\mu(\omega)^2) + m) = \mathbb{E}(\mu(\omega)^2). \end{cases}$$

Ratkaisu on

$$b = \frac{\mathbb{E}(\mu(\omega)^2) - \mu^2}{n\mathbb{E}(\mu(\omega)^2) + m - n\mu^2} = \frac{s}{m + ns} = \frac{s}{s + m},$$

$$a = \mu(1 - bn) = \frac{m\mu}{m + ns} = (1 - z)\mu,$$

Lause on siis todistettu, jos  $m > 0$ . Jos taas  $m = 0$ , niin  $IP(\delta^2(w) = 0) = 1$  ja

$$IP(\bar{X}_n = \mu(w)) = E \left( \underbrace{IP(\bar{X}_n = \mu(w) | w)}_{=1 \text{ m.v.}} \right) = 1.$$

Sis keskineliöpoikkeama (5.2) minimoituu (ja on  $= 0$ ), kun valitaan  $z = 1$ . Lauseen väite pätee siis myös, kun  $m = 0$ .  $\square$

Odotusarvoa  $\mu$  kutsutaan usein maximalimaksuksi. Lauseen 5.1 nojalla credibility-määrä on muotoa

$$(1-z)\mu + z\bar{\mu},$$

missä  $\bar{\mu} = (\bar{X}_1 + \dots + \bar{X}_n)/n$  on kohteen vahinkomäärien otoskeskiarvo. Kertoimetta  $z$  kutsutaan credibility-kertoimeksi. Nähdään, että aina  $0 \leq z \leq 1$  ja että  $z$  kasvaa kohti ykköstä, kun havaintovuosien määrä  $n$  kasvaa kohti ääretöntä.

Crediteilily - kurtinmen muoto on mielekäs seuraavasti:

- 1) Jos kollektiivi on hyvin epähomogeeninen eli  $\mu(\omega)$  vaihtelee suuresti kohteesta toiseen, on intuitiivisesti perusteltua antaa painoa kohteen omalle vahinkohistorialle. Tässä tilanteessa

$$s = \text{Var } \mu(\omega)$$

on suuri ja siis myös

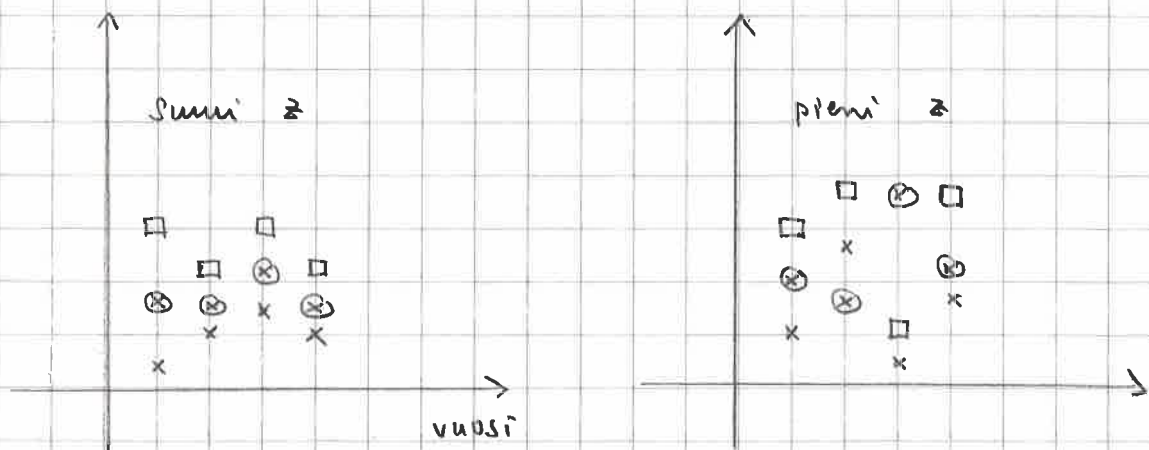
$$z = \frac{n}{n+m/s}$$

on suuri luden pitääkin.

- 2) Jos yksittäisen kohteen vahinkomäärän vuotuinen heilahtelu on vähäistä, on myös perusteltua antaa painoa kohteen omalle vahinkohistorialle. Nyt

$$m = E(\delta^2(\omega))$$

on pieni ja siis jälleen  $z$  on suuri.



Vahinkomäärät kohteille  $\square$ ,  $\otimes$  ja  $x$ .



Esimerkki 5.1. Olkoon  $\underline{X}$ ,:n ehdollisella ketymä-

lunktiolla  $F(\cdot|v)$  tiheysfunktio,

$$f(x|v) = \frac{\partial}{\partial x} F(x|v) = \frac{a(x)e^{-vx}}{g(v)}, \quad x \in A = (\underline{x}, \bar{x}),$$

missä  $v \in \Theta = (\underline{v}, \bar{v})$  ja  $a(x) \geq 0, \forall x$ . Päätepisteiden  $\underline{x}, \bar{x}, \underline{v}$  ja  $\bar{v}$  ei tarvitse olla äärellisiä. Luonnollisesti oletetaan, että

$$g(v) = \int_A a(x) e^{-vx} dx \in (0, \infty), \quad \forall v \in \Theta.$$

Olkoon myös rakennefunktiolla  $U$  tiheysfunktio,

$$u(v) = U'(v) = \frac{e^{-\alpha v}}{h g(v)^\beta}, \quad v \in \Theta,$$

missä  $\alpha, \beta > 0$  ovat vakioita ja

$$h = \int_{\Theta} \frac{e^{-\alpha v}}{g(v)^\beta} dv \in (0, \infty).$$

Oletetaan vielä, että  $\forall x \in A$ ,

$$u(\underline{v} + |x|) = u(\bar{v} - |x|) = 0,$$

missä

$$u(v|x) = \frac{u(v) f(x|v)}{f(x)}$$

ja

$$f(x) = \int_{\Theta} u(v) f(x|v) dv \in (0, \infty).$$

Bayesin kaavan nojalla  $u(\cdot|x)$  on  $\Theta$ :n ehdollinen tiheysfunktio ehdolla  $\underline{X}_1 = x$ . Kokonaistodennäköisyysteoreeman nojalla  $f$  on  $\underline{X}_1$ :n tiheysfunktio.

olleoon  $\mu = \mathbb{E}(\bar{X}_n)$  äärellinen.

Määritään vakuummaksun Bayes-estimaattori

$$R_0 = \mathbb{E}(\mu(v) | \bar{X}_n)$$

olettaen siis, että  $n=1$ . Näytetään, että

$$(5.8) \quad R_0 = \frac{\alpha}{\beta+1} + \frac{\bar{X}_1}{\beta+1}.$$

Koska tämä on lineaarinen  $\bar{X}_1$ :n suhteen, on se minimaalisuusominaisuutensa perusteella myös credibi-  
lity-maksu (edellyttäen, että lauseen 5.1 oletukset  
ovat täytetyt). Erikyisesti credibi-  
lity-maksun on

$$Z = \frac{1}{\beta+1}.$$

Todistetaan (5.8). Ehdollinen odotusarvo voidaan  
määrittää käyttäen ehdollista tiheysfunktioita, joten

$$\mathbb{E}(\mu(v) | \bar{X}_1 = x) = \int_{\Theta} \mu(v) u(v|x) dv.$$

Todetaan, että  $\forall v \in \Theta, x \in A$ ,

$$\frac{\partial}{\partial v} [u(v|x)] = \frac{a(x)}{h(x)} \frac{\partial}{\partial v} [g(v)^{-\beta-1} e^{-(\alpha+x)v}]$$

$$= \frac{a(x)}{h(x)} \left[ -(\beta+1)g'(v)g(v)^{-\beta-2} e^{-(\alpha+x)v} \right. \\ \left. - (\alpha+x)g(v)^{-\beta-1} e^{-(\alpha+x)v} \right]$$

$$= u(v|x) \left[ -\frac{(\beta+1)g'(v)}{g(v)} - (\alpha+x) \right].$$

Toisaalta

$$g'(v) = - \int_A x a(x) e^{-vx} dx$$

$$= -g(v) \int_A x f(x|v) dx = -g(v) \mu(v).$$

Siksi pä

$$\frac{\partial}{\partial v} [u(v|x)] = u(v|x) [(\beta+1)\mu(v) - (\alpha+x)].$$

Integroimalla puolittain saadaan

$$0 = u(\bar{v}-|x) - u(\underline{v}+|x)$$

$$= (\beta+1) \int_{\underline{v}}^{\bar{v}} u(v|x) \mu(v) dv - (\alpha+x) \int_{\underline{v}}^{\bar{v}} u(v|x) dv$$

$$= (\beta+1) E(\mu(v) | \bar{X}_1 = x) - (\alpha+x).$$

Tästä seuraa (5.8).



Esimerkki 5.2. Tarkastellaan esimerkitapausta, jossa

$$A = (0, \infty), \quad a(x) = 1, \quad \forall x \in A \quad \text{ja} \quad \Theta = (0, \infty).$$

Nyt

$$g(v) = \int_0^{\infty} e^{-vx} dx = \frac{1}{v},$$

$$f(x|v) = ve^{-vx},$$

$$h = \int_0^{\infty} v^{\beta} e^{-\alpha v} dv \in (0, \infty),$$

$$u(v) = h^{-1} v^{\beta} e^{-\alpha v},$$

$$f(x) = \int_0^{\infty} h^{-1} v^{\beta+1} e^{-(\alpha+x)v} dv \in (0, \infty),$$

$$\mu = \mathbb{E}(\bar{X}_n) = \int_0^{\infty} x f(x) dx \in (0, \infty),$$

$$u(v|x) = f(x)^{-1} h^{-1} v^{\beta+1} e^{-(\alpha+x)v},$$

$$u(0|x) = u(\infty|x) = 0, \quad \forall x \in A.$$

Sis  $\bar{X}_n$ :n ehdollinen jakauma ehdolla  $\vartheta = v$  on eksponenttijakauma parametreilla  $v$ . Riskiparametrilla  $\vartheta$  on gamma- $(\beta+1, \alpha)$ -jakauma. Esimerkin 5.1 nojalla

$$\mathbb{E}(\mu(v) | \bar{X}_n = x) = \frac{\alpha}{\beta+1} + \frac{x}{\beta+1}.$$

Todennetaan suoraan laskemalla, että  $z = \frac{1}{\beta+1}$  on lauseen 5.1 esityksen mukainen credibility-kennän.



Määritetään siis  $\mu$ ,  $m$  ja  $s$ :

$$\mu(v) = \int_0^{\infty} x v e^{-vx} dx = \frac{1}{v},$$

$$\zeta^2(v) = \int_0^{\infty} (x - \mu(v))^2 v e^{-vx} dx = \frac{1}{v^2},$$

$$m = \mathbb{E}(\zeta^2(v)) = \int_0^{\infty} v^{-2} u(v) dv$$

$$= \int_0^{\infty} v^{-2} x^{\beta+1} v^{\beta} e^{-xv} / \Gamma(\beta+1) dv$$

$$= \frac{x^2}{(\beta-1)\beta} \int_0^{\infty} x^{\beta-1} v^{\beta-2} e^{-xv} / \Gamma(\beta-1) dv = \frac{x^2}{(\beta-1)\beta},$$

$$\mu = \mathbb{E}(\mu(v)) = \int_0^{\infty} v^{-1} x^{\beta+1} v^{\beta} e^{-xv} / \Gamma(\beta+1) dv = \frac{x}{\beta}.$$

$$s = \text{var}(\mu(v)) = \int_0^{\infty} \mu(v)^2 u(v) dv - \mu^2$$

$$= \frac{x^2}{(\beta-1)\beta} - \frac{x^2}{\beta^2}.$$

Siis

$$z = \frac{s}{s+m} = \frac{1}{\beta+1}.$$

Tulee myös olla

$$(1-z)\mu = \frac{x}{\beta+1}.$$

Näin on, koska  $\mu = \frac{x}{\beta}$ .

Lauseen 5.1 uedileily-kaava on ongelmallinen, jos suuraahingot ovat mahdollisia. Mekaaninen kaavan soveltaminen johtaa tällöin toisinaan summiin vakuummaksun vuosivaihteluihin. Tämä ongelma voidaan ratkaista esimerkiksi katkaisemalla vuosuiset kokonaisvahinkomäärät. Oikean maksutason saavuttamiseksi katkaisut täytyy kompensoida esimerkiksi kollektiivisella lisämaksulla.

Toinen ongelma liittyy rakennefunktion  $U$  ja ehdollisten jakaumien  $F(\cdot|v)$  valintaan ja estimointiin. Lauseesta 5.1 riittää kuitenkin tun-  
 tea parametrit  $\mu$ ,  $m$  ja  $s$ . Näille voidaan esittää harkittomat estimaattorit perustuen usean vuoden aineistoon.

### 5.1.2. Taisteluvoiman krediittikysymykset

Tarkastellaan taisteluvoimaa, jossa on  $N$  kappaletta samanlaisista yksiköistä. Eriksittain ajatellaan, että yksiköiden on kaikkien sama a priori (enemmän vahinkomäärien havaitsemista), käsitystä korjataan havaintojen perusteella.

Oletetaan yksiköiden vuoksi, että tarkastelu-  
hetkellä kustakin yksiköstä on käytettävissä  $n$   
kappaletta vuotuisia vahinkohavaintoja. Yksiköiden  
havainnot olkoot

$$X_{i1}, \dots, X_{in}, \quad i=1, \dots, N.$$

Yksiköiden  $i$  liitetään riskiparametri  $\theta_i$ .  
Merkitään lyhyesti

$$Y_i = (X_{i1}, \dots, X_{in}).$$

Oletetaan, että vektorit

$$(v_1, Y_1), \dots, (v_N, Y_N), (v, Y)$$

ovat riippumattomia ja samoin jakautuneita. Muuten  
tehdään samat oletukset kuin kohdassa 5.1.1  
(merkitään lyhyesti vastaten generisistä vektoreita  $(v, Y)$ ).

Luonnollinen krediilitely-matka on muotoa

$$a + \sum_{i=1}^N \sum_{j=1}^n b_{ij} X_{ij}$$

kerroimet  $a, b_{ij}, i=1, \dots, N, j=1, \dots, n$ , määritään siten, että keskineliöpoikkeama

$$(5.8.1) \quad E \left( \left[ \mu(\omega_k) - \left( a + \sum_{i=1}^N \sum_{j=1}^n b_{ij} X_{ij} \right) \right]^2 \right)$$

minimoidaan,  $k=1, \dots, N$ .

Lause 5.1.1. Keskineliöpoikkeama (5.8.1) minimoidaan, kun  $b_{ij} = 0, \forall i \neq k, j=1, \dots, n$ , ja

$$a = (1-z)\mu, \quad b_{k1} = \dots = b_{kn} = z/n,$$

missä  $\mu$  ja  $z$  ovat kuten lauseessa 5.1.



Todistus. Jos kertoimien ovat kiinteitä, niin riippumattomuusoletuksen nojalla (5.8.1) on vähintään

$$\begin{aligned}
 & \text{Var} \left( \mu(x_k) - \left( a + \sum_{i=1}^N \sum_{j=1}^n b_{ij} x_{ij} \right) \right) \\
 & + \left[ \mathbb{E}(\mu(x_k)) - a - \sum_{i=1}^N \sum_{j=1}^n b_{ij} \mu \right]^2 \\
 & = \text{Var}(\mu(x_k) - \sum_{j=1}^n b_{kj} x_{kj}) \\
 & + \text{Var} \left( \sum_{i \neq k} \sum_{j=1}^n b_{ij} x_{ij} \right) \\
 & + \left[ \mathbb{E}(\mu(x_k)) - a - \sum_{i=1}^N \sum_{j=1}^n b_{ij} \mu \right]^2 \\
 & \geq \text{Var}(\mu(x_k) - \sum_{j=1}^n b_{kj} x_{kj}).
 \end{aligned}$$

Yhtäsuuruus pätee, kun

$$\begin{cases} b_{ij} = 0, & \forall i \neq k, j = 1, \dots, n, \\ a = \mu - \sum_{j=1}^n b_{kj} \mu. \end{cases}$$

Nähdään, että on yhtäpitävää minimoida keskenään riippumattomien

$$\mathbb{E} \left( \left[ \mu(x_k) - \left( a + \sum_{j=1}^n b_{kj} x_{kj} \right) \right]^2 \right).$$

Tämän ratkaisu löydetään lauseesta 5.1.  $\square$