Suppea SHV-harjoitustyö

Kokonaisvahinkojakauma

Jani Pajukangas

10.11.2016

Abstract

Individual and aggregate loss models are important concepts in non-life insurance mathematics. This work presents some basic ideas with compound model framework. Compound model allows the use of analytical model for aggregate losses. Compound Poisson model has computational benefits and is in favor of actuaries. Compound Poisson-gamma-model is part of the Tweedie family with parameter p between 1 and 2. This model has many useful applications. Although it does not have closed form for its CDF it has the benefit with connectedness to the risk theory.

Typically claims frequency is modelled by Poisson distribution and claim severity by gamma distribution. Sometimes claims frequency is overdispersed and the Poisson-Tweedie models are called upon. Important example is Poisson-Inverse Gaussian distribution. Tweedie models and Poisson-Tweedie models have ubiquitous presence among actuarial applications. Compound Poisson-gamma has also useful applications in several branch of science. Tweedie convergence theorem shows that for many probability distributions Tweedie distributions act as a foci of convergence. Because Tweedie models are a subset of exponential dispersion models the results that applies to exponential dispersion models are also applicable to Tweedie models.

Some crucial methods regarding calculation of collective risk model is presented. Convolution can be calculated by Fourier transform or more exactly by Fast Fourier Transform (FFT). It has been shown that this is faster option compared to the Panjer's recursion.

Sisältö

Merkit ja lyhenteet

Merkki	**Kuvaus**
$\mathcal{C}(t)$	Kumulantit generoiva funktio
$F_S(x)$	Satunnaismuuttujaan S liittyvä kertymäfunktio
$f_S(x)$	Satunnaismuuttujaan S liittyvä tiheysfunktio
$F_X^{*n}(x)$	Satunnaismuuttujaan X liittyvä n-kertainen konvoluutio
$\mathscr{F}(t)$	Fourier-muunnos
$\mathcal{G}(t)$	Todennäköisyydet generoiva funktio
$\mathcal{K}(\cdot)$	Kumulanttifunktio
$\mathscr{L}(t)$	Laplace-muunnos
$\mathcal{M}(t)$	Momentit generoiva funktio
$\mathbf{M}$	Matriisi
p_x	Vahinkotiheyden tiheysfunktio
$*$	Konvoluutio
$\sim$	Samoin jakautuneita
$\propto$	Verrannollinen

Lyhenne	**Kuvaus**
DÄJ	Diskreetti äärettömästi jaettavissa oleva
EDM	Eksponentiaaliset dispersiomallit
FFT	Nopea Fourier-muunnos (Fast Fourier Transform)
IBNR	Kollektiivivaraus (Incurred But Not Reported)
MC	Monte Carlo -simulointimenetelmä
ÄJ	Äärettömästi jaettavissa oleva

1 Johdanto

Työssä käydään aluksi läpi vakuutuskohtaisen ja kollektiivisen riskimallin idea luvussa (2). Vahinkotiheys eli vahinkojen lukumäärä aikayksikköä kohti ja vahingon suuruus mallinnetaan erikseen. Luvussa (3) käydään läpi sekajakauma, painotettu jakauma, yhdistetty jakauma ja yhdistetty painotettu jakauma. Luvussa (4) esitetään jakaumaperheet (eksponentiaaliset dispersiomallit, Tweedie-mallit ja Tweedie-Poisson-mallit) ja näiden ominaisuuksia. Luvussa (5) esitetään tunnettu menetelmä riippumattomien vahinkojen kokonaisvahinkomäärän laskemiseksi konvoluutiolla. Konvoluutio voidaan laskea esimerkiksi Panjerin rekursiolla tai Fourier-muunnoksen avulla. Työssä esitetään perusteluja Fourier-muunnoksen käytölle. Diskreetin Fourier-muunnoksen laskeva nopea Fourier-muunnos vaatii alkutietona vahingon suuruuden tiheysjakauman ja vahinkotiheyden sekajakauman sekä vahinkotiheyden todennäköisyydet generoivan funktion. Tiheysjakauman tuottaminen vaatii joko parametrisen jakauman tai empiirisen jakauman diskretisoinnin. Käydään läpi yleisimmin käytössä olevia jakaumia. Koska kyseessä on suppea työ, on tässä tehty merkittäviä rajauksia. Tehtyjä rajauksia on luetteloitu liitteessä (A).

2 Riskimallit

2.1 Vahinkojen ennustaminen

Vahinkomenoa ennustettaessa on tehtävä arvioita lopullisesta vahinkomenosta eli huomioitava IBNR-vahingot, huomioitava vahinkoinflaatio sekä muutokset vakuutusvuosissa ja muut mahdolliset vahinkomenoon liittyvät korjaukset. Aikadimensio jätetään muutoin huomiotta eli toisin kuin yleensä henkivakuutuksessa, tässä tarkastellaan vain yhtä ennustettavaa vuotta. Lisäksi vahinkojakauman odotusarvon ja varianssin lisäksi ollaan kiinnostuneita esim. Value-at-Risk arvosta. Tätä tarkoitusta varten on muodostettava ennustettavan vahinkojakauman tiheys- tai kertymäfunktio. [17, s. 3-7]

2.2 Vakuutuskohtainen riskimalli

Vakuutuskohtaisessa riskimallissa vakuutusyhtiön jokaista vakuutusta tarkastellaan erikseen ja kokonaisvahinkomäärä S on kyseisissä vakuutuksissa tapahtuvien korvausten summa. Malli voidaan esittää muodossa

$$S^{Ind} = X_1 + \cdots + X_n, \tag{2.1}$$

missä n on vakuutusten lukumäärä ja X on vahingon suuruusjakauma nollatodennäköisyys huomioiden. Muuttujat X_i ovat keskenään riippumattomia. Käytännössä simulaatiossa suuri osa termeistä on nollia vahingottomuuden vuoksi. Kokonaisvahinkomäärä voidaan vakuutuksille sattuneista vahingoista laskea konvoluutiolla. Tämä on kuitenkin kohtalaisen työlästä ja isolle vakuutuskannalle hidasta jopa nykyisille tietokoneille. On kuitenkin kehitetty menetelmiä, joilla voidaan S^{ind}:ä approksimoida. Yksinkertaisimmasta päästä on normaaliapproksimaatio, joka huomioi vain kaksi ensimmäistä momenttia. Vakuutustarkasteluissa tärkeä häntätodennäköisyys tulee paremmin huomioitua NP-approksimaatiossa (Normal Power), joka huomioi kolmannen momentin.

Vakuutuskohtaisen riskimallin yksittäistä satunnaismuuttujaa voidaan mallintaa soveltaen siihen yhdistettyä Poisson-gamma-jakaumaa, joka yhdistää sekä Poisson-jakauman että gamma-jakauman ominaisuuksia. Pisteessä nolla jakauma antaa Poisson-jakauman arvon $e^{-\lambda}$ ja positiivisessa alueessa gamma-jakauman arvoja. Vakuutuskohtaiselle mallille on kehitetty rekursiota helpottava malli, joka on saanut nimensä De Prilin mukaan. Fourier-muunnoksella ja riittävällä suuruusjakauman diskretisoinnilla saadaan arvio jakaumalle. Lisäksi on olemassa tärkeä approksimaatio vakuutuskohtaiselle mallille, joka on kollektiivinen riskimalli. [12, s. 17-33]

2.3 Kollektiivinen riskimalli

Kollektiivisessa mallissa dimensioina ovat vahingon suuruus ja vahinkojen lukumäärä. Puhuttaessa aikavälillä tapahtuneiden vahinkojen lukumäärästä puhutaan myös vahinkotiheydestä. Tässä työssä molempia termejä käytetään tarkoittamaan samaa asiaa eli yhtenä vuonna sattuneiden vahinkojen lukumäärää. Kollektiivisessa riskimallissa yksittäisen vakuutuksen ensimmäistä vahinkoa seuraavat vahingot tulevat huomioiduksi. Tämä poikkeaa vakuutuskohtaisesta riskimallista. Mallissa ei myöskään ole nollamuuttujia, kuten vakuutuskohtaisessa mallissa. Kollektiivinen riskimalli on muotoa

$$S^{Coll} = \begin{cases} 0, & \text{jos } N = 0 \\ \sum_{j=1}^{N} X_j, & \text{jos } N \geq 1. \end{cases} \tag{2.2}$$

Yhtälössä X_i:t ovat samoin jakautuneita satunnaismuuttujia, joita arvotaan n kappaletta. Muuttujaa S kutsutaan yhdistetyksi muuttujaksi. Kollektiivinen riskimalli mallintaa siten vahinkojen summaa, kun N vahinkoa sattuu. Vakuutusyhtiön vahinkojakaumaa simuloitaessa, satunnaismuuttuja X kuvaa vahingon suuruutta ja satunnaismuuttuja N vahinkojen lukumäärä vahinkokannassa. Sopiva valinta yhdistetyn muuttujan mallintamiseksi voi olla mallinnustarpeesta riippuen esim. vakuutuslaji tai tuote. Kollektiivinen malli on hyvä approksimaatio vakuutuskohtaiselle mallille, kun vahingon todennäköisyys on pieni. Approksimoinnin idea on hakea sopiva yhdistetty jakauma, joka voidaan laskea käyttäen olemassa olevia laskentamenetelmiä. [6, s. 3-4] [12, s. 41]

2.4 Vahinkotiheys

Vahinkojen lukumäärä N on diskreetti satunnaismuuttuja ja sitä voidaan mallintaa esimerkiksi Poisson-jakaumalla tai negatiivisella binomijakaumalla. Poisson-jakaumalla on ominaisuus $\mathbb{E}(N) = \mathrm{Var}(N)$. Kun vahinkojen lukumääräjakauman varianssi on suurempi kuin odotusarvo ts. kun kyseessä on ylihajontatilanne, voidaan käyttää negatiivista binomijakaumaa. Kun varianssi on pienempi kuin odotusarvo, voidaan käyttää jakaumaa, jolla on tälläinen ominaisuus. Luonnollinen valinta tällöin on binomijakauma.

2.4.1 Poisson-jakauma

Oletetaan, että vakuutettuna on n sopimusta ja näistä osalle sattuu vahinko todennäköisyyden p mukaan. Tällöin yksinkertainen malli kuvaamaan ilmiötä olisi binomijakauma. Ison kannan tapauksessa tätä voidaan Poissonin raja-arvo lauseen mukaan approksimoida Poisson-jakaumalla. Sillä jos $n \to \infty$ ja $p \to 0$ siten että $np \to \lambda$, niin binomijakauman pistetodennäköisyysfunktio lähenee Poisson-jakauman pistetodennäköisyysfunktiota. Tällöin Poisson-jakauman käyttö vahinkotiheyden approksimoinnissa on perusteltu, kun vakuutuskanta on iso ja vahingot ovat harvinaisia. Poisson(λ)-jakauman pistetodennäköisyysfunktio on muotoa

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}. \tag{2.3}$$

Poisson(λ)-jakaumalle pätee

$$\mathbb{E}(X) = \mathrm{Var}(X) = \lambda, \quad \gamma = \frac{1}{\sqrt{\lambda}} \tag{2.4}$$

ja

$$\mathcal{M}_N(t) = e^{\lambda(e^t-1)}, \tag{2.5}$$

missä $\mathcal{M}(t)$ on Poisson-jakauman momentit generoiva funktio. Poisson-jakauman todennäköisyydet generoiva funktio on muotoa

$$\mathcal{G}(t) = e^{\lambda(t-1)}. \tag{2.6}$$

Momentit generoiva funktio ja todennäköisyydet generoiva funktio ovat määritelty liitteessä (B). [21, s. 317]

2.5 Vahingon suuruus

Yksittäisen vahingon suuruutta voidaan mallintaa empiirisellä tai analyyttisellä menetelmällä. Empiirinen menetelmä soveltuu suuren kannan tarkasteluun, jossa vahinkojen kokoja taulukoidaan jakaumaksi. Suuressakin kannassa voidaan joutua tasoittamaan saatua jakaumaa. Analyyttisessä tarkastelussa haetaan tilanteeseen parhaiten sopiva jakauma. Analyyttisen tarkastelun etu on se, että harvinaiset suuret vahingot jakauman hännässä ovat edustettuina. Esimerkkejä analyyttisistä jakaumista ovat gamma-jakauma sekä paksuhäntäiset lognormaali- ja Pareto-jakauma.

2.5.1 Gamma-jakauma

Jatkuvan jakauman $Gamma(\alpha, \beta)$ tiheysfunktio on muotoa

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}. \tag{2.7}$$

Gamma-jakauman ensimmäiset momentit ovat

$$\mathbb{E}(X) = \frac{\alpha}{\beta} \quad \mathrm{Var}(X) = \frac{\alpha}{\beta^2} \quad \gamma = \frac{2}{\sqrt{\alpha}}. \tag{2.8}$$

Gamma-jakauman ominaisuuksista lisää lähteessä [21].

3 Jakaumatyypit

3.1 Sekajakauma

Tärkeä luokka jakaumia saadaan, kun valitaan sekajakauma, jossa voi olla sekä diskreettejä että jatkuvia osia. Esimerkki tästä on Tweedie-jakauma, jossa Poisson-jakauma antaa arvon pisteessä nolla ja ns. zero-truncated gamma-malli antaa nollaa suuremmat arvot. Toinen vaihtoehto on käyttää esimerkiksi kahden pisteen Poisson-jakaumaa, jossa λ_1 kuvaa hyviä riskejä ja λ_2 kuvaa huonoja riskejä. Tällaisessa kaksipisteisessä Poisson sekajakaumassa on kolme estimoitavaa muuttujaa. Pisteiden määrän kasvaessa, myös estimoitavien pisteiden lukumäärä kasvaa.

3.2 Painotettu Poisson-jakauma

Poisson-jakauman ongelma on se, että siinä on käytössä vain yksi estimointiparametri λ, joka vastaa sekä odotusarvoa että varianssia. Usein estimoitavat ryhmät ovat ainakin jossain määrin heterogeenisiä. Tämä johtaa ylihajontaan eli tilanteeseen, jossa varianssi ylittää odotusarvon. Kyseistä ilmiötä voidaan mallintaa painotetulla Poisson-jakaumalla. Tässä jakaumassa Poisson-jakauman parametri λ ei ole vakio, vaan se on satunnaismuuttuja, joka saa arvon satunnaismuuttujasta Λ, jolla on oma jakauma. Painotetun Poisson-jakauman toista jakaumaa Λ:aa kutsutaan struktuurijakaumaksi. Esimerkki painotetusta Poisson-jakaumasta on negatiivinen binomijakauma. [9, s. 3]

3.3 Yhdistetty jakauma

Yhdistetty jakauma vastaa tilannetta, jossa yhtälön (2.2) muuttujat X ja N ovat riippumattomia ja muuttujat X samoin jakautuneita. Yhdistetyn jakauman momentit saadaan kaavasta

$$\mathcal{M}_S(t) = M_N(M_X(s)). \tag{3.1}$$

Yhdistetyn jakauman ensimmäiset momentit voidaan laskea myös suoraan

$$\begin{aligned} \mathbb{E}(S) &= \mathbb{E}(N)\,\mathbb{E}(X) \\ \mathrm{Var}(S) &= \mathbb{E}(N)\,\mathrm{Var}(X) + \mathrm{Var}(N)\,\mathbb{E}(X)^2 \end{aligned} \tag{3.2}$$

Kuten edellä on käynyt ilmi, voidaan yhdistetyn jakauman S kertymäfunktio laskea konvoluutiolla. Lisäksi pätee

$$\mathscr{L}_S(t) = \mathcal{G}_N[\,\mathscr{L}_X(t)]. \tag{3.3}$$

Yhtälöissä $\mathscr{L}_S(t)$ tarkoittaa S:n Laplace-muunnosta, joka on määritelty liitteessä (B). Yhdistettyyn jakaumaan liittyy käsitteitä, jotka ovat hyödyllisiä luokittelumielessä tai niistä on laskennallisen tehokkuuden mielessä apua. Esimerkkejä ensimmäisestä ovat uskottavuusluokittelu (reliability classification) ja ääretön jaettavuus (infinite divisibility). Esimerkki jälkimmäisestä on vaihetyyppi (phase type), joka mahdollistaa matriisimetodien käytön. [21, s. 314-316]

3.4 Yhdistetty Poisson-jakauma

Mikäli vahinkojen lukumäärä on Poisson-jakautunut, kutsutaan yhtälöä (2.2) yhdistetyksi Poisson-jakaumaksi. Yhdistetylle Poisson-jakaumalle pätee tärkeä tulos: Jos $S_1, S_2, \ldots, S_m$ ovat riippumattomia yhdistettyjä Poisson-jakautuneita satunnaismuuttujia parametrilla λ_i ja vahinkojakaumalla P_i, niin summa $S = S_1 + S_2 + \cdots + S_m$ on yhdistetty Poisson -jakauma parametreilla

$$\lambda = \sum_{i=1}^{m} \lambda_i \quad ja \quad P(x) = \sum_{i=1}^{m} \frac{\lambda_i}{\lambda} P_i(x), \tag{3.4}$$

missä $P_i(x)$ on vahingon i kertymäfunktio. Mikäli mallinnuksessa hyödynnetään yhdistettyä Poisson-jakaumaa, on esimerkiksi erilaisten vakuutuslajien vahinkojakaumien yhdistäminen kohtuullisen suoraviivaista. Yhdistetyn Poisson-jakauman momentit ja todennäköisyydet generoiva funktio sekä Laplace-muunnos on annettu lähteessä [21]. [12, s. 47][21, s. 318-319]

3.4.1 Yhteys painotetun ja yhdistetyn Poisson-jakauman välillä

On olemassa tärkeä yhteys painotetun Poisson-jakauman ja yhdistetyn Poisson-jakauman välillä. Tätä varten esitetään käsitteet diskreetti äärettömästi jaettavissa oleva funktio (DÄJ) ja äärettömästi jaettavissa oleva funktio (ÄJ).

Tarkastellaan diskreettiä jakaumaa F. Jos jokaiselle $n = 1, 2, \ldots$ on olemassa riippumattomat ja identtisesti jakautuneet satunnaismuuttujat F_n siten että $F = F_n + F_n + \cdots + F_n = F_n^{(n)}$, niin jakauma F on DÄJ. Analyyttinen seuraus tälle on se, että todennäköisyydet generoiva funktio G_n, jolla pätee vastaavasti jokaiselle positiiviselle kokonaisluvulle n, $G_N(s) = G_n^{(n)}(s)$ on DÄJ. Jos muuttuja on jatkuva, määritellään muutoin kuten edellä ja puhutaan äärettömästi jaettavissa olevasta (ÄJ). Todennäköisyydet generoivan funktion paikalla voidaan käyttää myös karakteristista funktiota tai Laplace-muunnosta.

Feller on todistanut seuraavan tuloksen: diskreetti muuttuja on DÄJ jos ja vain jos se on yhdistetty Poisson-jakauma. Macedalta on saatu tulos: mikäli N on painotettu Poisson, on se DÄJ, jos struktuurimuuttuja on ÄJ. Esimerkiksi jos tarkastellaan painotettua Poisson-gamma-muuttujaa, on ensin tutkittava onko gamma-muuttuja ÄJ. Gamma-jakauman tiheysfunktion Laplace-muunnos $\mathscr{L}(t)$ on muotoa

$$\mathscr{L}(t) = \left(1 + \frac{t}{\beta}\right)^{-\alpha} = \left(\frac{\beta}{\beta + t}\right)^{\alpha}. \tag{3.5}$$

Tämä on selvästi ÄJ, joten painotettu Poisson-jakauma on DÄJ ja se on samalla yhdistetty Poisson-muuttuja. Voidaan osoittaa, että tämän yhdistetyn Poisson-muuttujan sekundäärijakauma on logaritmisesti jakautunut. Gamma-jakauman lisäksi jakaumia, jotka ovat ÄJ, ovat normaali-, log-normaali- ja yleistetyt gamma-konvoluutiojakaumat. [2, s. 69-71][9, s. 27,33-34]

3.5 Yhdistetty painotettu Poisson

Kuvattaessa vakuutuskantaa, jossa ilmenee ylihajontaa, on turvauduttava painotettuun Poisson-jakaumaan vahinkotiheyden osalta. Tällöin Poisson-jakauman λ saa arvon struktuurijakaumasta Λ. Jos samalla halutaan kuvata kollektiivista vahinkokantaa muodossa $S = X_1 + X_2 + ... + X_N$, tuloksena on yhdistetty painotettu Poisson-jakauma (λ, Λ, X).

Yhdistetyn painotetun Poisson-jakauman (λ, Λ, X) momentit voidaan laskea momentit generoivien funktioiden avulla:

$$\mathcal{M}_S(t) = \mathcal{M}_\Lambda(\lambda(\mathcal{M}_X(t) - 1)). \tag{3.6}$$

[5, s. 139][15, s. 141-142]

4 Jakaumaperheet

4.1 Eksponentiaaliset dispersiomallit

Eksponentiaaliset dispersiomallit ovat diskreettejä tai jatkuvia jakaumia, joiden pistetodennäköisyys- tai tiheysfunktio on muotoa

$$p(y; \theta, \phi) = \exp\left\{\frac{y\theta - \mathcal{K}(\theta)}{a(\phi)} + C(y; \phi)\right\}, \tag{4.1}$$

missä $y \in \mathcal{C}$ ja $\mathcal{C}$ on joukko, jossa y on määritelty. $\mathcal{K}(\cdot)$ on kumulanttifunktio, jolla tunnettu ominaisuus $\mathcal{K}'(0) = \mu$. Funktio $C(\cdot;\cdot)$ toimii normalisoivana funktiona pakottaen tiheysfunktion summan tai integraalin ykköseksi. Varianssifunktio $\nu(\cdot)$ määritellään kumulanttifunktion avulla:

$$Var(X) = K''(\mathcal{K}'^{-1}(\mu_i)) =: \nu(\mu_i). \tag{4.2}$$

Kaikille eksponentiaalisille dispersiomalleille (EDM) pätee

$$\frac{1}{n}\sum_{n} ED(\mu, \phi) \sim ED(\mu, \phi/n). \tag{4.3}$$

Esimerkkejä EDM:stä ovat normaali-, Poisson-, binomi-, negatiivinen binomi-, gamma-, käänteinen normaali- ja Tweedie-jakauma. [7, s. 4][11, s. 226-227]

4.2 Tweedie-mallit ja Tweedien suppenemislause

Tweedie-mallit ovat jakaumaperhe, joka sisältää diskreettejä ja jatkuvia jakaumia. Olkoon $p \in \mathbb{R}$ vakio. Tweedie-perheeseen kuuluvalle satunnaismuuttujalle pätee

$$\mathrm{Var}(Y) = \phi[\mathbb{E}(Y)]^p, \text{ kun } p \notin (0, 1), \tag{4.4}$$

missä ϕ on dispersio-parametri. Tweedie-mallia ei ole määritelty välillä $(0, 1)$ vaikka silloinkin parametriestimaatit voidaan estimoida maksimoimalla ns. kvasi-likelihood-yhtälöt. Tweedie-mallit erilaisilla p:n arvoilla esitetään taulukossa 1. Varianssifunktio määritellään $\nu(\mu) = \mathrm{Var}(Y)/\phi$, joten Tweedie-malleille pätee

$$\nu(\mu) = \mu^p. \tag{4.5}$$

Jakauma	p:n arvo
äärimmäinen stabiili	$p < 0$
normaali	0
Poisson	1
Tweedie (yhdistetty Poisson-gamma)	$1 < p < 2$
gamma	2
positiivinen stabiili	$2 < p < 3$
käänteinen normaalijakauma	3
positiivinen stabiili	$p > 3$

Taulukko 1: Tweedie-malleja erilaisilla p:n arvoilla

Tweedie-mallit on mielenkiintoinen luokka jakaumia, sillä sen jakaumat toimivat rajajakaumana EDM-perheen jakaumille. Tarkastellaan eksponentiaalisia dispersiomalleja $ED(\mu, \sigma^2)$, joille pätee $\nu(\mu) \propto c_0\mu^p$. Nyt jos $\mu \to 0$ tai $\mu \to \infty$, niin välillä $(c, c-1)$ malli $c^{-1}ED(c\mu, \sigma^2 c^{2-p})$ suppenee Tweedie-mallia $Tweedie(\mu, c_0\sigma^2)$ kohti, kun joko $c \to 0$ tai $c \to \infty$.

Edellisestä seuraa se, että monesti Taylorin laiksi kutsutulla suppenemisilmiöllä on matemaattinen tausta ja kyseessä on Tweedien suppeneminen. Kendal ja Jørgensen antavat lukuisia esimerkkejä ilmiöstä eri tieteenaloilta. [19, s. 36-37][13, s. 535]

4.2.1 Tweedie-mallien etuja

Tweedie-mallit ovat osajoukko eksponentiaalisia dispersiomalleja. Tästä seuraa se, että EDM:lle osoitetut lauseet pätevät myös Tweedie-malleille. Toisin sanoen jos Y_1 ja Y_2 kuuluvat samaan EDM-perheeseen, niin näiden painotettu summa $Y = (w_1Y_1 + w_2Y_2)/(w_1 + w_2)$ kuuluu samaan EDM-jakaumaan. On myös osoitettu, että EDM-jakaumista skaalainvariansseja ovat vain Tweedie-mallit. Skaalainvarianteille jakaumille pätee, että jos $c > 0$, niin Y ja cY seuraavat samaa jakaumaa. Tästä on etuna mm. se, että Tweedie-malleilla voidaan vaihtaa valuuttakurssia säilyttäen jakauma. Myös konversio promillen ja prosentin välillä onnistuu jakaumaa muuttamatta.

Varianssifunktio määrittää todennäköisyysjakaumaperheen yksikäsitteisesti EDM-luokassa. Tämä tarkoittaa sitä, että selvittämällä varianssifunktio tiedetään todennäköisyysjakauma EDM-luokassa. Varianssifunktio riippuu odotusarvosta. Odotusarvon ja varianssin selvittäminen on huomattavasti yksinkertaisempaa, kuin koko jakauman määrittely. [19, s. 34][16, s. 24]

4.2.2 Yhdistetty Poisson-gamma-jakauma

Kiinnostava jakauma Tweedie-mallien joukossa on yhdistetty Poisson-gamma-jakauma, jota kutsutaan myös Tweedie-jakaumaksi. Tweedie-mallien joukossa tämä vastaa jakaumaa, jolla $1 < p < 2$. Kun oletetaan kokonaisvahinkomallissa (2.2), että vahinkotiheys on Poisson-jakautunut ja vahinkojen suuruus on gamma-jakautunut, momentit generoiva funktio käyttäen kaavaa (3.1) on

$$\mathcal{M}_S(t) = e^{\lambda[(1-\frac{t}{\beta})^{-\alpha}-1]}. \tag{4.6}$$

Momentit generoivan funktion avulla voidaan osoittaa, että

$$\mathbb{E}(S) = \frac{\lambda\alpha}{\beta},\ \mathrm{Var}(S) = \frac{\lambda\alpha(1+\alpha)}{\beta^2}. \tag{4.7}$$

Yhdistettyä Poisson-gamma-jakaumaa ei voida esittää suljetussa muodossa, mutta se voidaan esittää päättymättömänä sarjana. Tweedie-jakaumalla on kiinnostavia sovelluskohteita myös vakuutusalan ulkopuolella. Tweedie-jakauma on yhdistetty Poisson-gamma-jakauma ja gamma on EDM-jakauma. Tästä seuraa se, että gamma-jakaumien summa ja samalla myös Tweedie-jakaumien summa on gamma-jakautunut. [7, s. 4-5][22, s. 16][23, s. 6-7]

4.3 Poisson-Tweedie-mallit

Jørgenssen ja Kokonendji ovat taulukoineet Poisson-Tweedie-jakaumia, joissa struktuurijakauma on Tweedie-perheestä. Esimerkki tällaisesta jakaumasta on Poisson-käänteinen normaalijakauma.

Poisson-Tweedie-konvergenssilause osoittaa, että monet jakaumat konvergoituvat Poisson-Tweedie-jakaumiin. Tämän perusteella voidaan olettaa, että Poisson-Tweedie-jakaumia esiintyy käytännön sovelluksissa. Vahinkomallinnuksessa erityisesti negatiivinen binomi-, Pólya-Aeppli-, Sichel- ja Poisson-käänteinen normaalijakauma ovat saaneet mainintoja kirjallisuudessa. [10, s. 1-2,13-14] [14, s. 688-689]

5 Kokonaisvahinkomäärän laskenta

5.1 Analyyttiset menetelmät

5.1.1 Momentit ja karakteristinen funktio

Kollektiivisessa riskimallissa (2.2) vahinkojen lukumäärän ja vahinkojen koon satunnaismuuttujat muodostavat yhdistetyn muuttujan. Oletuksena on, että vahinkojen lukumäärä ja vahinkojen suuruus ovat riippumattomia sekä vahinkojen suuruudet ovat keskenään riippumattomia. Tällöin kokonaisvahinkomäärän odotusarvo ja varianssi on laskettavissa:

$$\mathbb{E}(S) = \mathbb{E}(N)\,\mathbb{E}(X) \tag{5.1}$$

$$\operatorname{Var}(S) = \mathbb{E}(N)\operatorname{Var}(X) + \mathbb{E}(X)^2\operatorname{Var}(N). \tag{5.2}$$

Mikäli vahinkotiheys on Poisson-jakautunut, muuttuvat yhtälöt muotoon:

$$\mathbb{E}(S) = \lambda\,\mathbb{E}(X) \tag{5.3}$$

$$\operatorname{Var}(S) = \lambda\operatorname{Var}(X) + \mathbb{E}(X)^2\lambda = \lambda\,\mathbb{E}(X^2). \tag{5.4}$$

Lisäksi pätee

$$\mathcal{M}_S(t) = \mathcal{M}_N(\mathcal{C}_X(t)), \tag{5.5}$$

missä $\mathcal{C}_X(t)$ on kumulantit generoiva funktio. Oletetaan, että vahinkotiheys on Poisson-jakautunut. Silloin S vastaa yhdistettyä Poisson-muuttujaa. Sen momentit generoiva funktio on

$$\mathcal{M}_S(t) = e^{\lambda(e^{\mathcal{C}_X(t)}-1)} = e^{\lambda(\mathcal{M}_X(t)-1)}. \tag{5.6}$$

Vastaavasti voidaan esittää edellinen todennäköisyydet generoivan funktion avulla:

$$\mathcal{G}_S(t) = e^{\lambda(e^{\mathcal{C}_X(t)}-1)} = e^{\lambda(\mathcal{G}_X(t)-1)}. \tag{5.7}$$

Mikäli X:n jakauma on paksuhäntäinen, ei momentit generoivaa funktiota ole olemassa. Silloin on tutkittava karakteristisia funktioita

$$\phi_X(t) = \mathbb{E}(e^{itX}) = \mathbb{E}\big(\cos(tX) + i\sin(tX)\big). \tag{5.8}$$

Karakteristisen funktion käytön haasteena on se, että silloin on operoitava kompleksiluvuilla. Toisaalta karakteristinen funktio on aina olemassa. Tässä esitetyt kaavat löytyvät esimerkiksi lähteestä [21].

5.1.2 Konvoluutio

Muotoa $S = X_1 + X_2 + \cdots + X_N$ olevalle satunnaismuuttujalle voidaan laskea kertymäfunktio konvoluution avulla:

$$\begin{aligned} F_S(x) = \mathbb{P}(S \le x) &= \sum_{n=0}^{\infty} \mathbb{P}(S \le s \mid N = n)\,\mathbb{P}(N = n) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(X_1 + \cdots + X_N \le s)p_n \\ &= \sum_{n=0}^{\infty} p_n F_X^{*n}(x). \end{aligned} \tag{5.9}$$

Konvoluutio voidaan laskea vastaavasti tiheysfunktiolle:

$$f_S(x) = \mathbb{P}(S = x) = \sum_{n=0}^{\infty} p_n f_X^{*n}(x). \tag{5.10}$$

Tietyt jakaumat ovat suljettuja konvoluution suhteen. Kun jakauma on suljettu konvoluution suhteen, merkitään

$$\sum_n ED(\mu, \phi) \sim ED(\mu n, \phi). \tag{5.11}$$

Edellisen ehdon toteutuessa konvoluution laskemisen luonne helpottuu. Aktuaarisovelluksissa hyödyllisiä jakaumia, jotka ovat suljettuja konvoluution suhteen, ovat mm. Poisson-, gamma- ja käänteinen normaalijakauma. Kun vahinkojakauma on suljettu konvoluution suhteen, kaava (5.9) muuttuu muotoon:

$$F_S(x) = \sum_{n=0}^{\infty} p_n F_X^{*n}(x; a) = \sum_{n=0}^{\infty} p_n F_X(x; na). \tag{5.12}$$

Konvoluutiosta päästään eroon, mutta jäljellä jää ääretön summattavaa. Eräs käytännön ratkaisu laskurutiiniin on valita kynnys, jonka alittavia todennäköisyyksiä p_n ei oteta enää summaan mukaan. Lähteinä toimivat [12] ja [21].

5.2 Numeeriset menetelmät

Yleisesti sovellettuja numeerisia menetelmiä kokonaisvahinkojakauman ratkaisemiseksi on useita, joista kolme yleisintä ovat Monte Carlo, Panjerin rekursio ja nopea Fourier-muunnos (FFT).

5.2.1 Monte Carlo

Monte Carlo -simulaatio on yksinkertaisin numeerinen menetelmä. Simulaatio voidaan esittää prosessina:

1. Jokaiselle $k = 1, \ldots, K$
   (a) Arvotaan satunnaismuuttuja N.
   (b) Arvotaan N kappaletta satunnaismuuttujia X.
   (c) Lasketaan $S_k = \sum_{i=1}^{N} X_i$.
2. Seuraava k.

[18, s. 9-10]

5.2.2 Panjerin rekursio

Olkoon p_k diskreetin satunnaismuuttujan todennäköisyysfunktio. Tällöin se kuuluu luokkaan $(a, b, 0)$, jos on olemassa sellaiset vakiot a ja b, että

$$\frac{p_k}{p_{k-1}} = a + \frac{b}{k}, \quad k = 1, 2, 3, \ldots. \tag{5.13}$$

Luokkaa $(a, b, 0)$ kutsutaan myös Sundt-Jewell-perheeksi, mutta tässä käytetään kompaktia nimeä. Voidaan osoittaa, että diskreettejä jakaumia, jotka kuuluvat luokkaan $(a, b, 0)$ ovat Poisson-, binomi- ja negatiivinen binomijakauma.

Kun vahinkotiheyden N jakauma on luokkaan $(a, b, 0)$ kuuluva funktio, voidaan kertymäfunktio tuottaa Panjerin rekursiivisella algoritmilla, kun tunnetaan vahinkojen tiheysjakauma ja vahingon suuruusjakauma.

Olkoot yhtälössä (2.2) $f_S(x)$ S:n todennäköisyysfunktio ja $f_X(x)$ X:n todennäköisyysfunktio. Olkoon N luokkaan$(a, b, 0)$ kuuluva jakauma. Mikäli X:n jakauma on diskreetti, on Panjerin rekursio muotoa

$$\begin{aligned} f_S(0) &= \mathcal{G}(f_X(0)) \\ f_S(x) &= \sum_{y=1}^{x} \left(a + b\frac{y}{x}\right) f_X(y) f_S(x-y). \end{aligned} \tag{5.14}$$

Vastaavasti voidaan esittää Panjerin rekursiokaava tapauksessa, jossa X on jatkuvasta jakaumasta. Panjerin rekursio voi olla hidas, koska jatkuvassa tapauksessa on ratkaistava Volterran integraali numeerisesti. Laskentaa voidaan yksinkertaistaa approksimoimalla vahingon suuruusjakaumaa diskretoinnilla ja käyttämällä yhtälöä (5.14). Tässä tapauksessa Panjerin rekursio on numeerinen menetelmä. [14, s. 227-228][20, s. 38-39]

5.2.3 Nopea Fourier-muunnos

Nopea Fourier-muunnos on tehostettu menetelmä laskea diskreetti Fourier-muunnos. Fourier-muunnoksen avulla voidaan laskea konvoluutio tehokkaasti. Idea nähdään konvoluutiolauseesta

$$\mathscr{F}^{-1}(\mathscr{F}(X) \cdot \mathscr{F}(X)) = X * X. \tag{5.15}$$

Ja vastaavasti n:n konvoluution tapauksessa:

$$\mathscr{F}^{-1}(\mathscr{F}(X)^n) = X^{*n}. \tag{5.16}$$

Laskennallisesti hyödynnetään diskreettiä Fourier-muunnosta, jossa jakauma diskretisoidaan riittävän moneen pisteeseen ja tehdään laskelmat tällä. Tehokkain tapa laskea diskreetti Fourier-muunnos on tehdä se nopealla Fourier-muunnoksella. Tämä muunnos hyödyntää matriisilaskennassa modulo-laskentaa vähentääkseen erilaisia kompleksilukukertolaskuja. Laskentaperiaatteet esitti ensimmäisenä Gauss vuonna 1805. Ne löydettiin uudelleen kuitenkin vasta 1900-luvulla ja koottiin yhdeksi kokonaisuudeksi 1960-luvulla.

Fourier-muunnos on määritelty muodossa

$$\mathscr{F}(s) = \mathbb{E}(e^{-2\pi itS}) = \int_{-\infty}^{\infty} f(t)e^{-2\pi its}dt. \tag{5.17}$$

Käänteinen Fourier-muunnos määritellään

$$\mathscr{F}^{-1}(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-2\pi its}dt. \tag{5.18}$$

Mikäli kollektiiviselle riskimallille (2.2) lasketaan kaksiulotteisen kertymäjakauman k-kertainen konvoluutio käyttäen nopeaa Fourier-muunnosta, tapahtuu se kaavan

$$\phi_S(t) = \mathcal{G}_N[\phi_X(t)] \tag{5.19}$$

mukaisesti. Tämä voidaan merkitä diskreetin Fourier-mallin avulla

$$\mathbf{M}_T = \frac{1}{n}\mathscr{F}^{-1}(\mathcal{G}_N(\mathscr{F}(\mathbf{M}_L))), \tag{5.20}$$

jossa $\mathbf{M}_L$ tarkoittaa lähtömatriisia ja $\mathbf{M}_T$ tulosmatriisia. Molempien matriisien dimensio on $n \times n$. Mikäli on kyse yksiulotteisesta matriisista ts. vektorista, on matriisin ulottuvuus $n \times 1$. Esimerkiksi kun kyseessä on yhdistetty Poisson-malli, jonka todennäköisyydet generoiva funktio on muotoa (2.6) tulee lauseke muotoon

$$\mathbf{M}_T = \frac{1}{n}\mathscr{F}^{-1}\left(e^{\lambda(\mathscr{F}(\mathbf{M}_L)-1)}\right). \tag{5.21}$$

Kun vahinkojen lukumäärä N on $NegBin(r, p)$-jakautunut, saadaan konvoluutio kaavasta

$$\mathbf{M}_T = \frac{1}{n}\mathscr{F}^{-1}\left(\left(\frac{p}{1-(1-p)\mathscr{F}(\mathbf{M}_L)}\right)^r\right). \tag{5.22}$$

Algoritmi on tehokkaimmillaan, kun diskretisoidun vahingon suuruusjakauman osien lukumäärä on muotoa 2^r, jollekin $r \in \mathbb{N}$. R-koodi konvoluution laskemiseen yhdistetylle Poisson-jakaumalle löytyy liitteestä (C). [1, s. 8-9,172-182][4, s. 61-65][18, s. 17-20]

5.3 Muita approksimointimenetelmiä

Vastaavasti kuten vakuutuskohtaisessa riskimallissa, niin kollektiivista riskimallia voidaan approksimoida ensimmäiset momentit huomioivalla normaaliapproksimaatiolla ja kolmannen momentin huomioivalla NP-approksimaatiolla tai kolmeparametrisella gamma-approksimaatiolla. Näiden lisäksi on kehitetty esim. Wilson-Hilferty-approksimaatio, joka NP-approksimaatiota paremmin ottaa huomioon koko jakauman muodon. Haittapuolena on analyyttinen kompleksisuus. Nämä approksimointimenetelmät olivat tärkeitä ennen moderneja tietokoneita, jolloin rekursiomenetelmät ja FFT olivat epäkäytännöllisiä. Edelleen nämä approksimointimenetelmät ovat hyödyllisiä esim. silloin, kun ei tunneta jakauman tarkkaa muotoa. Lisäksi approksimaatiolla voidaan saada nopea arvio tilanteesta tai sitä voidaan hyödyntää tarkempien metodien suuruusluokkatarkistuksessa.

Approksimoinnit voidaan jakaa kahteen luokkaan. Ensimmäinen luokka on momenttien täsmäytystä tunnettuihin jakaumiin. Toinen luokka on kokonaisvahinkojakauman muuntamista tunnetuksi jakaumaksi. Esimerkkejä ensimmäisestä luokasta ovat normaaliapproksimaatio, kolmeparametrinen gamma-approksimaatio, Bowerin gamma-approksimaatio, NP-approksimaatio, Haldanen metodi ja Wilson-Hilferty-metodi. Esimerkiksi Hardy sovitti kaikkia edellisiä metodeita erilaisiin yhdistettyihin Poisson-Pareto-jakaumiin ja ilmeni, että normaaliapproksimaatio oli selkeästi huonoin ja Haldane-jakauma oli paras myös hyvin vinoissa jakaumissa. Toisen luokan esimerkki on Esscherin approksimaatio, jota käytetään kun konvoluutio on liian kompleksinen laskea. Kun ensimmäisen luokan metodeissa riitti tuntea momentteja, niin Esscherin approksimaatiossa on tunnettava koko jakauma. Metodi on jäänyt vähemmälle käytölle, kun rekursio- ja FFT-menetelmistä on tullut tehokkaampia nopeiden koneiden myötä. Esscherin approksimaatio tunnetaan tilastotieteessä eksponentiaalisena kallistusmenetelmänä tai satulapisteapproksimaationa. [21, s. 76-85]

5.4 Menetelmien tehokkuudesta ja käytettävyydestä

Konvoluutio perustuu todennäköisyysjakaumien yhteenlaskuun. Tämän laskennan tehokkuus on $\mathcal{O}(n^3)$. Tehokkaampia ratkaisuja ovat Panjerin algoritmi ja inversiometodit (mm. nopea Fourier-muunnos). Panjerin algoritmin tarvitsema laskujen lukumäärä on kertaluokkaa $\mathcal{O}(n^2)$. Esimerkiksi, kun n=1000, laskenta-aika konvoluution ja Panjerin algoritmin välillä pienenee $99,9\%$. Nopea Fourier-muunnos tarvitsee vain ajan $\mathcal{O}(n \log(n))$. Tehokkuuden lisäksi täytyy huomioida se, että rekursioilla negatiivisten vahinkojen käsittely on hankalaa, kun taas FFT-algoritmille ne eivät tuota ongelmia. Bühlmann havaitsi, että edellä mainittu FFT:n nopeusetu on voimassa, kun vahinkojakauma on diskretoitu riittävästi ($s \geq 255$). Kuitenkin rekursio on nopeampi, kun diskretointi on matala ($s \leq 63$).Täten voidaan päätellä FFT:n voittavan Panjerin rekursion tehokkuudessa, kun diskretointi on viety riittävän pitkälle. Toisaalta molemmat metodit ovat merkittävästi nopeampia kuin konvoluution laskenta suoraan. Huomattava on, että Bühlmannin tarkastelu perustui laskutoimitusten määrään.

Embrechts ja Frei totesivat rajan 255 olevan nykystandardien mukaan arkaainen, koska laskenta sekä Panjerin rekursiolla että FFT:llä tällaisella diskretoinnilla kestävät sekunnin murto-osia. He laskivat jakaumaa yhdistetylle Poisson-Pareto-jakaumalle ja totesivat aikaeron FFT:n eduksi olevan merkittävä. Toinen tuore tietotekninen tarkastelu Shevchenkon toimesta osoittaa, että laskettaessa yhdistettyä Poisson-logaritmista-jakaumaa Panjerin algoritmilla ja FFT:llä, jälkimmäinen on selvästi nopeampi. Shevchenkon esimerkissä, kun vahinkojen lukumäärän odotusarvo on tuhat, Panjerin rekursio kestää tunteja ja FFT sekunteja. [3, s. 116-126][8, s. 497-508][14, s. 252-253][18, s. 26-28]

5.5 Metodien vertailua

Tutkitaan kolmen edeltävän metodin (Monte Carlo, Panjerin rekursio ja FFT) toteutusta R-ohjelmointikielellä. Otetaan esimerkiksi yhdistetty Poisson-gamma($\lambda, \alpha = 5, \beta = 0.01$)-jakauma. Lasketaan eri metodien mukaiset $99,5\%$:n kvantiilit. Tutkitaan metodien antamaa lopputulosta ja verrataan käytettyä laskenta-aikaa.

Panjerin metodissa haasteeksi muodostui underflow-ilmiö johtuen siitä, että laskennassa $f_S(0) = e^{-1\,000}$ pyöristyy nollaksi. Tällöin rekursio antaa nollavektorin. Ilmiö voidaan kiertää joko sopivalla skaalauksella tai laskemalla osajakaumia sekä yhdistämällä nämä konvoluutiolla. Liian suurella skaalauksella haasteeksi voi ilmetä vastavuoroisesti overflow-ilmiö. Taulukossa (2) on vertailtu metodeita erilaisilla λ:n arvoilla. Nopeasti havaitaan, että Panjerin rekursio ei ole yksinään toimiva ratkaisu, kun λ on suuri. Tässä työssä Panjerin rekursiossa on käytetty lähteessä ([12]) mainittua R-koodia.

	λ	10	100	500	1 000	10 000
MC(10^5)	$Q_{0.995}$	10 156	64 611	281 830	545 031	5 141 543
	aika	0.9s	1.9s	7.1s	12.1s	1.9min
MC(10^6)	$Q_{0.995}$	10 091	64 730	282 094	545 292	5 142 020
	aika	10.5s	22.7s	1.2min	2.2min	21.6min
MC(10^7)	$Q_{0.995}$	10 086	64 754	282 238	545 272	-
	aika	1.9min	3.6min	17.5min	26.1min	-
Panjer	$Q_{0.995}$	10 103	64 879	282 753	-	-
	aika	1.6s	14.1s	2.3min	-	-
FFT	$Q_{0.995}$	10 103	64 879	282 755	546 349	5 152 116
	aika	0.0s	0.3s	1.2s	2.2s	23.3s

Taulukko 2: Metodien vertailua erilaisilla λ:n arvoilla.

MC-menetelmän antamat arvot poikkeavat hieman Panjerin ja FFT:n antamista arvoista. Kyseessä onkin simulaation tulos, jossa tarkkuus paranee, kun simulaatioiden lukumäärää lisätään. Satunnaiselementistä johtuen näin ei välttämättä yksittäisissä laskentatilanteissa käy. MC-menetelmässä on käytetty R-kielen funktioita $rpois(x, \lambda)$ ja $rgamma(x, \alpha, \beta)$. Simulaatioiden lukumäärille on käytetty kolmea eri arvoa (10^5, 10^6 ja 10^7).

Oletuksena on, että Panjerin rekursio on tarkin menetelmä ja tuloksia verrataan suhteessa siihen. Nopea Fourier-muunnos osoittautuu nopeaksi ja tarkaksi menetelmäksi. MC-menetelmän etuna on mm. se, että siinä voidaan rakentaa muuttujien välille riippuvuuksia yksinkertaisemmin kuin Panjer- ja FFT-menetelmässä.

5.6 Riippumattomuusolettama

Riskiteoriassa oletetaan perinteisesti, että riskit ovat toisistaan riippumattomia. Tästä oletuksesta on edellä nähtyjä laskennallisia etuja. Katastrofitilanteessa riippumattomuusolettama ei välttämättä päde. Suomalaisia vakuutusyhtiöitä koskevia esimerkkejä tällaisista katastrofeista 2000-luvulla ovat myrskyt (Janika- ja Tapanimyrsky vuosina 2001 ja 2011), sääilmiöistä johtuvat liikenteen kolarisumat vuosina 2005 ja 2012 ja Thaimaan tsunami vuonna 2004. Kaikki esimerkit aiheuttivat monelle samalla alueella olevalle vakuutetulle yhtäaikaisesti vahinkoja. Katastrofivahinkoja voidaan mallintaa käyttäen vahinkojen riippumattomuusmallia ja lisäten satunnaiset katastrofivahingot. Mikäli tyydytään kokonaisvahinkojakauman momenttien laskemiseen, saadaan riippuvuuksia sisältävän kollektiivimallin momentit laskettua yksinkertaisesti olettaen että suuruusjakauman momentit ovat olemassa. Paksuhäntäisillä jakaumilla tällaista vaihtoehtoa ei momenttien puuttuessa välttämättä olemassa.

6 Yhteenveto

Lähtökohta työhön oli selvittää käyttökelpoisia parametrisia malleja kokonaisvahinkojakauman mallinnukseen, tutkia nopean Fourier-muunnoksen tehokkuutta suhteessa muihin menetelmiin ja ymmärtää yhdistetyn Poisson-jakauman ja painotetun Poisson-jakauman teoriaa osana kokonaisvahinkomallia. Parametrisia kokonaisvahinkojakaumia, joiden tiheysfunktio on suljetussa muodossa, on lopulta harvoja. Tämä aiheuttaa sen, että sovelluksissa hyödynnetään usein helposti käsiteltäviä jakaumia. Tilasto-ohjelmistoilla kokonaisvahinkojakauman laskenta on nopeaa käyttäen approksimointimenetelmiä, joiden tarkkuutta voidaan säätää. Tässä työssä nostettiin esille nopean Fourier-muunnoksen tehokkuus. Mainittakoon, että jokaisella metodilla on omat edut ja heikkoutensa. Työssä ei käsitelty nopeaan Fourier-muunnokseen sisältyviä haasteita, jotka ilmenevät erityisesti liian pienellä diskretoinnilla. Panjerin menetelmä on todennäköisesti tarkin menetelmä ja se antaa myös vasemman puoleisen hännän vertailussa selkeästi parhaiten. Embrechts ja Frei osoittivat, että nopea Fourier-muunnos pystyy eksponentiaalisella kallistuksella tämänkin puutteen korjaamaan. Sovelluksissa ollaan kuitenkin yleensä kiinnostuneita vahinkojakauman oikeanpuoleisesta hännästä. Lopuksi voi todeta, että tässä työssä tehty laskentaesimerkki tukee vallitsevaa käsitystä metodien välillä olevista eroista. [8, s. 497-508]

Viitteet

[1] Birgham E. O., *The Fast Fourier Transform*, Prentice-hall, 1974 (Viitattu sivulla 15.)

[2] Bühlmann H., *Mathematical methods in risk theory*, Springer, 2005 (2nd print) (Viitattu sivulla 9.)

[3] Bühlmann H., *Numerical Evaluation of the Compound Poisson Distribution: Recursion or Fast Fourier Transform*, Scandinavian Actuarial Journal No.2, 1984 (Viitattu sivulla 16.)

[4] Cerchiara R.R., *FFT, Extreme Value Theory and Simulation to Model Non-Life Insurance Claim Dependencies*, in book: Perna, C., Sibillo M., *Mathematical and Statistical Methods in Insurance and Finance*, Springer, 2008 (Viitattu sivulla 15.)

[5] Chadjiconstantinidis S.,Antzoulakos D. L, *Moments of Compound Mixed Poisson Distribution*, Scandinavian Actuarial Journal No.3, 2002 (Viitattu sivulla 9.)

[6] Dufrense F., *Between the individual and collective models, revisited*, University of Lausanne, 2002 (Viitattu sivulla 5.)

[7] Dunn P.K., Smyth, G.K. *Series Evaluation of Tweedie Exponential Dispersion Model Densities*, Statistics and Computing, 2005 (Viitattu sivuilla 9 ja 11.)

[8] Embrechts P., Frei M. *Panjer recursion versus FFT for compound distributions*, Mathematical Methods of Operations Research 69 No.3, 2008 (Viitattu sivuilla 16 ja 18.)

[9] Grandell J., *Mixed Poisson Processes*, Chapman & Hall, 1997 (Viitattu sivuilla 7 ja 9.)

[10] Jørgensen, B., Kokonendji, C.C., *Discrete Dispersion Models and their Tweedie Asymptotics*, arXiv:1409.7482v1, 2014 (Viitattu sivulla 11.)

[11] Jørgensen, B., Martínez, J.R., Tsao, M., *Asymptotic behaviour of the variance function*, Scand. J. Stat., 1994 (Viitattu sivulla 9.)

[12] Kaas R., Goovaerts M., Dhaene J., Denuit M. *Modern Actuarial Risk Theory*, Springer, 2008 (Viitattu sivuilla 5, 8, 13, ja 17.)

[13] Kendal W.S., Jørgensen B., *A Scale Invariant Distribution of the Prime Numbers*, Computation No.3, 2015 (Viitattu sivulla 10.)

[14] Klugman S.A., Panjer H.P., Willmot G.E., *Loss Models*, Wiley, 2008 (Viitattu sivuilla 11, 14, ja 16.)

[15] Nadarajah S., Kotz S., *Compound mixed Poisson distributions I*, Scandinavian Actuarial Journal No.3, 2006 (Viitattu sivulla 9.)

[16] Ohlson E., Johansson B., *Non-life Insurance Pricing with Generalized Linear Models*, Verlag, 2010 (Viitattu sivulla 11.)

[17] Parodi P., *Pricing in General Insurance*, CRC Press, 2015 (Viitattu sivulla 4.)

[18] Shevchenko P.V., *Calculation of aggregate loss distributions*, arXiv:1008.1108v1, 2010 (Viitattu sivuilla 13, 15, ja 16.)

[19] Song P.X.-K., *Correlated Data Analysis: Modeling, Analytics, and Applications*, Springer, 2007 (Viitattu sivuilla 10 ja 11.)

[20] Sundt B., Vernic R., *Recursions for Convolutions and Compound Distributions with Insurance Applications*, Springer, 2009 (Viitattu sivulla 14.)

[21] Teugels J.L., Sundt B., *Encyclopedia of Actuarial Science*, Wiley, 2004 (Viitattu sivuilla 6, 7, 8, 12, 13, ja 16.)

[22] Withers S., Nadarajah S. *On the Compound Poisson-Gamma Distribution*, Kybernetika No. 42, 2011 (Viitattu sivulla 11.)

[23] Yi Yang, Wei Qianyand, Hui Zouz *A Boosted Tweedie Compound Poisson Model for Insurance Premium*, arXiv:1508.06378v2 [stat.ME], 2016 (Viitattu sivulla 11.)

Liitteet

A Työhön liittyviä rajauksia

Esitetty työssä	**Rajattu pois**
yksiulotteiset jakaumat	moniulotteiset jakaumat
riippumattomat vahingot	riippuvuudet, kopulat
yhdistetty painotettu Poisson	laskentaesimerkit
kollektiivimallin jakauman laskenta	laskennan tarkkuus
keskeisimmät kollektiivimallin numeeriset menetelmät	esim. suora numeerinen integrointi ja momenttipohjainen tiheysapproksimaatio
FFT	laskennan yksityiskohdat, aliasing, eksponentiaalinen kallistus

B Muunnokset

Karakteristinen funktio Karakteristinen funktio vastaa käänteistä Fourier-muunnosta. Se määritellään:

$$\phi_X(t) = \mathbb{E}\left[e^{itX}\right] = \mathscr{F}^{-1}(t), \quad \forall t \in \mathbb{R}\,. \tag{B.1}$$

Edelliselle pätee hyödyllinen ominaisuus:

$$\phi_{aX+b}(t) = e^{itb}\phi_X(at). \tag{B.2}$$

Laplace-muunnos

$$\mathscr{L}_X(t) = \mathbb{E}\left[e^{-tX}\right] = \int_0^\infty e^{-tx} f(x)\,\mathrm{d}x. \tag{B.3}$$

Laplace-muunnokselle pätee

$$\mathscr{L}\{af(x) + b\,g(x)\} = a\,\mathscr{L}\{f(x)\} + b\,\mathscr{L}\{g(x)\}. \tag{B.4}$$

ja

$$\mathscr{L}\{f(x) * g(x)\} = f(x)g(x). \tag{B.5}$$

Momentit generoiva funktio

$$\mathcal{M}_X(t) = \mathbb{E}\left[e^{tX}\right] = \int_0^\infty e^{tx} f(x)\,\mathrm{d}x. \tag{B.6}$$

Momentit generoivalle funktiolle ja riippumattomille satunnaismuuttujille X ja Y pätee:

$$\begin{aligned}\mathcal{M}_{X+Y}(t) &= \mathbb{E}(e^{(X+Y)t}) = \mathbb{E}(e^{Xt}e^{Yt}) \\ &= \mathbb{E}(e^{Xt})\,\mathbb{E}(e^{Yt}) = \mathcal{M}_X(t)\,\mathcal{M}_Y(t)\end{aligned} \tag{B.7}$$

Kumulantit generoiva funktio

$$\mathcal{C}_X(t) = \ln(\mathcal{M}(t)) \tag{B.8}$$

Joten kumulantit generoivalle funktiolle pätee, kun X ja Y ovat riippumattomia:

$$\begin{aligned}\mathcal{C}_{X+Y}(t) &= \ln(\mathcal{M}_{X+Y}(t)) = \ln(\mathcal{M}_X(t)\,\mathcal{M}_Y(t)) \\ &= \ln(\mathcal{M}_X(t)) + \ln(\mathcal{M}_Y(t)) \\ &= \mathcal{C}_X(t) + \mathcal{C}_Y(t)\end{aligned} \tag{B.9}$$

Todennäköisyydet generoiva funktio

$$\mathcal{G}_X(t) = \mathbb{E}\left[t^X\right] = \int_0^\infty t^x f(x)\,\mathrm{d}x. \tag{B.10}$$

Painotetun Poisson-jakauman todennäköisyydet generoivan funktion ja Laplace-muunnoksen välillä vallitsee yhteys:

$$\mathcal{G}_X(t) = \mathscr{L}_X(1-t). \tag{B.11}$$

Edellinen yhteys nähdään kaavoista

$$\begin{aligned}
\mathcal{G}_X(t) = \mathbb{E}(t^X) &= \sum_{x=0}^{\infty} t^x p_x \\
&= \sum_{x=0}^{\infty} t^x \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!} g(\lambda)\,\mathrm{d}\lambda \\
&= \int_0^\infty \sum_{x=0}^{\infty} \frac{(t\lambda)^x}{x!} e^{-\lambda} g(\lambda)\,\mathrm{d}\lambda \\
&= \int_0^\infty e^{\lambda(t-1)} g(\lambda)\,\mathrm{d}\lambda \\
&= \mathscr{L}_X(1-t).
\end{aligned} \tag{B.12}$$

Muunnosten välinen yhteys Edellä käsiteltyjen muunosten välillä voidaan osoittaa olevan seuraava yhteys, mikäli muunnos on olemassa:

$$\mathcal{M}_X(t) = \mathscr{L}_X(-t) = \mathcal{G}_X(e^t) = \phi(-it) \tag{B.13}$$

Kun puhutaan painotetusta Poisson-jakaumasta, on muunnosten välillä yhteys:

$$\mathcal{M}_X(t) = \mathscr{L}_X(-t) = \mathcal{G}_X(t-1) \tag{B.14}$$

C Nopean Fourier-muunnoksen R-koodi

```r
library(actuar)

# Konvoluutiofunktio yhdistetylle Poisson-jakaumalle
konvoluutioPois <- function(x, y) {
     tulo <- length(x) * y
     n <- (2^(ceiling(log2(tulo))))
     vek <- rep(0, n)
     length(p) <- length(vek)
     vek <- vek + p
     apu = c(0, vek[-n])
     apu[is.na(apu)] <- 0
     palaute = Re(fft(exp(y * (fft(apu) - 1)), T)/n)
     return(palaute)
}

# Suuruusjakauman diskretisointi
p <- discretize(pgamma(x, 50, 0.05), method = "
    unbiased", lev = levgamma(x, 50, 0.05), from = 0,
    to = 3000, step = 1)

# Konvoluution laskeminen
k <- konvoluutioPois(p, 100)
```