Stockholms universitet

Utvärdering av GLM-baserad reservsättning som tillåter separata skattningar av IBNR och RBNS

En studie på en av Kammarkollegiets försäkringsprodukter

Amanda Wiman

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2018:1
Försäkringsmatematik
Februari 2018

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Matematisk statistik
Stockholms universitet
Masteruppsats **2018:1**
http://www.math.su.se/matstat

Utvärdering av GLM-baserad reservsättning som tillåter separata skattningar av IBNR och RBNS

En studie på en av Kammarkollegiets försäkringsprodukter

Amanda Wiman*

Februari 2018

Sammanfattning

I detta masterarbete undersöks vilken metod som är bäst lämpad för att reservsätta en av Kammarkollegiets försäkringsprodukter. Vi strävar efter att hitta en metod som enbart kräver lättillgänglig data, är enkel att implementera och ger stabila skattningar. Samtidigt vill vi hitta en modell som smidigt går att anpassa på flera olika försäkringsprodukter. Produkten vi tittar på i detta arbete är en personförsäkring som Kammarkollegiet tillhandahåller där vi har information om alla rapporterade skador för åren 2002 - 2016.

Vi kommer att titta på reservsättningsmetoden beskriven i (Verrall et al. , 2010), i detta arbete kommer denna refereras till som VNJ-metoden. Metoden baserar reservskattningar på två aggregerade skadetrianglar. En triangel med skadebelopp byggs upp och vi betingar sedan denna på en aggregerad triangel uppbyggd av antalet skador, denna teknik gör att vi fångar upp både information om fördröjningstid för kända och okända skador. Fördelen med detta är att vi då kan skatta och analysera reserverna för kända och okända skador separat. Eftersom VNJ-metoden inte tillåter flera utbetalningar per skada har två olika sätt att hantera flera utbetalningar undersökts. Dessa metoder har sedan jämförts mot den välanvända chain ladder-metoden.

Då vi skattar reserven för de första åtta skadeåren, som vi då antar är slutreglerade, och jämför mot den faktiska reserven ser vi ingen märkbar skillnad mellan de olika metoderna. Vidare analyser visar dock att VNJ-metoden ger betydligt stabilare skattningar än chain ladder-metoden oavsett stickprovsstorlek. De två olika sätten att hantera utbetalningar i VNJ-metoden ger liknande resultat i de flesta testerna.

VNJ-metoden visar ett flertal fördelar jämfört med chain ladder-metoden vilket indikerar att den kan vara ett lämpligt alternativ som reservsättningsmetod.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: amandawiman12@gmail.com . Handledare: Mathias Lindholm.

Abstract

This master thesis aims to find a suitable method for claim reserving one of Kammarkollegiets insurance products. The goal is to find a method which works using simple and easily accessed data, is relatively simple to implement and gives stable estimations. The method should also be suitable for other types of insurance products. In this thesis, the product is a personal insurance provided by Kammarkollegiet with data ranging from 2002 - 2016.

This thesis will examine the reserving method described in (Verrall *et al.* , 2010) (from now on referred to as the VNJ method). The VNJ method bases its estimates on two aggregated development triangles. One development triangle is constructed using paid claims, which is conditioned on a second development triangle based on the number of claims. This technique gathers information on delays for both known and unknown claims. An advantage of doing this is that it's possible to estimate and analyse reserves for known and unknown claims separately. Since the VNJ method does not allow multiple payment per claim, two different ways of dealing with multiple payments are explored and compared to the well-known chain-ladder method.

When claim reserves for the first eight years are estimated (which are assumed to be settled), and compared to the actual reserve, there are only small differences to be found when the different methods are compared. However, further analysis reveals that the claim estimates produced by the VNJ method are more stable than those produced by the Chain-Ladder method, regardless of sample size. The two different ways of dealing with multiple payments per claim behave in a similar fashion in most of the tests.

The VNJ method shows a number of benefits compared to the Chain-Ladder method, which indicates that it may be a viable option as a claim reserving method.

Förord och tack

Detta arbete utgör ett examensarbete på masternivå om 30 hp inom försäkringsmatematik vid matematiska institutionen på Stockholms Universitet.

Jag vill tacka Fabian Österman på Kammarkollegiet som har hjälpt till med framtagning av försäkringsdata samt svarat på alla mina frågor om deras verksamhet. Även ett stort tack till min handledare Mathias Lindholm på Stockholms Universitet som har gett värdefull feedback och som hela tiden har kommit med nya idéer för att driva detta arbete framåt.

Innehåll

Sammanfattning **i**

Abstract **ii**

Förord och tack **iii**

1 Inledning **1**
1.1 Syfte och metod 1
1.2 Bakgrund 1
1.2.1 Solvens II-regelverket 1
1.2.2 Kammarkollegiet 2

2 Teori **3**
2.1 Reservsättning 3
2.2 Chain ladder-metoden 4
2.2.1 Antaganden i chain ladder-metoden 6
2.3 VNJ-metoden 6
2.3.1 Antal skador 7
2.3.2 Skadekostnad 8
2.3.3 Skattning av reserver 13

3 Dataanalys **15**
3.1 Brister och justering av data 16
3.2 Beroende mellan rapporteringstid och storleken på betalningen . . . 16
3.3 Storleken på utbetalningar 18

4 Modellering **19**
4.1 Reservsättning med chain ladder-metoden 19
4.1.1 Utvärdering chain ladder-metoden 20
4.2 Reservsättning med VNJ-metoden 21
4.2.1 Hantering av fler utbetalningar per skada 22
4.2.2 Approximation av utbetalningar 22
4.2.3 Parameterskattningar 24
4.2.4 Analys av parameterskattningar 24
4.2.5 Osäkerhet i skattningen av ψ 26
4.2.6 Införande av intercept 27
4.3 Modellutvärdering av reserversättningsmetoder 28
4.3.1 Jämförelse mot faktisk reserv 29
4.4 Stabilitetstest reserver 32
4.5 Robusthetsanalys 33

4.5.1 Avvikelse från faktisk reserv för olika stickprovsstorlekar . . 34
4.5.2 Reservskattningar då vi inte har en dominant utbetalning . . 35
4.5.3 Större andel skador med fler utbetalningar 37
4.6 Olika metoder för att approximera utbetalningar 39

5 Diskussion **42**
5.1 Reservsättningsmetoderna 42
5.2 Resultat från modelleringen 43

6 Slutsats **44**

7 Vidare utveckling av arbetet **45**

Referenser **46**

1 Inledning

I detta avsnitt beskrivs syftet med arbetet samt metoden som används. Det finns även en beskrivning till bakgrunden som lett fram till frågeställningen.

1.1 Syfte och metod

Detta arbete går ut på att försöka hitta en reservsättningsmetod till en av Kammarkollegiets försäkringsprodukter. Vi testar två olika metoder för att se vilken som passar bäst. En noggrannare titt tas även på brister i de olika modellerna och hur stabila skattningar de båda metoderna producerar.

Vi kommer att anpassa en metod som använder information om de utbetalade skadebeloppen, och som även betingar på antalet rapporterade skador som beskrivs i (Verrall *et al.* , 2010) och jämföra denna metod med den mer klassiska och välanvända chain ladder-metoden.

När vi sedan tittar på vilken reservsättningsmetod som passar Kammarkollegiets data bäst kommer flera saker spela in. Vi vill självklart ha en metod som ger trovärdiga och inte allt för volatila skattningar av reserven. Men vi kommer även värdera modellens simplicitet samt dess förmåga att anpassas på flera olika typer av försäkringsprodukter utan allt för mycket manuellt arbete.

1.2 Bakgrund

1.2.1 Solvens II-regelverket

Solvens II är ett riskbaserat regelverk som trädde i kraft 2016-01-01, dess syfte är bland annat att ge ett ökat skydd till försäkringstagare. Regelverket sätter krav på att försäkringsbolagen ska vara tillräckligt solventa, det vill säga att de har tillräckligt med kapital för att infria sina framtida åtaganden gentemot försäkringstagarna.

I direktivet står det bland annat “tillsynsordningen bör innefatta ett riskkänsligt krav, baserat på en prospektiv beräkning för att säkerställa korrekta ingripanden i god tid av tillsynsmyndigheterna (solvenskapitalkravet) och en miniminivå för säkerheten som medlen inte får underskrida (minimikapitalkravet). Båda dessa kapitalkrav bör harmoniseras inom hela gemenskapen så att en enhetlig skyddsnivå för försäkringstagarna kan uppnås. För att detta direktiv ska fungera väl bör det

finnas lämpliga ingripandenivå mellan solvenskapitalkravet och minimikapitalkravet." (Eiopa, 2009, s.6).

Alla försäkringsbolag måste följa regelverket och behöver därför vara medvetna om vilka risker de står inför, vilka åtaganden de har och alltså även se till att ha tillräckligt mycket kapital.

1.2.2 Kammarkollegiet

Kammarkollegiet är en statlig förvaltningsmyndighet som utför olika uppdrag åt regeringen. Deras uppgifter består bland annat av att representera den allmänna arvsfonden, att effektivisera den offentliga förvaltningen samt att de försäkrar hela den statliga verksamheten i Sverige (Kammarkollegiet, n.d.). Myndigheter i Sverige får inte teckna försäkringar hos privata försäkringsbolag, utan får istället efter eget beslut eller efter uppdrag från regeringen teckna försäkring hos Kammarkollegiet (Sjöberg, 2007, s.1). Med privata försäkringsbolag menas de bolag som försäkrar privatpersoner och företag och som inte finansieras med statliga medel.

I det statliga försäkringssystemet finns det ett tjugotal försäkringsprodukter, de tillhandhåller försäkringar inom olika områden. Exempelvis fordon-, verksamhet- och personförsäkringar. Systemet ämnar att skydda statliga myndigheter och de som är anställda vid myndigheterna, men även personer som har andra anknytningar till statlig verksamhet (däribland studenter vid statliga högskolor) (Sjöberg, 2007, Förord).

Privata försäkringsbolag står under Finansinspektionens tillsyn och har strikta krav på sig att följa försäkringslagstiftningen. Kammarkollegiets försäkringsverksamhet har inte krav på sig att följa samma regelverk som privata försäkringsbolag. Anledningen till detta är att i försäkringsrörelselagens mening är inte det statliga försäkringssystemet en försäkringsverksamhet (Sjöberg, 2007, s.3).

Verksamheten finansieras av försäkringspremier som myndigheterna betalar. De har även en kreditram på 100 MSEK som de kan använda om enskilt stora skador skulle inträffa, eller om det totala skadebeloppet under ett år blir väldigt högt. Om ett enskilt skadebelopp överstiger 50 MSEK är det bestämt att det överstigande beloppet kommer att täckas utanför systemet och det är regeringen som tar beslut om hur beloppet ska finansieras (Sjöberg, 2007, s.3). Detta kan liknas med ett återförsäkringsskydd som privata försäkringsbolag ofta tecknar.

Även om det statliga försäkringssystemet inte ses som en försäkringsverksamhet

och därför inte behöver följa samma regelverk som privata försäkringsbolag finns det många likheter i de båda verksamheterna. På samma sätt som privata försäkringsbolag får även det statliga systemet in anmälningar om skador som ska regleras och där en rimlig reserv behöver sättas av. Det är på samma sätt viktigt för det statliga systemet att vara konsoliderat, att ha tillräckliga medel för att kunna infria sina åtaganden. Det finns därför ett intresse av att ta fram en metodik för att bestämma storleken på Kammarkollegiets försäkringsfond så att det statliga försäkringssystemet är tillräckligt konsoliderat (där konsolideringsgraden baseras på det regelverk som privata försäkringsbolag måste följa, Solvens II).

2 Teori

I detta avsnitt ges en övergripande presentation av teorin som används i detta arbete, det ges även förslag på vidare läsning där man kan fördjupa sig mer inom respektive ämne. Teorin om reservsättning är främst hämtad från (Wüthrich, 2017), mer om chain ladder-metoden kan bland annat läsas i (Mack, 1993) och om VNJ-metoden kan man läsa i (Verrall *et al.* , 2010).

2.1 Reservsättning

För ett försäkringsbolag ska intäkter i form av premier under en viss period täcka de kostnader som uppstår för skador som sker under den perioden. Ett problem är att det i slutet av en period är svårt att veta hur mycket kostnader som uppstått. Det är vanligtvis inte känt då alla skador ofta inte har rapporterats vid periodens slut, samt att det ofta inte finns komplett information om hur mycket de skador som kommit in under perioden kommer att kosta totalt. (Wüthrich, 2017)

Bolaget behöver därför sätta av pengar, reserver, för dessa skador. Reserven kan delas upp i två huvudsakliga poster, IBNR och RBNS. Där IBNR står för "incurred but not reported" som innebär en reserv för de skador som inträffat under perioden men som inte ännu har rapporterats. RBNS står för "reported but not settled" vilket är en reserv för de skador som har rapporterats, men där man ännu inte vet vilket belopp den totala skadekostnaden kommer att uppgå till.

Det finns olika metoder för att skatta reserven, den kanske mest använda metoden kallas chain ladder-metoden och beskrivs i nästa avsnitt. Denna metod skattar hela reserven och delar alltså inte upp den mellan IBNR och RBNS, mer om chain ladder-metoden kan bland annat läsas i (Mack, 1993). Vi kommer även titta på en

metod som tillåter separat skattning av IBNR och RBNS lite längre fram, teorin om denna metod är hämtad från (Verrall *et al.* , 2010).

2.2 Chain ladder-metoden

En vanligt förekommande metod för att skatta total skadekostnad är chain ladder-metoden. Metoden antar att en viss typ av försäkringsprodukt kommer att bete sig liknande för olika skadeår, och skattar därför utvecklingsfaktorer som beskriver förändringen från ett utvecklingsår till ett annat. Dessa utvecklingsfaktorer används sedan för samtliga skadeår.

Data sätts upp på en triangelform, se exempel i Tabell 1, raderna anger skadeår och kolumnerna utvecklingsår. Vilken typ av data triangeln bygger på kan variera, man kan använda antalet skador (i detta fall får en medelskada skattas separat), utbetalat belopp eller känd skadekostnad. Där känd skadekostnad består av utbetalat belopp plus individuellt satta reserver som skadereglerarna uppskattat då skadan inkommit. (Wüthrich, 2017)

Hur många utvecklingsår som behövs bestäms efter hur försäkringsprodukten ser ut, antalet utvecklingsår ska motsvara hur många år efter ett skadeår alla skador från det året är slutreglerade och stängda.

i\j	0	1	2	3	4	5	6	7	8	9
1	451288	339519	333371	144988	93243	45511	25217	20406	31482	1729
2	448627	512882	168467	130674	56044	33397	56071	26522	14346	
3	693574	497737	202272	120753	125046	37154	27608	17864		
4	652043	546406	244474	200896	106802	106753	63688			
5	566082	503970	217838	145181	165519	91313				
6	606606	562543	227374	153551	132743					
7	536976	472525	154205	150564						
8	554833	590880	300964							
9	537238	701111								
10	684944									

Tabell 1: Exempel på inkrementella skadekostnader för tio skadeår och tio utvecklingsår (Verall et al. , 2010).

I exemplet i Tabell 1 låter vi $X_{i,j}$ beteckna värdet i cell i och kolumn j, här mot-

svarar i skadeår och j utvecklingsår. Varje element i triangeln bygger på skadekostnader där $X_{i,j}$ motsvarar hur mycket som har betalats ut för skador som inträffat under skadeår i under utvecklingsår j. Den undre triangeln motsvarar tiden som ännu ej har passerat och där har vi därför ingen information. Utifrån informationen i den övre triangeln skattas reserver för den undre triangeln. Observera att Tabell 1 visar inkrementella värden, i formlerna nedan för chain ladder-metoden kommer $X_{i,j}$ vara ackumulerade skadekostnader, det vill säga hur mycket som totalt har betalats ut efter j utvecklingsår för skadeår i.

Chain ladder-metoden tittar på hur den ackumulerade skadekostnaden förändras från ett utvecklingsår, j, till nästa år, $j+1$, för samtliga skadeår och skattar en utvecklingsfaktor, f_j, som beskriver denna förändring. Man antar sedan att oavsett vilket skadeår i vi tittar på kommer skadekostnaden mellan utvecklingsår j och $j+1$ ha samma procentuella ökning. Vi skattar utvecklingsfaktorerna, $f_1, ..., f_{I-1} > 0$, enligt

$$\hat{f}_j = \frac{\sum_{i=1}^{I-j} X_{i,j+1}}{\sum_{i=1}^{I-j} X_{i,j}}, \tag{2.1}$$

där $I \leq j \leq I-1$, (Mack, 1993). Den totala skadekostnaden för år i, alltså den totala kostnaden skadorna som skedde år i uppgick till när alla skador är slutreglerade, kallas ultimo skadebelopp. Ultimo skadebeloppen anges som X_{iJ} och skattas som

$$\hat{X}_{iJ} = X_{i,J+1-i} \cdot \hat{f}_{J+1-i} \cdot ... \cdot \hat{f}_{J-1}, \tag{2.2}$$

(Mack, 1993). Den totala reserven för skadeår i kommer sedan att ges av ultimo skadebelopp minus känt skadebelopp. Vi skriver chain ladder skattningen av reserven som

$$\hat{R}_i^{CL} = \hat{X}_{i,J} - X_{i,J-i+1} \tag{2.3}$$

$$\hat{R}^{CL} = \sum_{i=2}^{n} \hat{R}_i^{CL} \tag{2.4}$$

2.2.1 Antaganden i chain ladder-metoden

För att vi ska kunna använda chain ladder-metoden behöver vissa antaganden vara uppfyllda. Antagandena ges av

(1) $E[X_{i,j+1}|X_{i1},...,X_{ij}] = X_{ij}f_j, \quad 1 \leq i \leq I, 1 \leq j \leq I-1$

(2) $\{X_{i1},...,X_{iI}\}, \{X_{j1},...,X_{jI}\}, i \neq j$, är oberoende (Mack, 1993).

Det första antagandet säger att det förväntade värdet av det ackumulerade skadebeloppet utvecklingsår $j+1$, givet att vi vet alla ackumulerade skadebelopp upp tills utvecklingsår j, kommer ges av det ackumulerade skadebeloppet år j multiplicerat med en utvecklingsfaktor, f_j. Där f_j är oberoende av vilket år skadan har inträffat, alltså oberoende av i.

Det andra antagandet säger att de ackumulerade utbetalade skadebeloppen de kommande åren för skador som skett år i ska vara oberoende av de ackumulerade utbetalade skadebeloppen de kommande åren för skador som skett år j för alla $i \neq j$.

Mer om chain ladder-metoden kan läsas i bland annat (Mack, 1993) och (Wüthrich, 2017).

2.3 VNJ-metoden

Vidare ska vi titta på en metod som tillåter oss att skatta IBNR- och RBNS-reserverna separat. Vi kommer att utgå ifrån metoden som beskrivs i (Verrall *et al.* , 2010). Härefter kommer vi i detta arbete att kalla denna metod för VNJ-metoden efter författarna Verrall, Nielsen och Jessen som skrivit artikeln (Verrall *et al.* , 2010). Fördelen med denna metod är att den inte har fullt lika många begränsningar som chain ladder-metoden, samtidigt som den enbart använder lättillgänglig data och dessutom är relativt enkel att implementera.

Till skillnad från chain ladder-metoden baseras VNJ-metoden på två inkrementella skadetrianglar, en med skadeutbetalningar (antingen baserad på utbetalt belopp eller känd skadekostnad) och en med antalet skador. I detta arbete väljer vi att basera vår skadetriangel på utbetalt belopp. Metoden utnyttjar att vi ofta har information både om hur mycket som har betalats ut men även om hur många skador som har rapporterats in. Genom att skatta framtida utbetalningar då vi betingat på antalet rapporterade skador kommer modellen att fånga upp både IBNR-fördröjning (tiden från att en skada sker tills att den rapporteras), samt RBNS-fördröjning (tiden från att en skada rapporterats tills att den är färdigutbetald).

Det leder till att vi kan skatta reserven för IBNR- och RBNS-skador separat.

I denna metod tar vi fram två inkrementella skadetrianglar, en med skadebelopp och en med antalet rapporterade skador. Precis som i chain ladder-metoden har vi bara observerat informationen i den övre triangeln. Vi låter först $\mathcal{A}_m$ beteckna området som vi har observerat och skriver därför $\mathcal{A}_m = \{(i,j) \in \mathbb{N} \times \mathbb{N}_0 : 1 \leq i + j \leq m\}$. Våra skadetrianglar kommer att byggas upp av de skadebelopp och skador som ligger inom området $\mathcal{A}_m$, se (Verrall *et al.* , 2010).

Nedan beskrivs hur vi konstruerar vår modell för att kunna skatta IBNR- och RBNS-reserven med VNJ-metoden.

2.3.1 Antal skador

Totala antalet skador som inträffat år i och rapporterats efter j år ges av N_{ij} där $(i,j) \in \mathcal{A}_m$. Vi betecknar den stokastiska variabeln med $\aleph_m = \{N_{ij} : (i,j) \in \mathcal{A}_m\}$, denna antas ha en överspridd poissonfördelning, det vill säga tillhöra en Tweedie-fördelning. Vissa av skadorna N_{ij} kommer att vara slutreglerade och kommer därför inte att bidra till någon reserv, medan resterande skador kommer att vara våra RBNS-skador. Vi låter N_{ijk}^{paid} beteckna de skador som skett under period i, rapporterats efter j perioder, och som är helt utbetalda k perioder efter att den rapporterats. Vi får sambandet mellan N_{ij} och N_{ijk}^{paid} genom att summera över alla k enligt nedan

$$N_{ij} = \sum_k N_{ijk}^{paid}. \tag{2.5}$$

Sedan antas att antalet slutreglerade skador betingat på att vi vet totala antalet rapporterade skador kommer att ha en multinomial fördelning på följande form

$$(N_{ij0}^{paid}, ..., N_{ijd}^{paid}) \sim \text{Multi}(N_{ij}; p_0, ..., p_d), \tag{2.6}$$

för $(i,j) \in \mathcal{A}_m$. Där p_k är sannolikheten att en skada är helt utbetald k perioder efter att den rapporterats, p_k uppfyller $p_0 + ... + p_d = 1$. Här låter vi k gå från 0 till d, där d alltså är den längsta RBNS-fördröjningstiden vi tillåter, se (Verrall

et al. , 2010).

IBNR-skadorna kommer att ges av de skador som ännu ej är rapporterade, alltså N_{ij} så att $(i, j) \notin \mathcal{A}_m$. Dessa är inte observerade och behöver skattas, detta kan göras genom att utveckla skadetriangeln för antalet skador med chain ladder-metoden beskriven ovan.

2.3.2 Skadekostnad

En begränsning i denna modell är att den enbart kan hantera en utbetalning per skada. Detta eftersom vi vill att det totala skadebeloppet i en viss cell $X_{i,j}$ ska motsvaras av antalet skador som ligger i samma cell i vår andra skadetriangel, $N_{i,j}$. Detta är något vi kommer behöva förhålla oss till när vi skapar vår triangel med utbetalade belopp och vi kommer återvända till detta i modelleringen.

De individuella skadeutbetalningarna ges av $Y_{ij}^{(k)}$, där k anger vilken skada, i anger vilket år skadan inträffat, och j anger vilket utvecklingsår utbetalningen skett. De aggregerade inkrementella skadebeloppen betecknas med X_{ij}, vilket är det totala skadebeloppet för skador som inträffat år i och betalats ut efter j år, detta kan skrivas som

$$X_{ij} = \sum_{k=1}^{N_{ij}^{paid}} Y_{ij}^{(k)} \tag{2.7}$$

för $(i, j) \in \mathcal{A}_m$.

Här antas det att de individuella utbetalningarna inte beror på antalet skador och inte heller på fördröjningstiderna. Det antas även att $Y_{ij}^{(k)}$ är oberoende och lika fördelade. Dessa antaganden kommer inte alltid vara helt rimliga och vi kommer därför diskutera kring detta i dataanalysen som följer.

Vi vill skatta de framtida skadebeloppen betingat på att vi vet antalet rapporterade skador, alltså kommer $X_{ij}|\aleph_m$ att vara responsvariabel. Nedan visar vi hur de två första momenten av responsvariabeln, $E[X_{ij}|\aleph_m]$ och $V[X_{ij}|\aleph_m]$, kan skrivas. Dessa härledningar är hämtade från (Verrall *et al.* , 2010).

$$
\begin{aligned}
E[X_{ij}|\aleph_m] =& E[E[X_{ij}|N_{ij}^{paid}]|\aleph_m] \\
=& E\left[E\left[\sum_{k=1}^{N_{ij}^{paid}} Y_{ij}^{(k)}|N_{ij}^{paid}\right]|\aleph_m\right] \\
=& E[N_{ij}^{paid}E[Y_{ij}^{(k)}|N_{ij}^{paid}]|\aleph_m] \\
=& E[Y_{ij}^{(k)}]E[N_{ij}^{paid}|\aleph_m]
\end{aligned} \tag{2.8}
$$

$$
\begin{aligned}
V[X_{ij}|\aleph_m] =& E[V[X_{ij}|N_{ij}^{paid}]|\aleph_m] + V[E[X_{ij}|N_{ij}^{paid}]|\aleph_m] \\
=& E[V[\sum_{k=1}^{N_{ij}^{paid}} Y_{ij}^{(k)}|N_{ij}^{paid}]|\aleph_m] + V[N_{ij}^{paid}E[Y_{ij}^{(k)}]|\aleph_m] \\
=& E[N_{ij}^{paid}V[Y_{ij}^{(k)}|\aleph_m]] + V[N_{ij}^{paid}E[Y_{ij}^{(k)}]|\aleph_m]
\end{aligned} \tag{2.9}
$$

Här står N_{ij}^{paid} för antalet skador som har inträffat år i och är helt utbetalda efter j utvecklingsår, vi kan skriva dessa som

$$
N_{ij}^{paid} = N_{ij0}^{paid} + N_{i,j-1,1}^{paid} + ... + N_{i,j-min\{j,d\},min\{j,d\}}^{paid} = \sum_{k=0}^{min\{j,d\}} N_{i,j-k,k}^{paid}, \tag{2.10}
$$

för $(i,j) \in \mathcal{A}_m$.

Eftersom vi antagit att de individuella utbetalningarna inte beror på antalet skador kan vi skriva variansen som

$$
V[X_{ij}|\aleph_m] = E[N_{ij}^{paid}|\aleph_m]V[Y_{ij}^{(k)}] + V[N_{ij}^{paid}|\aleph_m]E[Y_{ij}^{(k)}]^2 \tag{2.11}
$$

Ett antagande är även att $Y_{ij}^{(k)}$ är oberoende och lika fördelade kan vi kan därför låta $V[Y_{ij}^{(k)}] = \sigma^2$ och $E[Y_{ij}^{(k)}] = \mu$ och vi får då

$$
V[X_{ij}|\aleph_m] = E[N_{ij}^{paid}|\aleph_m]\sigma^2 + V[N_{ij}^{paid}|\aleph_m]\mu^2 \tag{2.12}
$$

Det återstår då att ta fram väntevärde och varians av N_{ij}^{paid}, väntevärdet följer enligt nedan

$$\begin{aligned} E[N_{ij}^{paid}|\aleph_m] =& E\left[\sum_{k=0}^{min\{j,d\}} N_{i,j-k,k}^{paid}|\aleph_m\right] \\ =& \sum_{k=0}^{min\{j,d\}} E[N_{i,j-k,k}^{paid}|\aleph_m] \\ =& \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k \end{aligned} \tag{2.13}$$

Med hjälp av ekvation (2.13) kan (2.8) skrivas som

$$E[X_{ij}|\aleph_m] = \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k\mu \tag{2.14}$$

Vi tittar även på variansen

$$\begin{aligned} V[N_{ij}^{paid}|\aleph_m] =& V\left[\sum_{k=0}^{min\{j,d\}} N_{i,j-k,k}^{paid}|\aleph_m\right] \\ =& \sum_{k=0}^{min\{j,d\}} V[N_{i,j-k,k}^{paid}|\aleph_m] \\ =& \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k(1-p_k) \end{aligned} \tag{2.15}$$

Sätter vi sedan in ekvation (2.13) och (2.15) i (2.12) får vi

$$
\begin{aligned}
V[X_{ij}|\aleph_m] &= \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k\sigma^2 + \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k(1-p_k)\mu^2 \\
&= \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k(\sigma + \mu^2(1+p_k)) \qquad (2.16) \\
&\approx \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k(\sigma + \mu^2)
\end{aligned}
$$

Med denna approximation kan vi slutligen skriva variansen som en funktion av väntevärde genom

$$
Var(X_{ij}|\aleph_m) \approx (\sigma^2 + \mu^2) \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k = \rho E[X_{ij}|\aleph_m] \qquad (2.17)
$$

där $\rho = \frac{\sigma^2+\mu^2}{\mu}$.

Vi introducerar även $\psi_k = \mu p_k$ så att formel (2.14) kan skrivas på följande form

$$
E[X_{ij}|\aleph_m] = \sum_{k=0}^{min\{j,d\}} N_{i,j-k}p_k\mu = \sum_{k=0}^{min\{j,d\}} N_{i,j-k}\psi_k \qquad (2.18)
$$

Tittar vi på formel (2.18) ser vi att denna kan ställas upp på matrisform och att vi kan låta X_{ij} vara vår responsvariabel, $\sum_{k=0}^{min\{j,d\}} N_{i,j-k}$ bildar vår designmatris och ψ_k våra regressionskoefficienter som vi vill skatta och vi får alltså formen

$$
\begin{bmatrix} X_{10} \\ X_{20} \\ \vdots \\ X_{m0} \\ X_{11} \\ \vdots \\ X_{m-1,1} \\ \vdots \\ X_{1,m-1} \end{bmatrix}
=
\begin{bmatrix}
N_{10} & 0 & 0 & \dots & 0 \\
N_{20} & 0 & 0 & \dots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
N_{m0} & 0 & 0 & \dots & 0 \\
N_{11} & N_{10} & 0 & \dots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
N_{1,m-1} & N_{1,m-2} & N_{1,m-3} & \dots & N_{1,0}
\end{bmatrix}
\cdot
\begin{bmatrix} \psi_0 \\ \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{m-1} \end{bmatrix}
$$

Vi får då ett ekvationssystem som går att lösa med en generaliserad linjär modell. Generaliserade linjära modeller (GLM) är en generalisering av vanlig linjär regression med normalfördelade feltermer. GLM tillåter, till skillnad från en enkel linjär regressionsmodell, responsvariabeln att ha en annan fördelning än normalfördelningen. För att använda en GLM behöver vi en länkfunktion, länkfunktionen är en monoton, deriverbar funktion som beskriver sambandet mellan väntevärdet av responsvariabeln och de förklarande variablerna, i uppställningen ovan har en identitetslänk används. Mer om generaliserade linjära modeller kan bland annat läsas i (McCullagh & Nelder, 1989) samt på s. 116 i (Agresti, 2002).

För att kunna använda en GLM behöver vi anta en fördelning för $Y_{ij}^{(k)}$, i (Verrall *et al.* , 2010) antas en kvasilikelihood. Detta eftersom då man använder en generaliserad linjär modell krävs bara de två första momenten för en kvasilikelihood och vi behöver alltså inte känna till hela fördelningen. I (McCullagh & Nelder, 1989) introduceras följande funktion

$$U = u(\mu; Y) = \frac{Y - \mu}{\sigma^2 V(\mu)}, \tag{2.19}$$

som uppfyller samma egenskaper som derivatan av en log-likelihood, det vill säga att

$$E(U) = 0, \tag{2.20}$$

$$var(U) = 1/\{\sigma^2 V(\mu)\}, \tag{2.21}$$

$$-E(\frac{\partial U}{\partial \mu}) = 1/\{\sigma^2 V(\mu)\} \tag{2.22}$$

Denna teori bygger sedan på att vi antar, att eftersom $u(\mu; Y)$ uppfyller samma villkor som en log-likelihood, att om denna funktion existerar kommer den även bete sig som en log-likelihood. Man låter därför

$$Q(\mu; y) = \int_y^\mu \frac{y - t}{\sigma^2 V(t)} dt, \tag{2.23}$$

vara den logaritmerade kvasilikelihood. I (McCullagh & Nelder, 1989) skriver det sedan att eftersom Y är oberoende (vilket även vi har antagit) kan den totala kvasilikelihooden skrivas

$$Q(\mu; y) = \sum Q_i(\mu_i; y_i) \tag{2.24}$$

Vi behöver alltså inte ha specifik information om fördelningen av responsvariabeln för att kunna sätta upp en likelihoodfunktion. Istället räcker det med att visa att variansen kan skrivas som en funktion av väntevärdet, det vill säga att $var(Y_i) = V(\mu_i)$. Det vi behöver göra är därför enbart att ta fram de två första momenten och visa att variansen kan skrivas som en funktion av väntevärdet. Härledningen för de två första momenten, som kommer ifrån (Verrall *et al.* , 2010), visades ovan och relationen mellan väntevärde och varians ses i ekvation (2.17).

Vi har då allt vi behöver för att sätta upp en GLM för vår kvasilikelihood och får då ut skattningar av regressionskoefficienterna ψ_k där $\psi_k = p_k\mu$, se även ekvation (2.18). Summerar vi över alla k på båda sidorna kan vi skriva

$$\sum_k \psi_k = \sum_k p_k\mu = \mu \sum_k p_k = \mu \tag{2.25}$$

vilket i sin tur leder till att p_k kan skrivas som

$$p_k = \frac{\psi_k}{\sum_k \psi_k}. \tag{2.26}$$

Vi har då uttryckt μ och p_k i termer av enbart ψ_k vilket gör att vi även kan få ut skattningar av μ och p_k.

2.3.3 Skattning av reserver

VNJ-metoden tillåter oss, till skillnad från chain ladder-metoden, att addera fler utvecklingsår än skadeår. Eftersom vi har skattat p_k, som anger sannolikheten att en skada betalas ut k år efter att den rapporterats, kommer vi kunna tillåta skador som rapporterats under det sista kända utvecklingsåret att utvecklas i ytterligare k år.

För att skatta de framtida skadebeloppen för IBNR- och RBNS-skador används nedanstående formler

$$\hat{R}_i^{IBNR} = \mu \sum_{k=0}^{min\{i-m+j-1,d\}} p_k \hat{N}_{i,j-k} \tag{2.27}$$

$$\hat{R}_i^{RBNS} = \mu \sum_{k=i-m+j}^{min\{j,d\}} p_k N_{i,j-k}, \tag{2.28}$$

där de totala IBNR- och RBNS-reserverna tas fram genom att summera över alla i. Den totala reserven ges sedan av summan av IBNR- och RBNS-reserven för varje i, se nedanstående formler

$$\hat{R}^{IBNR} = \sum_{i=1}^{m} \hat{R}_i^{IBNR} \tag{2.29}$$

$$\hat{R}^{RBNS} = \sum_{i=1}^{m} \hat{R}_i^{RBNS} \tag{2.30}$$

$$\hat{R}^{VNJ} = \hat{R}^{RBNS} + \hat{R}^{IBNR} \tag{2.31}$$

Observera att i formel (2.29) och (2.30) kommer vi, till skillnad från den skattade chain ladder-reserven (se formel (2.4)), få en reserv även för det första skadeåret, då $i = 1$. Chain ladder-metoden kommer anta att det första året är färdigutvecklat eftersom metoden inte tillåter fler utvecklingsår än skadeår. VNJ-metoden tillåter däremot flera års utveckling och kommer skatta en reserv även för det första skadeåret.

Eftersom RBNS-reserven sätts upp för skador som redan är rapporterade använder den enbart information om de skador som redan är kända, N_{ij}, medan IBNR-reserven tillhör de skador som ännu ej rapporterats och använder sig istället av de skattade antalet skador som kommer att rapporteras i framtiden, $\hat{N}_{ij}$.

Formeln för att skatta reserven för RBNS-skador, (2.28), säger att då vi vill skatta reserven för ett visst år i vid en viss tidpunkt j tittar vi på antalet skador som har inträffat under år i och multiplicerar det med sannolikheten att antalet år mellan att skadan rapporteras och betalas ut är k år, alltså med p_k. Där k bestäms av antalet år mellan att skadorna rapporterades till tidpunkten vi nu skattar reserverna för. Vi har då fått en skattning av antalet skador som kommer att stängas under

denna tidsperiod ,och multiplicerar sedan det med den skattade medelskadan μ.

På samma sätt kommer IBNR-reserven att skattas genom att titta på de skador vi förväntar oss rapporteras för skadeår som redan skett och multiplicerar dem med sannolikheten att deras slutregleringstid sedan är k år, se formel (2.27).

3 Dataanalys

I detta avsnitt kommer vi att göra en inledande analys av det data som kommer att användas. Vi kommer att undersöka vad för slags beroenden vi kan hitta och se om vi kan dra några slutsatser utifrån detta.

Datat som används i denna rapport kommer ifrån Kammarkollegiet och består av skadedata för en personförsäkring som innehåller alla rapporterade skador för denna försäkring från 15 skadeår (år 2002-2016). Datat innehåller närmare 32 000 unika skador, vi har information om när skadorna skett, när de anmäldes, om skadan är slutreglerad och i sådana fall när den blev slutreglerad. Vi har även information om vilket/vilka belopp som har betalats ut för skadan och när dessa utbetalningar har skett. En stor del av skadorna kommer även ha stängts utan någon utbetalning alls, så kallade nollskador. Ett utdrag ur datat ses i Figur 1.

	NR	skadedatum	rapportdatum	slutdatum	betaldatum	utbetalning	totalut	moms	reserv	regress	skadekostnad
1	000002-16	2013-12-03	2016-01-15	2016-11-02	2016-01-21	130	48561	1413.36	0	0	47147.64
2	000002-16	2013-12-03	2016-01-15	2016-11-02	2016-06-20	4250	48561	1413.36	0	0	47147.64
3	000002-16	2013-12-03	2016-01-15	2016-11-02	2016-06-27	41400	48561	1413.36	0	0	47147.64
4	000002-16	2013-12-03	2016-01-15	2016-11-02	2016-11-03	2781	48561	1413.36	0	0	47147.64
5	000003-13	2012-08-20	2013-01-22	2015-05-18	2013-02-18	2400	2400	0.00	0	0	2400.00
6	000003-14	2013-10-16	2014-01-08	2014-07-15	2014-01-29	3700	3700	0.00	0	0	3700.00
7	000004-14	2013-09-03	2014-01-08	2014-01-21	2014-01-22	3690	3690	0.00	0	0	3690.00
8	000004-16	2014-04-25	2016-01-05	2016-03-03	2016-06-29	391	7770	1494.00	34037	0	40313.00
9	000004-16	2014-04-25	2016-01-05	2016-03-03	2017-01-12	6988	7770	1494.00	34037	0	40313.00
10	000004-16	2014-04-25	2016-01-05	2016-03-03	2017-01-18	391	7770	1494.00	34037	0	40313.00

Figur 1: Ett utdrag ur datat från Kammarkollegiet över en personförsäkring.

I Figur 1 ser vi till exempel att de fyra första observationerna tillhör samma skada och varje rad motsvarar en utbetalning som gjorts för denna skada. I datat finns även en kolumn som heter regress vilket motsvarar en slags justering/återbetalning.

3.1 Brister och justering av data

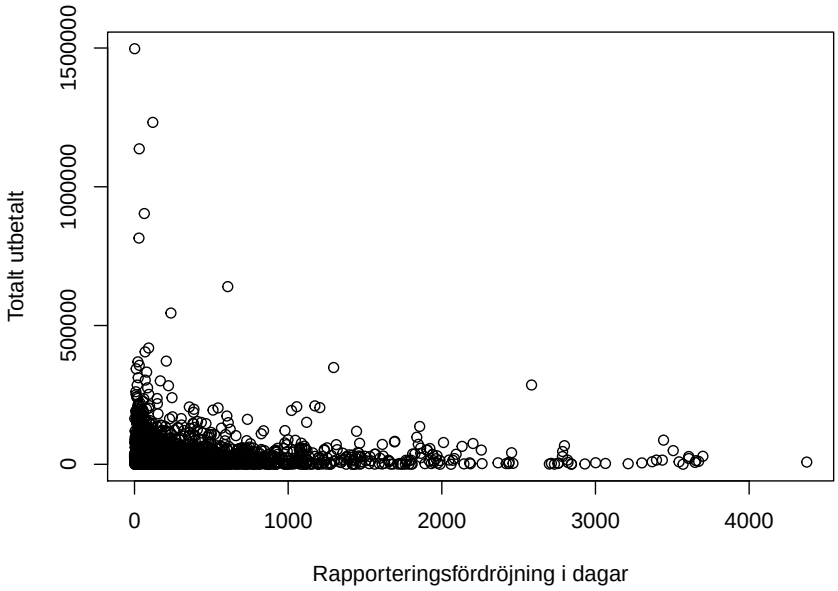
En brist i datat är att vi inte har något volymmått, alltså ingen information om hur stor affär försäkringen utgör och hur detta har förändrats mellan åren. Detta kan bli ett problem om man till exempel vill undersöka hur många rapporterade skador som inkommit under årens gång. Om vi ser en variation mellan år kan vi då inte avgöra om detta beror på att storleken av affären har förändrats eller om det faktiskt finns en förändrad skadefrekvens.

I handläggningsprocessen har man valt att stänga skador då man inväntar ny information, det kan ta ett tag innan det exempelvis går att göra en bedömning om invaliditet. Den skadelidna får vanligtvis höra av sig igen om besvär från skadan finns kvar efter till exempel ett år. Detta innebär att en stor del skador kommer stängas kort efter att de rapporterats utan någon utbetalning. Vissa av dessa skador kommer sedan öppnas igen efter ett antal år och i samband med detta få ett uppdaterat slutregleringsdatum samt en eller flera utbetalningar. Vi ser bara det senaste slutregleringsdatumet och vet därför inte om skadan varit stängd tidigare. Detta leder till att vi kommer observera ett flertal nollskador som egentligen inväntar mer information och kan komma att öppnas igen vid ett senare skede.

Innan vi börjar med analysen kommer mindre justeringar behöva göras. Vissa skador har ett rapporteringsdatum som ligger före skadedatum, eftersom vi inte kan avgöra vilket datum som är korrekt kommer dessa skador att exkluderas, det rör sig om ca 30 stycken utbetalningar vi exkluderar av de drygt 45 000 utbetalningar vårt data består av. I vissa fall summerar inte den totala skadekostnaden till de totala utbetalningarna plus reserv minus regress, i dessa fall justerar vi regressen så att beloppen stämmer överens. Det rör sig om ca 1% av observationerna som behöver denna justering. Vi justerar även beloppen för moms.

3.2 Beroende mellan rapporteringstid och storleken på betalningen

Ett vanligt förekommande problem vid reservsättning är att det kan finnas ett beroende mellan tiden från att en skada inträffat till att den rapporteras (rapporteringsfördröjning) och det utbetalda beloppet. Det innebär exempelvis att större skador som genererar större utbetalningar kan ha en kortare rapporteringsfördröjning då de kan vara lättare att upptäcka. Mindre skador kan vara svåra att upptäcka, samt att det kanske inte heller är självklart att de kommer ge ersättning. Vi undersöker hur förhållandet mellan rapporteringsfördröjningen och skadebeloppen ser ut för vårt data i Figur 2.



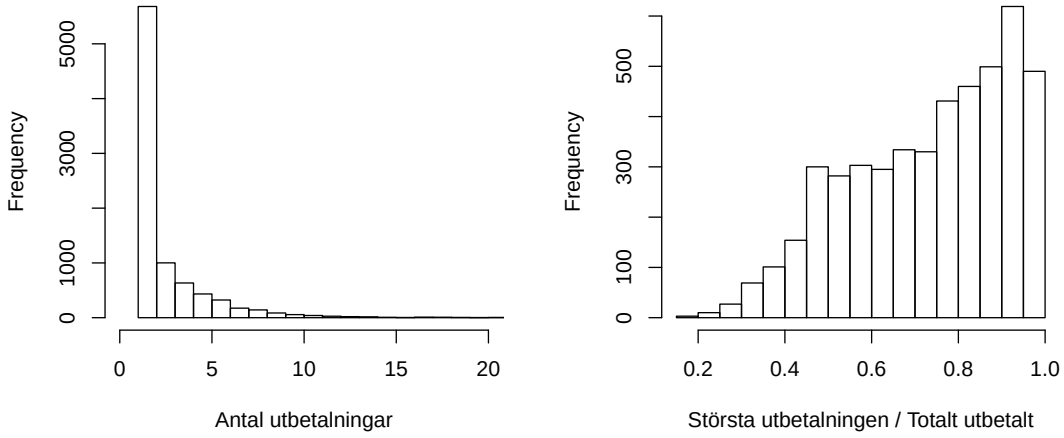
Figur 2: Antal dagar från rapportering till utbetalning plottat mot utbetalt belopp.

Vi ser att de allra flesta skadebelopp ligger långt till vänster i bilden vilket indikerar att de hör ihop med en kort fördröjningstid. Beloppen som har en längre rapporteringsfördröjning och därför ligger längre till höger i bilden tenderar att ha mindre skadebelopp. Bilden kan dock vara lite vilseledande då den kan ha påverkats av högercensurering. Två skador som skett vid samma tidpunkt men som sedan har rapporterats vid olika tidpunkter kommer ha olika långa utbetalningstider fram tills att vi slutat observera vid slutet av år 2016. Om vi enbart tittar på stängda skador kommer dock även detta ge ett liknande resultat.

Som vi nämnde i teorin om VNJ-metoden är ett av antagandena i denna modell att de individuella utbetalningarna inte beror på fördröjningstiderna. Det är kanske inte helt rimligt att säga att detta antagande är uppfyllt efter vad vi sett i Figur 2. Vi ser dock i denna figur att majoriteten av skadorna som inkommit kommer att ha en rapporteringsfördröjning på några få år och ett skadebelopp på några tusen kronor. Vi har sedan några extremfall där skadebeloppen bli väldigt höga alternativt att rapporteringstiden blir väldigt lång. Förhoppningsvis är dessa skador tillräckligt få för att inte ska ha en större inverkan på modellen.

3.3 Storleken på utbetalningar

Vi har data på en finkornig nivå där vi har information om varje utbetalning som skett, när den inträffade och dess storlek. De separata utbetalningarna kan ge ytterligare information till modellen, dock tillåter inte VNJ-metoden vi kommer undersöka fler utbetalningar per skada. Vi undersöker därför nedan hur betalningarna förhåller sig till varandra för att avgöra om vi kan göra något förenklat antagande om dessa.



Figur 3: Bilden till vänster visar ett histogram över antalet utbetalningar per skada (givet minst en utbetalning). Bilden till höger visar ett histogram över förhållandet mellan största skadan och totalt utbetalt per skada (givet fler än en utbetalning).

Båda histogrammen i Figur 3 är enbart baserad på skador som är slutreglerade. I det vänstra histogrammet ser vi antalet utbetalningar per skada. Vi har valt att exkludera de skador med fler än 20 utbetalningar för att få en tydligare bild (färre skador än 0.5% har fler utbetalningar än så, maximalt antal utbetalningar för en skada är 46 st). Vi ser att det allra vanligaste är en utbetalning per skada, ytterst få skador har fler än tio utbetalningar.

Givet att vi har fler än en utbetalning tittar vi på hur storleken på dessa förhåller sig till varandra. I det högra histogrammet i Figur 3 visas den största utbetalningen

delat på totalt utbetalt belopp. Vi ser att det är många skador med flera utbetalningar vars största utbetalning motsvarar över 80% av det totala skadebeloppet.

Histogrammen i Figur 3 visar att det vanligast förekommande är en betalning per skada, samt att då vi har fler än en utbetalning är det vanligt att den största utbetalningen utgör en stor del av det totalt utbetalda beloppet. Detta tyder på att det inte borde ha allt för stor modellpåverkan om vi antar att vi enbart ha en utbetalning per skada.

Vi kommer i modelleringen undersöka flera olika alternativ för att hantera fler utbetalningar per skada.

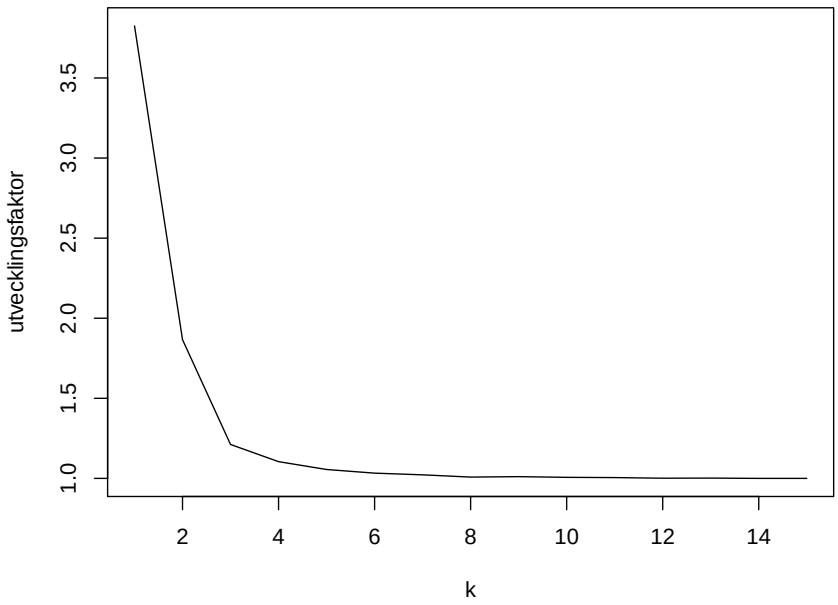
4 Modellering

I detta stycke kommer vi att reservsätta skadorna med både chain ladder-metoden och VNJ-metoden. Vi kommer att skatta reserver med VNJ-metoden då vi gör olika antaganden om utbetalningarna. Resultaten kommer sedan jämföras mot varandra då vi testar hur väl metoderna skattar den faktiska reserven samt titta på volatiliteten i skattningarna, detta kommer testas med hjälp av bootstrap.

Vi kommer att bygga skadetrianglar på perioder om år, raderna representerar alltså olika skadeår och kolumnerna utvecklingsår. Vi kommer att utgå ifrån en signifikansnivå på 5% då vi diskuterar signifikans i våra resultat.

4.1 Reservsättning med chain ladder-metoden

Vi börjar med att skatta våra reserver med chain ladder-metoden. Vi sätter upp en matris med ackumulerade utbetalda belopp, X_{ij}, som representerar det totala beloppet som betalats ut för år i vid år j. Vi baserar skadetriangeln på samtliga 15 år vi har data för. Utvecklingsfaktorer skattas sedan utifrån vår skadetriangel med formel (2.1) som presenterades i teorin. Vi tillåter 15 års utveckling vilket leder till 14 stycken skattade utvecklingsfaktorer, dessa presenteras i Figur 4.



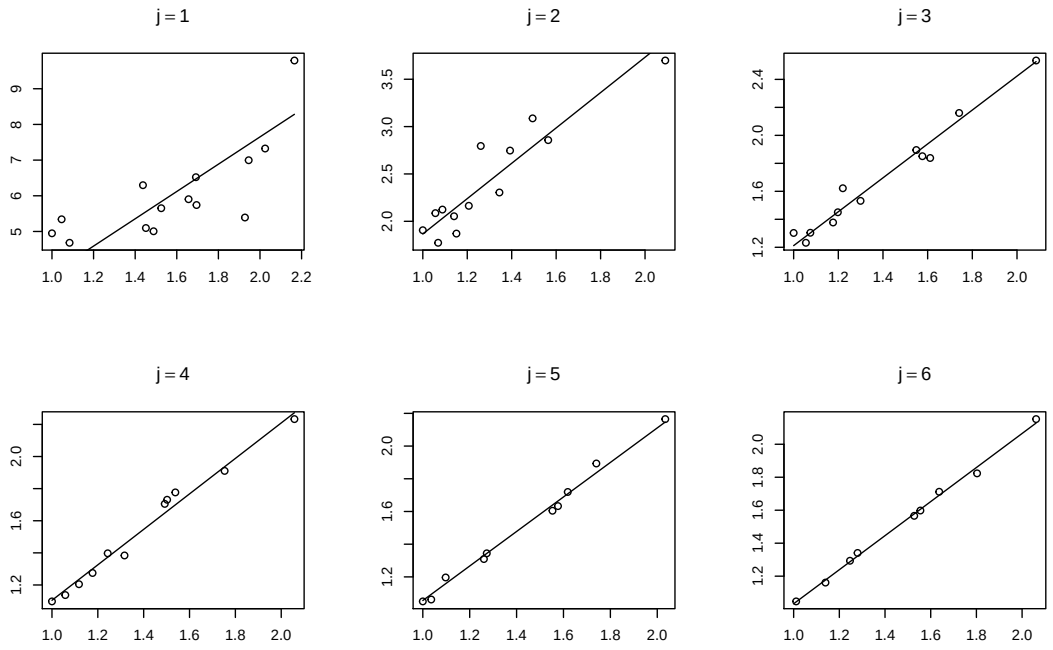
Figur 4: Skattade utvecklingsfaktorerna $\hat{f}_j$ för $j = 1, .., 14$ med chain ladder-metoden.

I Figur 4 ser vi hur de skattade utvecklingsfaktorerna beter sig, vi har en stor utveckling de första åren och att den sedan stannar av efter ca sex år.

Reserven skattas sedan enligt formel (2.3). Reservskattningen kommer jämföras mot de andra metoderna längre ner i arbetet.

4.1.1 Utvärdering chain ladder-metoden

Ett sätt att sedan undersöka om chain ladder-metoden passar för datat är att se om de skattade utvecklingsfaktorerna beskriver den faktiska utvecklingen vi ser för de observerade åren. Vi plottar därför normaliserade ackumulerade årliga utbetalningar, $X_{i,j}$, mot $X_{i,j+1}$ för ett visst utvecklingsår j. Sedan inkluderar vi en linje som har lutningen av utvecklingsfaktorn för år j, f_j, i enlighet med artikel (Mack, 1994). Vi tittar på de första sex utvecklingsåren, alltså då $j = 1, 2, .., 6$, bilden ses i Figur 5.



Figur 5: Normaliserade ackumulerade årliga utbetalningar $X_{i,j}$ plottade mot $X_{i,j+1}$ tillsammans med en linje med lutning av utvecklingsfaktorerna $\hat{f}_j$ för $j = 1, 2, ..., 6$.

Figur 5 visar att de skattade utvecklingsfaktorerna från chain ladder-metoden verkar beskriva den faktiska utvecklingen bra. För det första utvecklingsåret verkar inte linjen riktigt kunna beskriva de faktiska punkterna då de finns en del spridning. För övriga utvecklingsår verkar dock linjen fånga upp mönstret punkterna bildar bra och ligger i samma nivå. Vi ser därför inte något som tyder på att ett intercept skulle behöva inkluderas i modellen.

4.2 Reservsättning med VNJ-metoden

Som tidigare nämnts tillåter VNJ-metoden enbart en utbetalning per skada. Därför inleds detta avsnitt med att diskutera kring vilka olika valmöjligheter som finns för att hantera detta. Vi går sedan in på hur vi använder teorin beskriven ovan om VNJ-metoden för att kunna skatta våra reserver, vi undersöker även osäkerheten i parameterskattningarna.

4.2.1 Hantering av fler utbetalningar per skada

Vanligtvis kan fler utbetalningar per skada ske hos ett försäkringsbolag och detta är, som vi sett i dataanalysen, även fallet i vårt data. Vi behöver därför bestämma hur vi ska hantera detta.

Ett alternativ är att på något sett approximera de olika utbetalningarna till en utbetalning per skada. Detta innebär dock att vi medvetet kommer justera det faktiska datat för att det ska passa in i vår modell. Det andra alternativet är att låta varje enskild utbetalning hamna på dess enskilda utbetalningsår då vi bygger vår skadetriangel. I denna metod ignoreras alltså det faktum att metoden bara tillåter en utbetalning per skada. Detta kommer dock medföra att cellerna med skadebelopp inte motsvaras av antalet skador i samma cell i vår andra skadetriangel.

Inget av dessa alternativ är mer självklart än det andra och vi väljer därför att titta på båda alternativen för att se vilken metod som verkar ge den bästa skattningen. Vidare kommer metoden då vi tillåter fler utbetalningar per skada kallas VNJ-utan approximation då vi i detta fallet inte gör någon approximation utan låter varje enskild utbetalning hamna på dess utbetalningsår. Då approximation till en utbetalning per skada utförs kallas metoden VNJ-med approximation, hur denna approximation görs beskrivs i nästa avsnitt.

4.2.2 Approximation av utbetalningar

Vi börjar med att undersöka vilka valmöjligheter vi har när det kommer till att approximera flera utbetalningar till en utbetalning per skada. Vi vill ta det totala skadebeloppet varje skada genererat och säga att hela beloppet betalas ut vid en viss tidpunkt, frågan är var denna tidpunkt bör vara.

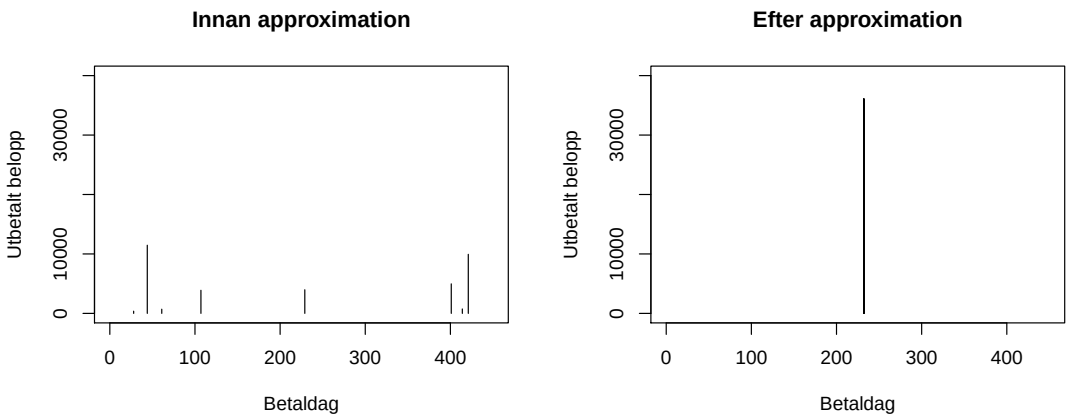
Ett alternativ är att sätta denna tidpunkt som en viktad tid som baseras på storleken på de enskilda utbetalningarna. Vi kommer då alltid att få en tidpunkt som ligger någonstans mellan den första och den sista utbetalningstidpunkten. Men det är inte helt problemfritt att göra en sådan approximation eftersom detta leder till att vi successivt kommer att uppdatera vår historik. Vi kommer säga att vi inte har någon information om skadebeloppet innan skadan är helt stängd, när skadan sedan är helt stängd kommer vi säga att hela skadebeloppet betalades ut vid en tidigare tidpunkt. Om vi bara tittar på en skadetriangel vid en viss tidpunkt kommer detta inte synas, men om vi skulle titta på samma skadetriangel vid olika tidpunkter kommer vi kunna se att celler som tillhör vår historik kommer att ha

uppdaterats i efterhand vilket kan verka lite märkligt.

Ett annat alternativ skulle vara att istället lägga hela skadeutbetalningen när skadan stängs, vilket skulle lösa problemet med att vi uppdaterar historiskt data. Vi får också en korrekt handläggningstid eftersom vi kommer stänga skadan vid den tidpunkten skadan faktiskt stängdes. Detta leder dock till andra problem, exempelvis säger vi då att vi inte har någon information alls om skadebeloppet innan skadan är helt stängd. Det i sin tur leder till att vi i allmänhet kommer att överskatta tiden mellan att en skada rapporterats och att en utbetalning sker.

Vi väljer fortsättningsvis att fokusera på metoden då vi lägger hela skadebeloppet vid en viktad tidpunkt och det är alltså denna metod som kommer kallas VNJ-med approximation. Men vi kommer senare även att titta på hur reservskattningen påverkas om vi hade lagt det totala skadebeloppet vid tidpunkten för den sista betalningen istället.

Den approximativa tidpunkten för varje skada skattas genom att summera hela skadebeloppet för en skada och sedan placera ut den vid en viktad tidpunkt som baseras på storleken av de enskilda utbetalningarna, ett exempel för en utvald skada ses i Figur 6.



Figur 6: Vänster: enskilda utbetalningarna för en skada. Höger: approximerad utbetalning för samma skada.

I den vänstra bilden i Figur 6 ses ett exempel på en skada och dess utbetalningar. Denna skada har åtta utbetalningar av varierande storlek som ligger vid olika tidpunkter, dagarna på x-axeln motsvarar antalet dagar mellan det att skadan har rapporterats och utbetalningen har skett. I den högra bilden ses istället hur

tidpunkten för utbetalningen kommer att ligga efter vår approximation. Där har hela skadebeloppet lagts vid 232 dagar efter att skadan rapporterats.

För att kunna göra en approximation som ovan kan vi enbart titta på stängda skador eftersom det är först då vi har fullständig information om samtliga utbetalningar och det totala skadebeloppet.

4.2.3 Parameterskattningar

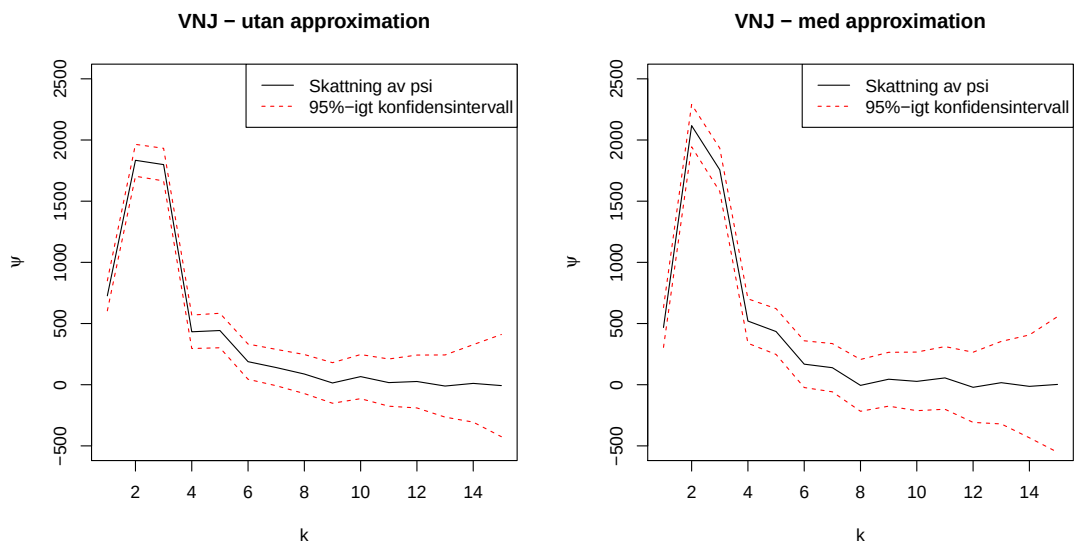
De två olika metoderna för att hantera fler utbetalningar kommer enbart att påverka uppbyggnaden av skadetriangeln bestående av utbetalade skadebelopp. När vi använder oss av VNJ-utan approximation kommer $X_{i,j}$ att byggas upp på samma sätt som med chain ladder-metoden, förutom att den kommer att bestå av inkrementella värden. När VNJ-med approximation används kommer $X_{i,j}$ istället att byggas upp av enbart stängda skador. Skadetriangeln med antalet skador, $N_{i,j}$, skapas på samma sätt oavsett vilken av metoderna som används.

För att skatta ψ_k sätts en generaliserad linjär modell upp i programvaran `R` (R Core Team, 2016), vi antar en kvasilikelihood och en identitetslänk. Identitetslänken säger att medelvärdet μ och den linjära skattaren är identiska. Vi har i detta fall inget som hindrar skattningarna av ψ_k att bli negativa, vilket i sin tur leder till att vi kan få negativa skattningar av p_k. Får vi negativa värden på p_k kan detta vara ett tecken på att modellanpassningen inte är bra då vi i detta fallet säger att sannolikheten att en skada betalas ut k år efter att den rapporterats är negativ. Det är mer troligt att större k skulle bli negativa då dessa baseras på mindre data och då de ligger längre i framtiden får vi en större osäkerhet.

Med en generaliserad linjär modell kan vi få ut skattningar av regressionskoeffecienterna, $\hat{\psi}_k$, vilket även ger oss skattningar av μ och p_k. Då har vi alla parametrar som behövs för att kunna skatta IBNR- och RBNS-reserven enligt formlerna (2.27) och (2.28).

4.2.4 Analys av parameterskattningar

För skattningarna av ψ_k får vi en del osignifikanta resultat, speciellt för större k. För VNJ-utan approximation är endast de fem första skattningarna av ψ signifikanta. För VNJ-med approximation är det istället de första sex skattningarna som är signifikanta. Skattningarna av ψ för de olika metoderna ses tillsammans med konfidensintervall i Figur 7.



Figur 7: Skattningar av ψ för samtliga skadeår med VNJ-utan approximation (vänster) och VNJ-med approximation (höger), tillsammans med 95%-iga konfidensintervall.

Här har konfidensintervallen för ψ beräknats med följande formel:

$$KI(\psi_k) = \hat{\psi}_k \pm Z_{1-\alpha/2} SE(\hat{\psi}_k) \tag{4.1}$$

I Figur 7 ser vi att båda metoderna ger liknande mönster när det kommer till skattningarna av ψ och dess konfidensintervall. Det finns en stor osäkerhet i skattningarna av ψ_k för större värde på k för båda metoderna. Konfidensintervallen kommer, för större värden på k, sträcka sig under noll vilket tyder på att negativa skattningar kan förekomma. Annars ses att skattningarna planar ut efter cirka åtta år vilket tyder på att vi inte förväntar oss någon utveckling efter det. Konfidensintervallet är aningen bredare för VNJ-med approximation och vi ser även att skattningarna är mer volatila för de senare åren. Båda metoderna ger i övrigt liknande resultat, vi får exempelvis medelskador på 5 709 (utan approximation) och 5 767 (med approximation). Spridningsparametern kan skattas på samma sätt som beskrivet på s.11 i (Miranda *et al.* , 2012), det vill säga enligt nedan

$$\hat{\rho} = \frac{1}{n-(d+1)} \sum_{i,j} \frac{(X_{ij} - \hat{X}_{ij}^{VNJ})^2}{\hat{X}_{ij}^{VNJ} \hat{\gamma}_i} \tag{4.2}$$

där $n = m(m+1)/2$ och $\hat{\gamma}_i$ är en inflationsparameter som vi kommer sätta konstant till 1. Vidare utgör $\hat{X}_{ij}^{VNJ}$ VNJ-skattningen av $E[X_{ij}|\aleph_m]$. Då fler utbetalningar per skada tillåts skattas spridningsparametern till 109 044, när approximation till en utbetalning per skada gjorts blir skattningen istället 159 466. Med dessa resultat kan vi ta fram en variansskattning enligt nedanstående formel presenterad i (Miranda *et al.* , 2012)

$$\hat{\sigma}^2 = \hat{\mu}\hat{\rho} - \hat{\mu}^2 \tag{4.3}$$

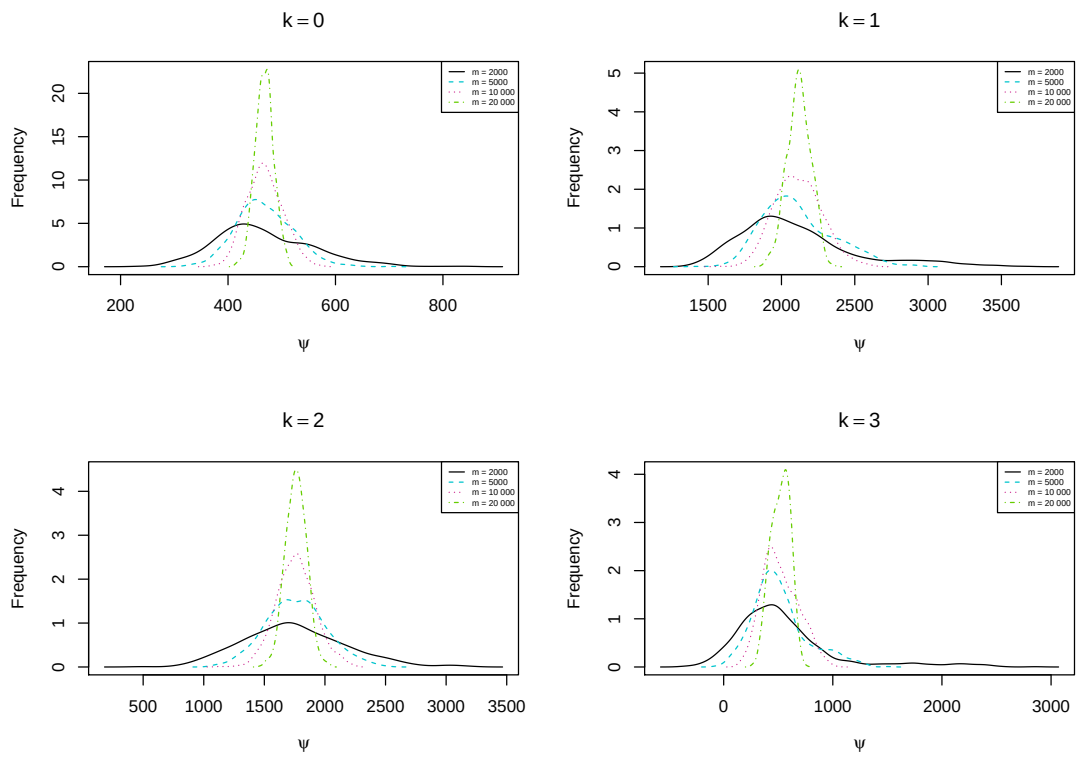
Variansskattningen för VNJ-utan approximation beräknad enligt formel ovan blir 595 631 401 och med VNJ-med approximation fås 877 855 087. Variansen blir alltså mindre då vi inte har gjort någon approximation på utbetalningarna. Som nämnts tidigare byggs VNJ-med approximation upp enbart av stängda skador. Denna metod kommer därför få en lägre genomsnittlig regleringstid då den inte fångar upp de skador som har som längst regleringstid och därmed fortfarande är öppna. I dataanalysen sågs även att de skador som genererar stora skadebelopp i regler har kortare rapporteringstid. Detta kan vara en anledning till att VNJ-med approximation genererar högre variansskattning.

4.2.5 Osäkerhet i skattningen av ψ

En närmare titt tas nu på några av de ψ som vi fick signifikanta skattningar för. Vi är intresserade av att undersöka hur stor variation vi får i skattningarna då vi har färre observationer att basera skattningar på. Detta är ett viktigt test för att kunna avgöra om metoden även lämpar sig för en försäkringsprodukt av mindre storlek än den personförsäkring med drygt 30 000 skador vi hittills har tittat på.

Vi tittar på data bestående av 2 000, 5 000, 10 000 samt 20 000 skador, dessa skador samplas från vår personförsäkring. Vi samplar skador 1 000 gånger för varje stickprovsstorlek och för varje sample skattar vi en uppsättning ψ:n.

Vi tittar sedan på resultatet av skattningarna för de fyra första ψ:na, då $k = 0, 1, 2, 3$. Detta görs för båda VNJ-metoderna som diskuterats ovan, men då resultaten blir så pass lika väljer vi att enbart visa resultatet från skattningarna med VNJ-utan approximation. Resultatet ses i Figur 8.



Figur 8: Skattning av ψ_k för $k = 0, 1, 2, 3$ baserat på 1000 skattningar för stickprovsstorleken $m = 2\ 000$, 5 000, 10 000, 20 000. Skattningarna är framtagna med VNJ-utan approximation.

Figur 8 visar att en ökad stickprovsstorlek ger mindre variation i skattningarna, vilket är förväntat. Vi ser att skattningarna ligger lägre för $k = 0$ vilket är rimligt då dessa skattningar hör ihop med skador som blir helt utbetalda samma år som de rapporterats. I regel kommer inte den tiden kunna vara ett helt år då skador kommer rapporteras en bit in på året vilket då rimligtvis bör ge en lägre skattning av ψ_0 än för övriga k.

Vid framtagning av skattningar presenterade i Figur 8 med VNJ-med approximation kommer spridningen i skattningarna bli aningen större än resultatet som visas ovan.

4.2.6 Införande av intercept

En egenskap som VNJ-metoden har är att den tillåter oss att införa ett intercept. Detta kan vara behjälpligt om vi till exempel i Figur 5 hade sett en konstant

underskattning av reserverna, den skattade linjen i figuren hade då legat under punkterna som baseras på den faktiska utvecklingen. Detta hade varit ett tecken på att ett intercept behöver införas, vilket inte är möjligt i chain ladder-metoden.

Detta görs genom att inkludera ett intercept i den generaliserade linjära modellen när vi skattar parametrarna, den kommer nu ställas upp på formen nedan istället

$$\begin{bmatrix} X_{10} \\ X_{20} \\ \vdots \\ X_{m0} \\ X_{11} \\ \vdots \\ X_{m-1,1} \\ \vdots \\ X_{1,m-1} \end{bmatrix} = \begin{bmatrix} 1 & \dots & 1 & N_{10} & 0 & 0 & \dots & 0 \\ 1 & \dots & 1 & N_{20} & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & N_{m0} & 0 & 0 & \dots & 0 \\ 1 & \dots & 1 & N_{11} & N_{10} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \dots & 1 & N_{1,m-1} & N_{1,m-2} & N_{1,m-3} & \dots & N_{1,0} \end{bmatrix} \cdot \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_{m-1} \\ \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{m-1} \end{bmatrix}$$

Där α_j är interceptet för utvecklingsår j och motsvarar den lägsta möjliga betalningen vi förväntar oss.

Då ett intercept inkluderas i vårat fall blir dock samtliga skattningar av α_j osignifikanta, resultatet är alltså i linje med det vi fick med chain ladder-metoden (se Figur 5) och det verkar därför inte finnas ett behov av att införa ett intercept. Det är dock en värdefull egenskap att modellen tillåter införande av ett intercept. Notera att då ett intercept införs påverkar detta inte bara parameterskattningarna utan även skattningarna av reserverna, α_j skulle därför även behöva inkluderas i formlerna (2.27) och (2.28).

4.3 Modellutvärdering av reserversättningsmetoder

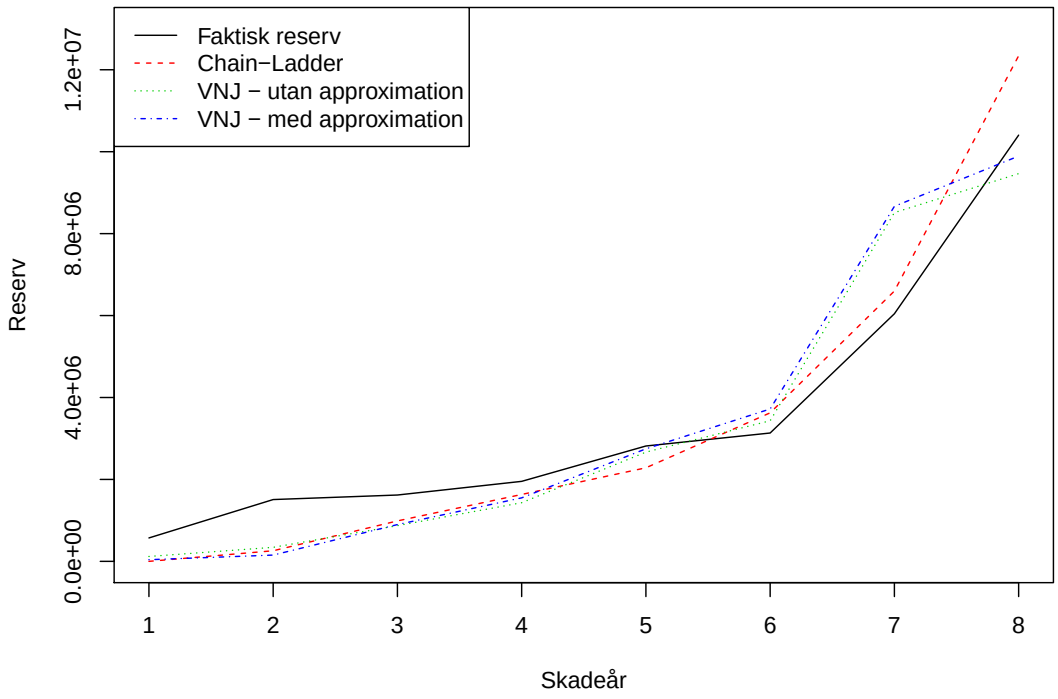
Vi har undersökt tre olika metoder för att skatta reserven. Chain ladder-metoden, samt två olika implementationer av VNJ-metoden, där flera utbetalningar per skada har hanterats på två olika sätt. I det här avsnitt jämförs dessa metoder och vi försöker ta reda på vilken som är bäst lämpad att använda på Kammarkollegiets produkter.

4.3.1 Jämförelse mot faktisk reserv

Eftersom vi har data från 15 år och med tanke på utvecklingsmönstret vi sett tidigare, kan vi anta att några av de första åren kan ses som färdigutvecklade. Ett test för att se hur bra de olika metoderna verkar passa vårt data är att beräkna vad den faktiska reserven blev för de första åtta skadeåren (år 2002-2009). Vi antar sedan att vi bara har data för den övre triangeln för de första åtta skadeåren och skattar reserver med hjälp av de olika metoderna. Resultaten jämförs sedan mot den faktiska reserven.

Chain ladder-metoden kommer inte beräkna någon reserv för det första skadeåret. Detta eftersom metoden inte tillåter fler utvecklingsår än skadeår och kommer därför anta att det första skadeåret är färdigutvecklat. VNJ-metoden kommer istället att kunna räkna med upp till femton års utveckling för det första skadeåret. Men, som ses i Figur 7, kommer skattningarna av ψ_k för $k > 8$ ligga väldigt nära noll vilket tyder på att femton års utveckling kanske inte är nödvändigt. I det faktiska datat ses att det fortfarande finns en viss utveckling om än liten för det första skadeåret efter åtta utvecklingsår, vilket då är i linje med VNJ-skattningarna.

I Figur 9 ser vi hur den faktiska reserven såg ut i slutet av år 2009 för skador som skett under 2002-2009, det går även att se hur den skattade reserven ser ut då den skattats med chain ladder-metoden samt de två olika implementationerna av VNJ-metoden.



Figur 9: Den faktiska reserven för skadeår 2002-2009 vid årskiftet 2009/2010 samt reserven för samma period skattad med tre olika metoder.

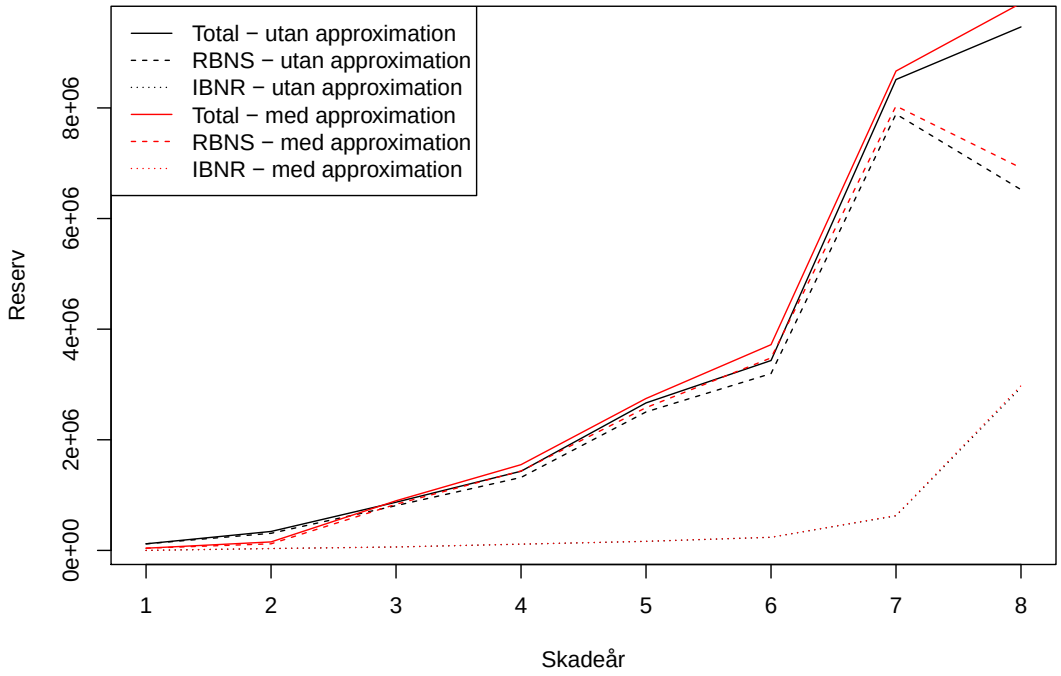
Samtliga metoder underskattar reserven för de tre första skadeåren och följer sedan den faktiska reserven för de tre efterföljande skadeåren. Chain ladder-metoden ger en bra skattningen även för det sjunde året men överskattar reserven för det sista skadeåret, för VNJ-metoden gäller tvärtom för de två sista skadeåren. De två olika sätten att hantera flera utbetalningar i VNJ-metoden ger snarlika resultat. VNJ-utan approximation ger skattningar som ligger aningen närmare den faktiska reserven för de flesta skadeåren. Skattningarna här är dock såpass lika att det är svårt att säga att någon av dessa två metoder skulle vara bättre än den andra. Rent grafiskt är det svårt att avgöra vilken av de tre metoderna som följer reserven bäst, därför studeras medelkvadratfelet (MSE) för de olika metoderna, vilket beräknas enligt

$$MSE(\hat{R}^l) = \frac{1}{m}\sum_{i=1}^{m}(R_i - \hat{R}_i^l)^2 \tag{4.4}$$

där R är den faktiska reserven, $\hat{R}^l$ är skattningen där l anger vilket metod. Då

medelkvadratfelet beräknas för de olika metoderna visar det, precis som Figur 9, att skillnaden mellan de olika metodernas avvikelse från den faktiska reserven är små. Chain ladder-metoden ger dock minst avvikelse från den faktiska reserven, följt av VNJ-utan approximation och sist VNJ-med approximation.

Som tidigare nämnts är en fördel med VNJ-metoden att den kan skatta RBNS- och IBNR-reserven separat. Skattningarna kan därför analyseras var för sig, vi studerar därför skattningarna av reserven för båda VNJ-metoderna som visas i Figur 9 uppdelade mellan reserv för kända och okända skador i Figur 10.



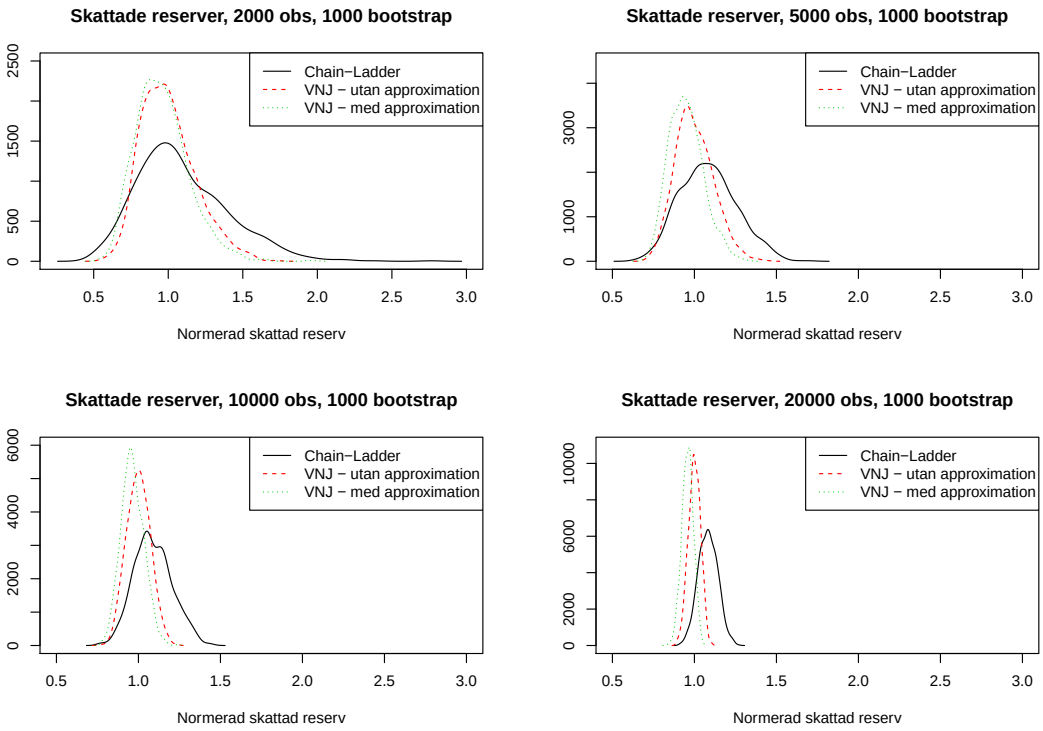
Figur 10: Totala skattningen för de första åtta skadeåren med VNJ-metoden, samt skattningarna uppdelade för IBNR och RBNS.

Som vi ser i Figur 10 utgör RBNS-reserven den absolut största delen av den totala reserven. Modellen fångar upp att skador rapporteras relativt snabbt och att majoriteten av inträffade skador har rapporterats efter några få utvecklingsår. För det sista skadeåret predikterar modellen att ett flertal skador inte har rapporterats än, vi får då en ökad reserv för okända skador samtidigt som vi sätter av mindre för kända skador.

4.4 Stabilitetstest reserver

För att en metod ska vara bra vill vi att den ska ge stabila skattningar. Som nämnts tidigare har försäkringsprodukten vi undersöker en hel del observationer. Detta kommer inte alltid att vara fallet, ibland kommer vi att vilja analysera exempelvis en nystartad produkt där inte mycket data finns att tillgå, alternativt en produkt som av någon annan anledning innehåller få observationer. Detta test ämnar undersöka hur pass stabila resultat våra tre metoder producerar och även se hur känsliga modellerna är då vi minskar antalet observationer i datat.

För att undersöka detta använder vi bootstrap, vi samplar data av en viss stickprovsstorlek och skattar sedan reserverna för samtliga metoder. För varje stickprovsstorlek och metod gör vi sedan detta 1 000 gånger och undersöker variationen i skattningarna vi får ut. Vi gör detta för fyra olika stickprovsstorlekar: 2 000, 5 000, 10 000 och 20 000. Vi normerar resultaten genom att för varje stickprovsstorlek dela samtliga skattade reserver med medelvärdet av alla bootstrappade skattningar av VNJ-metoden (utan approximation) för samma stickprovsstorlek. Detta leder till att resultaten för de olika stickprovsstorlekarna kommer att vara jämförbara då variationen inte kommer att bero på att ett större sample ger större reserver. Resultatet ses i Figur 11.



Figur 11: Normaliserade reserver för samtliga skadeår skattad 1 000 gånger med tre olika metoder, för stickprovsstorlek 2 000, 5 000, 10 000 och 20 000.

I Figur 11 ser vi att skattningarna blir stabilare för samtliga metoder då stickprovsstorleken ökas, vilket är förväntat. För samtliga storlekar på stickprovet kommer chain ladder-metoden att ha störst spridning i dess reservskattningar. För en mindre stickprovsstorlek ger de två olika varianterna av VNJ-metoden liknande resultat. För en större stickprovsstorlek verkar dock den implementation av VNJ-med approximation ge ett mer stabilt resultat. Vi observerar även att VNJ-metoderna är centrerade runt ett för samtliga stickprovsstorlekar samtidig som chain ladder-metodens centrering flyttas mer åt höger för större stickprov. Det verkar alltså som att chain ladder-metoden systematiskt skattar en högre reserv då stickprovsstorleken ökar. För mindre stickprovsstorlek får fördelningen för chain ladder skattningarna även en lång svans åt höger.

4.5 Robusthetsanalys

Något som är viktigt att komma ihåg är att en metod kan passa bra för en typ av data, men inte alls för en annan. Som tidigare nämnts är syftet att hitta en re-

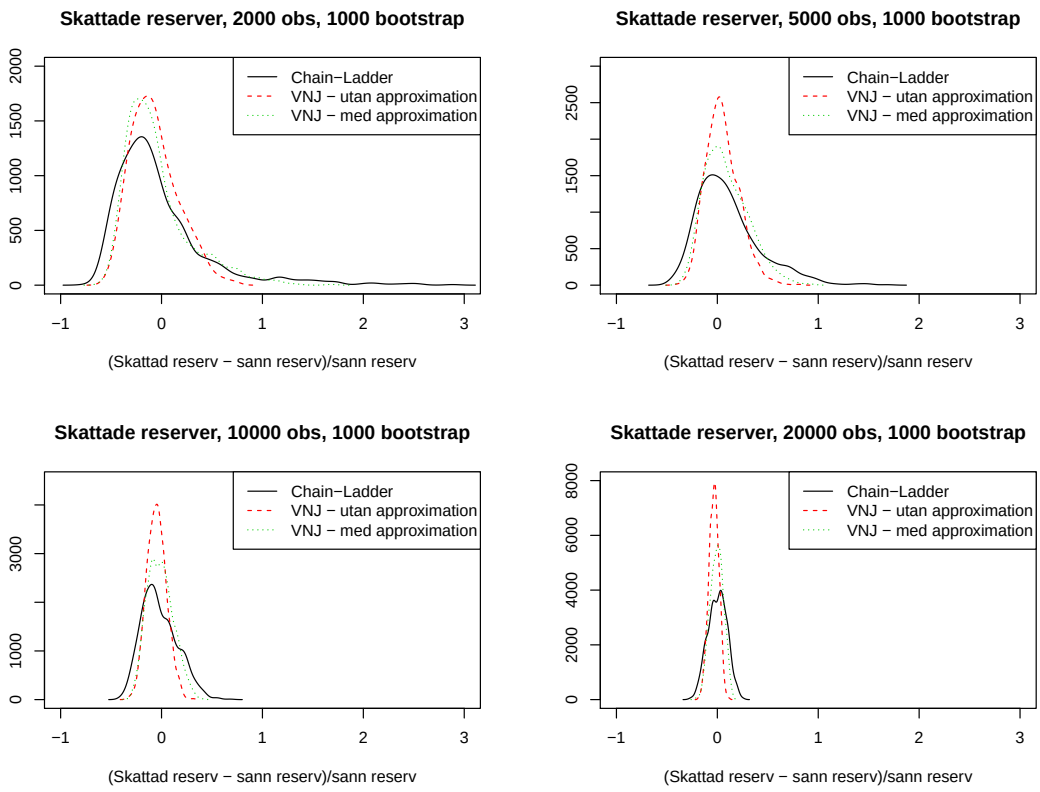
servsättningsmetod som kan passa fler av Kammarkollegiets försäkringsprodukter utan att det ska kräva allt för mycket manuellt arbete. Det är därför av intresse att undersöka hur reservskattningarna för de olika metoderna förändras då små justeringar i data görs.

I detta avsnitt kommer det därför undersökas hur mycket våra reservskattningar avviker från den faktiska reserven då vi antar att vårt data innehåller färre observationer. Vi kommer även titta på hur skattningarna för de olika metoderna påverkas då vi manipulerar data så att vi inte längre kan anta en dominant utbetalning per skada. Ett test utförs också då vi antar att en större del av skadorna har fler än en utbetalning.

4.5.1 Avvikelse från faktisk reserv för olika stickprovsstorlekar

Vi undersökte tidigare hur väl de olika metoderna kan skatta den faktiska reserven för de första åtta åren. Vi vill nu se hur mycket reservskattningarna för de första åtta skadeåren avviker från den faktiska reserven för olika stickprovsstorlekar, vi använder samma bootstrapmetod som användes för stabilitetstestet i avsnitt 4.4. Den faktiska reserven beräknas för de åtta första skadeåren för samma stickprovsstorlek. De skattade reserverna normeras sedan genom att subtrahera den faktiska reserven för samma stickprovsstorlek och resultatet delas sedan med den faktiska reserven.

Återigen skattar vi reserven 1 000 gånger för varje metod och stickprovsstorlek. Vi använder oss av samma storlekar som tidigare, vi drar från alla observationer och det är sedan de skadorna som inträffade under första åtta skadeåren som skattningarna baseras på. Resultatet ses i Figur 12.



Figur 12: Normaliserade reserver för de åtta första skadeåren skattade 1 000 gånger med tre olika metoder minus den faktiska reserven. För stickprovsstorlekar 2 000, 5 000, 10 000 och 20 000.

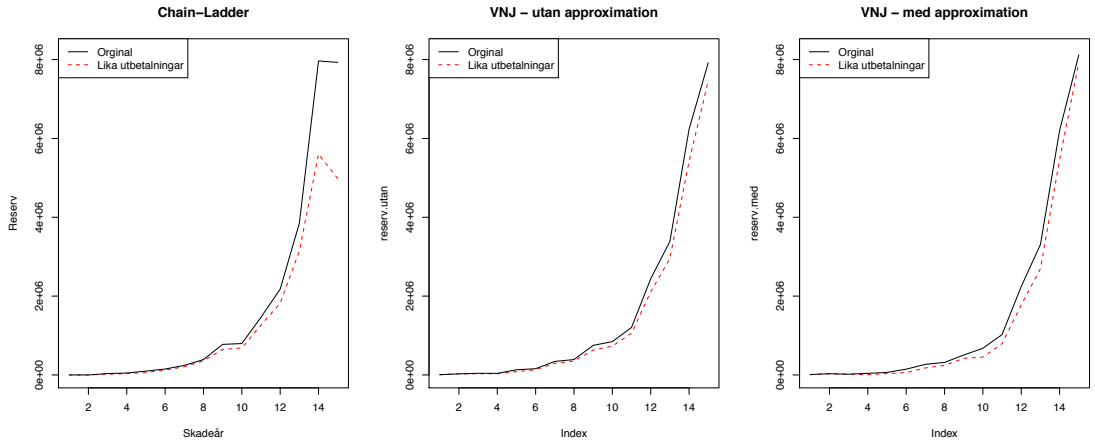
I Figur 12 ser vi att skattningarna vi får från chain ladder-metoden ger störst spridning för samtliga stickprovsstorlekar och tenderar alltså att avvika mest från den faktiska reserven. Mest stabilt resultat får vi från VNJ-utan approximation. Även här ser vi att för små stickprovsstorlekar får chain ladder-metoden en svans åt höger. Vilket innebär att metoden kan ge grova överskattningar av reserven då vi har få observationer.

4.5.2 Reservskattningar då vi inte har en dominant utbetalning

Tidigare i modelleringen bestämde vi oss för att approximera data till en utbetalning per skada för VNJ-metoden, och gjorde då ett antagande om en viktad tidpunkt då det totala skadebeloppet betalades ut. Detta verkade som ett rimligt antagande då vi i dataanalysen såg att det var vanligt förekommande att en av utbetalningarna utgjorde en betydande del av det totala skadebeloppet för varje skada. Vi är nu intresserade av att testa vad som sker då vi inte kan anta en do-

minant utbetalning men ändå gör samma approximation om ovan.

Vi manipulerar vårt data och säger att alla utbetalningar för en skada är lika stora. Vi fördelar alltså det totala skadebeloppet för varje skada jämnt mellan dess utbetalningstidpunkter. Vi skattar sedan reserven för de tre olika metoderna vi diskuterat tidigare. Resultatet presenteras i Figur 13 och 14 nedan.



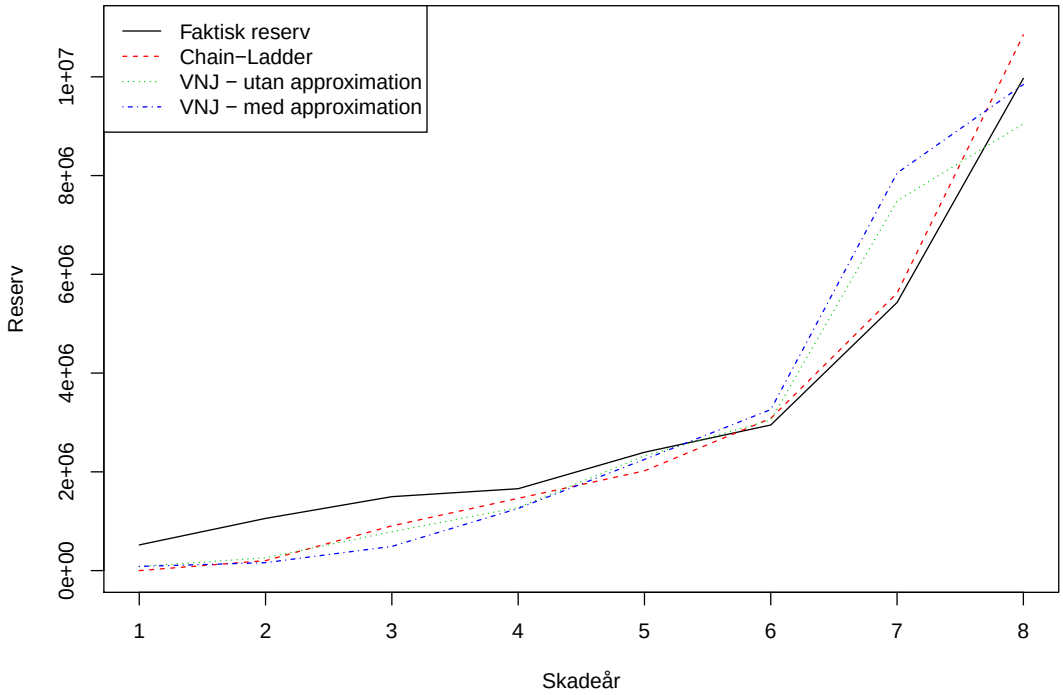
Figur 13: Reservskattningar med tre olika metoder jämfört mot reservskattningar med samma metoder då vi antagit lika fördelad utbetalningsstorlek.

I Figur 13 ser vi reservskattningarna för de tre olika metoderna då vi antagit att alla utbetalningar för en skada är av samma storlek. Vi jämför detta mot de ursprungliga reservskattningarna. För båda anpassningarna av VNJ-metoden blir skattningarna aningen lägre då vi antar att alla utbetalningar har samma storlek. Skattningarna är dock väldigt lika de vi fick ursprungligen.

För chain ladder-metoden får vi ett liknande resultat för de första tretton åren som vi fick med båda VNJ-metoderna, en skattning som ligger strax under den ursprungliga. För de två sista skadeåren får vi dock en betydligt lägre skattning när det antas att alla utbetalningar är av samma storlek. När vi antar att alla utbetalningar för samma skada har samma storlek leder detta till, för vårt data, att större del av det totala skadebeloppet betalas ut under tidigare utvecklingsår än tidigare. De första skattade utvecklingsfaktorerna blir därför mindre för denna metod vilket leder till att de sista skadeåren inte får lika stor utveckling som med våra ursprungliga skattningar.

Vi vill även se hur dessa reservskattningar förhåller sig till den faktiska reserven.

Vi beräknar den faktiska reserven för de första åtta åren när vi har manipulerat utbetalningarna till att vara lika stora. Dessa jämförs mot skattningarna vi fick fram med de olika metoderna i Figur 14.



Figur 14: Reservskattningar med tre olika metoder för de första åtta skadeåren jämfört mot den faktiska reserven när vi antagit att alla utbetalningar är lika stora.

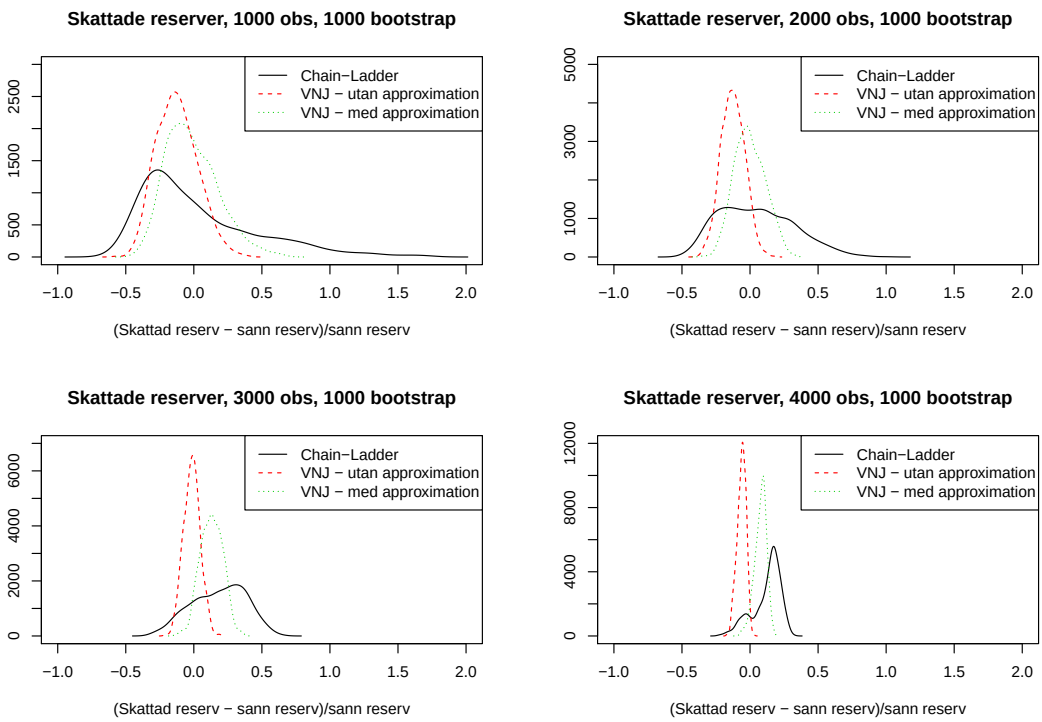
I Figur 14 ser vi att resultatet blir snarlikt det vi såg då vi jämförde våra ursprungliga skattningar mot den faktiska reserven för de första åtta skadeåren, se 9. Chain ladder-metoden verkar inte få samma effekt för det sista skadeåret då vi enbart tittar på de åtta första åren.

4.5.3 Större andel skador med fler utbetalningar

I tidigare analyser har de två olika implementationer av VNJ-metoden visat liknande resultat. Eftersom metoderna producerat så lika resultat verkar det gå bra att ignorera det faktum att modellen egentligen inte kan hantera fler utbetalningar

per skada. En anledning till de snarlika resultaten kan vara att många av skadorna enbart har en utbetalning och de skadorna kommer behandlas lika i båda metoderna. Därför studeras nu vad som sker om en större andel av skadorna antas ha flera utbetalningar än en.

Vi antar nu att vårt data enbart består av de skador som har fler än en utbetalning och gör sedan samma analys som i Figur 12. Vi drar alltså stickprov från data, denna gång av storlek 1 000, 2 000, 3 000 och 4 000 då vi har så pass få unika skador med fler än en utbetalning. Det hade kanske varit bättre att sampla med återläggning men detta blev för beräkningstungt, dock bör även mindre stickprovsstorlekar ge en indikation på hur modellerna skiljer sig åt då vi har större andel skador med fler utbetalningar än en. Varje stickprovsstorlek dras 1 000 gånger och en reserv skattas för varje metod för de första åtta skadeåren. Dessa jämförs sedan mot den faktiska reserven, resultatet ses i Figur 15.



Figur 15: Normaliserade reserver för de första åtta skadeåren skattade 1 000 gånger med tre olika metoder minus den faktiska reserven då vi antagit att alla skador har fler än en utbetalning. För stickprovsstorlekar 1 000, 2 000, 3 000 och 4 000.

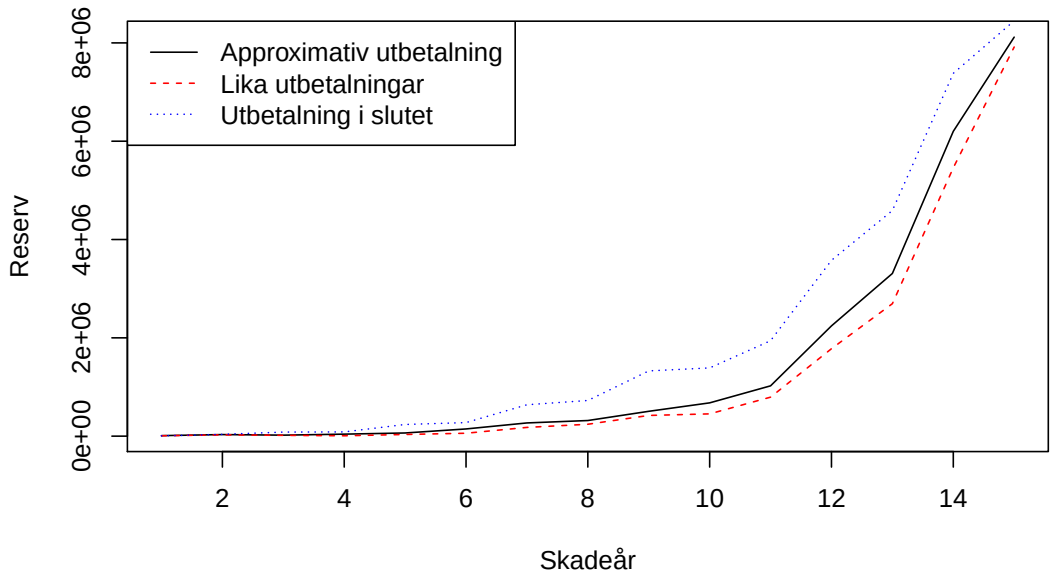
I Figur 15 ser vi en större skillnad mellan de båda VNJ-metoderna än vad vi sett

tidigare. VNJ-utan approximation har fortfarande minst spridning i dess skattningar men ger i allmänhet lägre skattningar än VNJ-med approximation. För mindre stickprovsstorlekar är VNJ-med approximation den metod som är mest centrerad runt noll, det vill säga den metod som ger skattningar mest likt den faktiska reserven. Chain ladder-metoden har fortfarande störst spridning för samtliga stickprovsstorlekar och tenderar att överskatta reserven.

4.6 Olika metoder för att approximera utbetalningar

Avslutningsvis vill vi återkoppla till approximationen av flera utbetalningar per skada. I modelleringen diskuterade vi kring hur man kan approximera flera utbetalningar per skada och valde då att gå vidare med metoden då vi placerade den totala skadekostnaden vid en viktad tidpunkt. I detta avsnitt väljer vi att titta på hur reservskattningen hade förändrats om vi istället valt en annan metod för denna approximation.

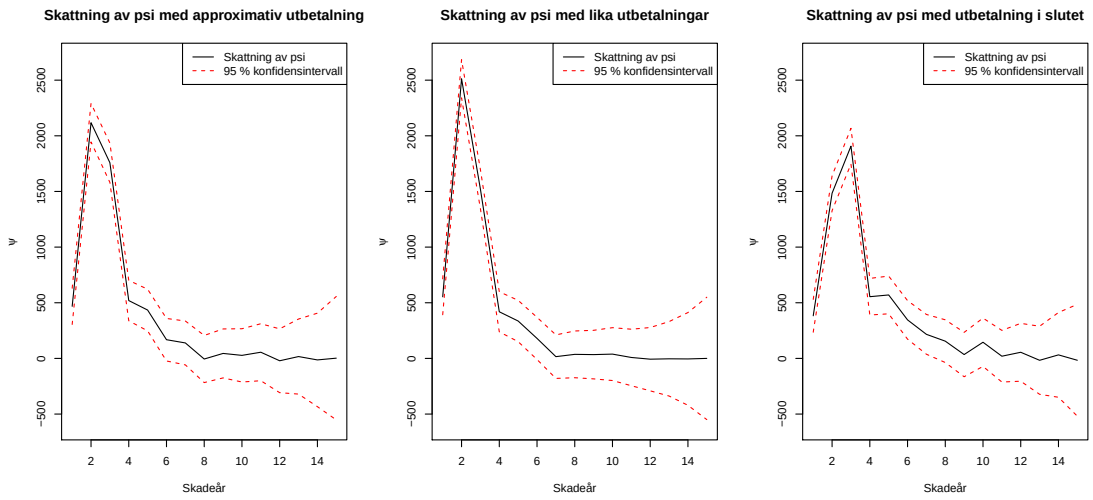
Vi väljer nu att placera det totala utbetalade beloppet vid den sista utbetalningen för varje skada. Vi undviker då att uppdatera vår historik och får även en korrekt handläggningstid. Denna skattning jämförs sedan mot VNJ-med approximation. Resultatet ses i Figur 16, där inkluderas även reservskattningen då vi antagit lika utbetalningar som vi tog fram i avsnitt 4.5.2.



Figur 16: Reservskattning för olika antaganden av utbetalningstidpunkt.

I Figur 16 ser vi att reservskattningen ökar en hel del för de senare skadeåren när vi istället lägger det totala skadebeloppet vid den sista utbetalningen. Att vi senarelägger utbetalningar leder till att vi förväntar oss längre handläggningstid än tidigare, vilket innebär att vi förväntar oss att mer ska betalas ut i framtiden vilket rimligtvis bör öka reservskattningarna. VI har tidigare sett att skattningarna med VNJ-metoden har legat i linje med den faktiska reserven. Resultatet i 16 tyder därför på att då vi lägger hela skadebeloppet vid den sista utbetalningen överskattas reserven.

Vi kan även se hur skattningarna av ψ och dess konfidensintervall kommer att se ut när vi har använt olika metoder för att approximera till en utbetalning per skada, se Figur 17.



Figur 17: Skattningar av ψ tillsammans med 95 %-iga konfindensintervall för olika antaganden av utbetalningstidpunkt.

I Figur 17 ses att skattningarna av ψ, för samtliga sätt att hantera utbetalningarna, närmar sig noll då $k \geq 8$. Likt resultatet i Figur 7 förväntar vi oss ingen större utveckling åtta år efter att en skada rapporterats. För metoden då lika utbetalningar antagits kommer skattningarna inte fluktuera lika mycket runt noll för större värden på k som med de andra metoderna. Vilket är ett mer rimligt mönster då vi förväntar oss att utvecklingen slutar efter ett visst antal utvecklingsår och då är konstant noll, inte att vartannat år har en negativ sannolikhet.

5 Diskussion

I detta arbete har vi tittat på två reservsättningsmetoder med målet att hitta en metod som lämpar sig för att reservsätta en av Kammarkollegiets försäkringsprodukter. Syftet är att hitta en metod som använder lättillgänglig data och som går att anpassa på flera olika försäkringsprodukter utan alltför mycket manuellt arbete, samtidigt som den ger stabila skattningar av reserven. Önskvärt är också att metoden är lätt att implementera. Vi kommer nedan att diskutera kring hur våra studerade reservsättningsmetoder förhåller sig till dessa kriterier utifrån resultaten som har framkommit i modelleringen.

5.1 Reservsättningsmetoderna

Vi har tittat på den klassiska och välanvända chain ladder-metoden, som har jämförts med VNJ-metoden som producerar separata skattningar av IBNR- och RBNS-reserverna tack vare att den betingar på antalet skador.

Av dessa metoder är chain ladder-metoden den som är lättast att implementera eftersom den enbart baserar sina skattningar på en aggregerad skadetriangel. Denna metod är också lätt att anpassa på flera produkter utan att göra omfattande manuella justeringar, vilket är en av anledningarna till att denna metod används i så stor utsträckning idag. Nackdelen med denna modells simplicitet märks då mer data finns att tillgå, modellen kommer inte att kunna fånga upp detta och värdefull information slängs eventuellt bort. Denna metod tillåter inte att man skattar reserver för fler utvecklingsår än antalet skadeår och den producerar inte heller separata skattningar för kända och okända skador.

VNJ-metoden är också en relativt enkel metod att implementera. Den använder två aggregerade skadetrianglar istället för en, samt en generaliserad linjär modell för att skatta parametrarna. Fördelen med att skatta parametrar med en GLM (i exempelvis `R`) är att man får ut mer information om de underliggande parametrarna. Då har man möjligheten att göra fler analyser och att lättare få förståelse över variationer i resultaten. Eftersom metoden även producerar separata skattningar för kända och okända skador kan dessa analyseras var för sig för att få bättre förståelse om skattningarna. I denna metod har man även möjlighet att göra justeringar, har vi exempelvis information om fördelningen av skadebeloppen kan detta inkluderas. Vi såg även i modelleringen att ett intercept kan inkluderas om det finns ett behov av detta, i vårt fall fick vi dock osignifikanta skattningar för samtliga α_j då vi testade detta och vi ser därför inget behov av att införa ett intercept i vår modell. Ett modellantagande i denna metod är dock att vi enbart

har en utbetalning per skada, vilket många gånger inte kommer att vara uppfyllt.

5.2 Resultat från modelleringen

När vi antog att våra första åtta skadeår var färdigutvecklade och beräknade den faktiska reserven och jämförde med våra skattade reserver gav de olika metoderna snarlika resultat, se Figur 9. För de första sex skadeåren är samtliga skattningar väldigt nära varandra, för det sjunde året överskattar de båda implementationerna av VNJ-metoden reserven samtidigt som chain ladder-metoden hamnar nära den faktiska reserven. För det sista skadeåret verkar motsatsen gälla. Även då det beräknas hur mycket de olika metodernas skattningar avviker från den faktiska reserven är de olika metoderna jämlika.

Då vi studerade stabilitet i reservskattningarna gav chain ladder-metoden ett märkbart sämre resultat för samtliga stickprovsstorlekar. Detta gäller både då vi tittade på de första åtta skadeåren och jämförde mot den faktiska reserven (Figur 12) samt då vi tittade på samtliga skadeår (Figur 11). Vi ser att för de mindre stickprovsstorlekarna producerar chain ladder-metoden långsvansade skattningar åt höger. Detta innebär att för vissa simuleringar kommer chain ladder-metoden ge en grov överskattning av reserven. Då vi tittade på stabiliteten i reservskattningarna för samtliga 15 skadeår gav båda implementationerna av VNJ-metoden liknande resultat även om vi för en större stickprovsstorlek ser tendens till ett mer stabilt resultat med VNJ-med approximation. Då vi istället tittade på de första åtta skadeåren och jämförde mot den faktiska reserven ser vi tydligt att VNJ-utan approximation ger ett mer stabilt resultat.

Vi testade att manipulera storleken på utbetalningarna i vårt data och gjorde en analys då vi antog att varje utbetalning för en skada var av samma storlek. I Figur 13 såg vi då att ingen av VNJ-metoderna fick en märkbar förändring av detta när det kommer till skattningen av den totala reserven för de första åtta skadeåren, vilket är i linje med resultatet vi ser i den faktiska reserven. Vi såg däremot att chain ladder-metoden får betydligt lägre reservskattningar för de två sista skadeåren, vilket visar att metoden är känslig mot sådana förändringar i data.

När vi analyserar data med enbart skador med fler utbetalningar såg vi i Figur 15 en större skillnad mellan de två implementationerna av VNJ-metoden. VNJ-utan approximation ger fortfarande mest stabila skattningar men tenderar att underskatta reserven. Detta är ett tecken på att när man har data med få observationer alternativ då en stor andel av skadorna har fler utbetalningar kan det vara ett problem att ignorera modellantagandet om enbart en utbetalning per skada.

6 Slutsats

Båda sätten att hantera fler utbetalningar per skada i VNJ-metoden ger liknande resultat i samtliga tester. När fler utbetalningar per skada tillåts fås dock stabilare skattningar och mindre avvikelse från den faktiska reserven då vi tittar på de första åtta skadeåren, enligt Figur 12. Bortsett från detta ses ingen märkbar skillnad mellan de olika sätten att hantera utbetalning. Det verkar alltså i praktiken fungera väl, för produkten vi undersöker, att ignorera det faktum att mer än en betalning kan ske per skada vid implementationen av VNJ-metoden. Det verkar även som att metoden utan approximation är att föredra i vårt fall då den producerar mer stabila skattningar.

Att ignorera modellantagandet om fler utbetalningar per skada verkar dock inte fungera för alla produkter. Då en större andel skador har fler utbetalningar än en eller då data består av färre observationer bör detta ses över innan modellantagandet ignoreras, se exempel i Figur 15.

Chain ladder-metoden ger något mindre medelkvadrat fel än de båda implementationerna av VNJ-metoden då reservskattningar för de första åtta åren jämförs mot den faktiska reserven. Vi såg dock i Figur 11 att vi får mycket större osäkerhet i reservskattningarna för samtliga stickprovsstorlekar med chain ladder-metoden. Metoden tenderar även att överskatta reserven vid vissa tillfällen.

Bortsett från att VNJ-metoden inte hanterar fler utbetalningar per skada har modellen visat sig att fungera bra. Reservskattningarna är i nivå med de vi får med chain ladder-metoden. Vi får stabila skattningar samtidigt som metoden är relativt lätt att implementera. En fördel är även att modellen producerar separata skattningar för kända och okända skador.

7 Vidare utveckling av arbetet

I modelleringen kan vi se att det finns osäkerhet i parameterskattningarna för VNJ-metoden. Det hade varit intressant att studera alternativa metoder för att skatta våra parametrar. Istället för att använda en generaliserad linjär modell för att skatta ψ skulle man kunna skatta p, exempelvis med en metod baserad på överlevnadsanalys och sedan från denna skattning få fram ψ. Eventuellt skulle detta kunna leda till mer stabila parameterskattningar. Det hade även varit intressant att se hur pass mycket en annan metod för att skatta parametrarna hade påverkat resultatet.

Då målet delvis varit att hitta en metod som går att applicera på flera olika produkter hade det varit intressant att reservsätta flera av Kammarkollegiets försäkringsprodukter med diskuterade metoder. En bättre utvärdering av hur anpassningsbar VNJ-metoden är hade då kunnat uppnås.

I (Miranda *et al.* , 2012) gör de en utvidgning av VNJ-metoden, de använder en så kallad "double chain ladder method" vilket innebär att de utför chain ladder algoritmen på två skadetrianglar istället för en. Denna metod har fördelen att även den kan producera separata skattningar för kända och okända skador. Det hade varit intressant att testa att även implementera denna metod på Kammarkollegiets försäkringsprodukt och jämföra mot de tidigare diskuterade metoderna.

För personförsäkringen som analyserats finns data på skadenivå, samt information om alla utbetalningar som gjorts. Med denna information har vi möjlighet att ställa upp betydligt mer avancerade metoder än de diskuterade i detta arbete. Då data finns på skadenivå kan istället exempelvis så kallade mikromodeller användas, se bland annat (Antonio & Plat, 2014). I dessa modeller låter man varje skada utgöra en egen process istället för att enbart titta på data på en aggregerad nivå. För varje skada simulerar man tiden från en händelse till en annan, exempelvis från att skadan inträffar tills den rapporterats och sedan till den första utbetalningen och så vidare. Med det data vi har att tillgå skulle en liknande analys som görs i (Sigmundsdóttir & Lindholm, 2017) kunna utföras för att se om detta kan ge en bättre prediktion än tidigare analyserade metoder. Fördelen med dessa metoder är att vi tar vara på mer av den informationen vi har vilket kan leda till bättre skattningar, dessa modeller tillåter även fler utbetalningar per skada till skillnad från VNJ-metoden. Nackdelen med modeller på denna form är att de lätt kan bli beräkningstunga samt att de kan komma att kräva stora justeringar då vi vill anpassa modellen på en annan försäkringsprodukt.

Referenser

Agresti, Alan. 2002. *Categorical Data Analysis*. 2 edn. John Wiley Sons.

Antonio, Katrien, & Plat, Richard. 2014. Micro-level stochastic loss reserving for general insurance. *Scandinavian Actuarial Journal*, **2014**(7), 649–669.

Eiopa. 2009. *Europaparlamentets och rådets direktiv 2009/138/EG av den 25 november 2009 om upptagande och utövande av försäkrings- och återförsäkringsverksamhet (Solvens II)*. 2017-01-31. Europeiska unionens officiella tidning. `http://eur-lex.europa.eu/legal-content/SV/TXT/PDF/?uri=CELEX:32009L0138`.

Kammarkollegiet. *Om Kammarkollegiet*. 2017-01-22. `http://www.kammarkollegiet.se/om-kammarkollegiet`.

Mack, Thomas. 1993. Distribution-Free Calculation Of The Standard Error Of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, **23**(2).

Mack, Thomas. 1994. Measuring the variability of chain ladder reserve estimates. *In: Casualty Actuarial Society Forum*, vol. 1.

McCullagh, P, & Nelder, J.A. 1989. *Generalized Linear Models*. 2 edn. New York: Chapman & Hall.

Miranda, María Dolores Martínez, Nielsen, Jens Perch, & Verrall, Richard. 2012. Double Chain Ladder. *ASTIN Bulletin*, **42**(1).

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Sigmundsdóttir, Jóhanna, & Lindholm, Mathias. 2017. One-Year Non-Life Solvency II Risk Calculations for Point Process Micro Models: Methods and Applications. *Mathematical Statistics, Stockholm University*.

Sjöberg, Lars. 2007. *Rätten till ersättning från det statliga försäkringssystemet*. Thomson Fakta AB.

Verrall, Richard, Nielsen, Jens, & Hedegaard Jessen, Anders. 2010. Prediction of RBNS and IBNR claims using claim amounts and claim counts. *ASTIN Bulletin*, **40**(11).

Wüthrich, Mario V. 2017. *Non-Life Insurance: Mathematics & Statistics*. 2017-12-21. `https://ssrn.com/abstract=2319328`.