Stockholms universitet

# Reservsättning inom sakförsäkring

## - Analys av en version av Bornhuetter-Ferguson metoden

Yuqi Ma

Masteruppsats i försäkringsmatematik
Master Thesis in Actuarial Mathematics

Masteruppsats 2019:5
Försäkringsmatematik
Juni 2019

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Matematisk statistik
Stockholms universitet
Masteruppsats **2019:5**
http://www.math.su.se/matstat

# Reservsättning inom sakförsäkring
## - Analys av en version av Bornhuetter-Ferguson metoden

Yuqi Ma*

Juni 2019

**Sammanfattning**

En skadeförsäkring ger ersättning vid fallet att försäkrad egendom tar skada. Tiden mellan det att skadan inträffar och att skadeersättningen är slutut- betald kan variera beroende på vilken typ av försäkring kunden valt att teckna. En av aktuariens uppgifter är att välja en lämplig metod för att skatta återstående utbetalningar från redan inträffade skador. Detta kallas reservsättning.

I detta masterarbete undersöks vilken metod som är bäst lämpad för en sjuk- och olycksfallsförsäkring från ett svensk försäkringsbolag. Samtidigt vill vi undersöka hur de använda metoderna passar för andra typer av försäkringsprodukter. De reservsättningsmetoder vi kommer att titta på är Chain Ladder och Bornhuetter-Ferguson. Analysen bygger påsimulerad data och visar att för en försäkringsprodukt med kort utbetalningsperiod ger Chain Ladder metoden bättre skattning och mindre osäkerhet av skadereserven medan Bornhuetter-Ferguson metoden är att föredra för produkten med lång utbetalningsperiod.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: angelama94@hotmail.com. Handledare: Mathias Lindholm.

# Förord

Denna uppsats utgör mitt masterarbete i försäkringsmatematik vid Stockholms Universitet och omfattar 30 hp.

Ett stort tack till min handledare Mathias Lindholm som har gett värdefull feedback och uppmuntran under arbetets gång.

# Innehåll

**1 Inledning** **5**
1.1 Syfte och metod . . . . . 5
1.2 Dataanalys . . . . . 6

**2 Teori** **6**
2.1 Utvecklingstriangel . . . . . 7
2.2 Chain Ladder . . . . . 8
2.2.1 Antaganden i Chain Ladder-metoden . . . . . 8
2.2.2 Parameterskattningar i Chain Ladder-metoden . . . . . 9
2.2.3 Medelkvadratfelet i Chain Ladder-metoden . . . . . 10
2.3 Bornhuetter-Ferguson . . . . . 11
2.3.1 Antanganden i Bornhuetter-Ferguson metoden . . . . . 12
2.3.2 Parameterskattningar i Bornhuetter-Ferguson metoden 13
2.3.3 Medelkvadratfelet i Bornhuetter-Ferguson metoden . . 14

**3 Modellanpassning på riktig data** **16**
3.1 Chain Ladder . . . . . 17
3.1.1 Kontroll av Chain Ladder antaganden . . . . . 17
3.1.2 Reservsättning med Chain Ladder-metoden . . . . . 19
3.2 Bornhuetter-Ferguson . . . . . 20
3.2.1 Kontroll av Bornhuetter-Ferguson antaganden . . . . . 20
3.2.2 Reservsättning med Bornhuetter-Ferguson metoden . . 22
3.3 Medelkvadratfel för Chain Ladder-metoden och Bornhuetter-Ferguson metoden . . . . . 25

**4 Modellanpassning på simulerad data** **27**
4.1 Datasimulering . . . . . 27
4.2 Dataanalys . . . . . 28
4.3 Kontroll av antaganden i Chain Ladder-metoden och Bornhuetter-Ferguson metoden . . . . . 29
4.4 Reservsättning med Chain Ladder-metoden och Bornhuetter-Ferguson metoden . . . . . 30
4.5 Medelkvadratfel för Chain Ladder och Bornhuetter-Ferguson . 32
4.6 Avvikelse från faktisk reserv för olika försäkringstyper . . . . . 33

**5 Slutsatser och diskussion** **39**

**6 Vidareutveckling av arbetet** **41**

**7 Referenser** **43**

**8 Appendix** **44**

# 1 Inledning

En skadeförsäkring ger försäkringstagare ersättningar om det inträffar en viss typ av skador. Försäkringsavtalet löper vanligtvis på ett år, men utbetalningsperioden ifall det inträffar skador kan variera beroende på vilken typ av försäkringsprodukt kunden valt att teckna. Om vi skulle dela upp försäkringsprodukter utifrån deras utbetalningsperiod, så finns det två typer: långsvansade och kortsvansade. För långsvansade försäkringsprodukter innebär det att tiden mellan skadan inträffar och att den är slutreglerad är lång. Exempel på sådana produkter kan vara sjuk-, olycksfall- och barns-försäkring. Kortsvansade försäkringsprodukter, såsom djurförsäkring, definieras av att utbetalningsperioden är kort.

För att kunna säkerställa att det finns tillräckligt mycket pengar att betala ut för inträffade skador behöver aktuarien göra en korrekt uppskattning av skadereserven, det vill säga en uppskattning av framtida utbetalningar från redan inträffade skador. Valen av reservsättningsmetod beror på typen av försäkringsprodukt.

## 1.1 Syfte och metod

Syftet med denna uppsats är att hitta en reservsättningsmetod som är lämplig för en produkt från ett försäkringsbolag samt undersöka hur dessa metoder fungerar för andra typer av data. Metoderna som används är Chain Ladder och Bornhuetter-Ferguson.

Chain Ladder är den vanligaste metoden inom reservsättning, den antar att framtida utbetalningar har samma mönster som de historiska. Modellen skattar utvecklingsfaktorer som beskriver förändringen i den ackumulerade skadekostnaden från ett år till ett annat år, sedan använder man dessa faktorer för att prediktera framtida utbetalningar. Denna metod passar bra för försäkringsprodukter som har relativt stabil utveckling och fungerar sämre om man inte har tillräckligt stort dataunderlaget.

Bornhuetter-Ferguson metoden antar istället att den totala skadekostnaden är känd vid beräkningstidpunkten. I praktiken brukar man använda Bornhuetter-Ferguson metoden för de försäkringsprodukter som inte har tillräckligt med historiska data eller produkter som har instabilt utbetalningsflöde och lång

utbetalningsperiod.

Upplägget för uppsatsen kommer att vara följande. Vi skattar först reserver med både Chain Ladder och Bornhuetter-Ferguson på riktig data från ett försäkringsbolag. Vid utvärderingen kommer vi studera medelkvadratfelet(MSE) för reserver. För att en metod ska vara bra, är det viktig att den är enkel att implementera samt passar för flera typer av data. I denna uppsats kommer vi undersöka detta genom att även tillämpa Chain Ladder och Bornhuetter-Ferguson på simulerad data, där vi själva kan bestämma hur datamaterialet ska se ut. Vi kommer att simulera två typer av data, en med kort utbetalningsperiod och den andra med lång utbetalningsperiod, båda simuleras 10 000 gånger. För att få en bättre bild av hur de två reservsättningsmetoderna fungerar för olika datatyper, kommer vi att välja ut ett datamaterial från varje datatyp och tillämpa båda metoderna på de två datamaterialen. Sist kommer vi även skatta reserverna med båda metoderna på alla 10 000 simulerade datamaterial och undersöka hur mycket de skattade reserverna avviker från de simulerade sanna framtida utbetalningarna för olika datatyper.

## 1.2 Dataanalys

Data som används i denna uppsats kommer från ett svensk försäkringsbolag, produkten som analyseras tillhör segmentet sjuk- och olycksfall. Data innehåller information om skadedatum, utbetalningsdatum, utbetalningsbelopp och premie för en produkt med 15 skadeår bakåt i tiden. Av sekretesskäl kommer data tranformeras, informationen om produkten kommer vi inte presentera närmare.

# 2 Teori

I detta avsnitt presenterar vi teorin som används i arbetet. Teorin om Chain Ladder metoden och Bornhuetter-Ferguson metoden är hämtad från (Mack, 1993) samt (Mack, 2006), (Mack, 2008).

## 2.1 Utvecklingstriangel

Innan vi går igenom teorin om reservsättningsmetoder så vill vi först presentera en viktig del i reservsättningen, nämligen utvecklingstriangeln. Utvecklingstriangeln är ett vanligt sätt att redovisa skador där raderna anger skadeperiod och kolumnerna utvecklingsperiod. Utvecklingstriangeln i sig kan användas för att uttrycka både skadehändelser och skadebelopp men i denna uppsats kommer vi att bygga vår triangel med skadeutbetalningar.

Låt $S_{ik}$ beteckna den inkrementella skadeutbetalningen under utvecklingsår $k$ för skador inträffade år $i$, där $1 \leq i \leq n$ och $1 \leq k \leq n$. Vi antar att alla utbetalningar är slutreglerade vid utvecklingsår $n$ och $S_{ik}$ för $i+k \leq n+1$ är känd. De inkrementella utbetalningarna för observerade skadeutbetalningar kan skrivas enligt Tabell 1.

| **Skadeår i/ Utvecklingsår k** | 1 | 2 | 3 | ... | $n-1$ | $n$ |
|---|---|---|---|---|---|---|
| 1 | $S_{11}$ | $S_{12}$ | $S_{13}$ | ... | $S_{1,n-1}$ | $S_{1,n}$ |
| 2 | $S_{21}$ | $S_{22}$ | $S_{23}$ | ... | $S_{2,n-1}$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $S_{i,n+1-i}$ | | |
| $n-1$ | $S_{n-1,1}$ | $S_{n-2,1}$ | | | | |
| $n$ | $S_{n,1}$ | | | | | |

Table 1: Inkrementell triangel över observerade skadeutbetalningar.

För den kumulativa triangeln skriver vi $C_{ik}$ istället för $S_{ik}$. Värdet $C_{ik}$ motsvarar ackumulerat utbetalt skadebelopp upp till och med utvecklingsår $k$ för skador inträffade år $i$. $C_{i,n+1-i}$ står för den nuvarande ackumulerade skadeutbetalningen och $C_{in}$ är den totala skadeutbetalningen för skadeår $i$. Tabell 2 visar den kumulativa triangeln över observerade skadeutbetalningar.

| **Skadeår i/ Utvecklingsår k** | 1 | 2 | 3 | ... | $n-1$ | $n$ |
|---|---|---|---|---|---|---|
| 1 | $C_{11}$ | $C_{12}$ | $C_{13}$ | ... | $C_{1,n-1}$ | $C_{1,n}$ |
| 2 | $C_{21}$ | $C_{22}$ | $C_{23}$ | ... | $C_{2,n-1}$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $C_{i,n+1-i}$ | | |
| $n-1$ | $C_{n-1,1}$ | $C_{n-2,1}$ | | | | |
| $n$ | $C_{n,1}$ | | | | | |

Table 2: Kumulativ triangel över observerade skadeutbetalningar.

Vi ser från Tabell 1 och Tabell 2 att den nedre triangeln är tom, eftersom den delen motsvarar tid som ännu ej har upplevts. Det vi vill göra är att prediktera de framtida utbetalningarna med hjälp av olika metoder baserat på information som vi har i den övre triangeln. Vi noterar att man kan bygga triangeln på olika tidsenheter, vi kommer i detta arbete att använda årsvisa tidssteg.

## 2.2 Chain Ladder

Den mest kända metoden inom reservsättning är Chain Ladder, metoden antar att framtida utbetalningar har samma mönster som de historiska. Man skattar först utvecklingsfaktorer som beskriver förändringen i den ackumulerade skadekostnaden från ett år till ett annat, dessa faktorer används sedan för att prediktera framtida utbetalningar.

### 2.2.1 Antaganden i Chain Ladder-metoden

För att kunna använda Chain Ladder metoden, måste några viktiga antaganden vara uppfyllda. Dessa antaganden är

**(CL1)** $(C_{i,1}, ..., C_{i,n}), (C_{j,1}, ..., C_{j,n}), i \neq j$ är oberoende,

där $n$ är året då alla skador är utbetalda.

**(CL2)** Det existerar utvecklingsfaktorer $f_1, ..., f_{n-1} > 0$ så att

$$\mathrm{E}[C_{i,k+1}|C_{i1}, ..., C_{ik}] = C_{ik} f_k.$$

Vi kan tolka ovan som att metoden använder samma utbetalningsmönster för alla skadeår, alltså oberoende av vilket år skadan har inträffat.

**(CL3)** Det existerar $\sigma_k$ så att

$$\mathrm{Var}(C_{i,k+1}|C_{i1}, ..., C_{ik}) = C_{ik}\sigma_k^2$$

där $k = 1, 2, \ldots, n-1$.

### 2.2.2 Parameterskattningar i Chain Ladder-metoden

Låt $f_k$ vara utvecklingsfaktorer som beskriver förändringen i den ackumulerade skadeutbetalningen mellan två på varandra följande år. Vi skattar utvecklingsfaktorerna $f_k$ enligt

$$\hat{f}_k = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{ik}} \tag{1}$$

där $k = 1, ..., n-1$, (Mack, 1993).

Den totala skadekostnaden för skadeår $i$ skattas enligt

$$\hat{C}_{in} = C_{i,n+1-i} \cdot \hat{f}_{n+1-i}, ..., \hat{f}_{n-1}. \tag{2}$$

Och reserven för skadeår $i$ kan uttryckas som

$$\hat{R}_i^{CL} = \hat{C}_{in} - C_{i,n+1-i}. \tag{3}$$

För att beräkna variansen av $\hat{C}_{i,k+1}$ behöver man även skatta $\sigma_k^2$

$$\hat{\sigma}_k^2 = \frac{1}{n-k-1} \sum_{i=1}^{n-k} C_{ik}(\frac{C_{i,k+1}}{C_{ik}} - \hat{f}_k)^2, \quad k = 1, 2, \ldots, n-2.$$

För skattning av $\sigma_{n-1}^2$ använder man sig av följande formel (Mack, 1993)

$$\hat{\sigma}_{n-1}^2 = \min(\hat{\sigma}_{n-2}^4 / \hat{\sigma}_{n-3}^2, \min(\hat{\sigma}_{n-3}^2, \hat{\sigma}_{n-2}^2)).$$

### 2.2.3 Medelkvadratfelet i Chain Ladder-metoden

Ett sätt att mäta kvaliteten på en skattning är medelkvadratfelet. Medelkvadratfelet för estimatorn $\hat{C}_{i,n}$ definieras enligt (Mack, 1993)

$$\begin{aligned} \text{mse}(\hat{C}_{i,n}) &= \text{mse}(\hat{R}_i^{CL}) \\ &= \text{E}[(\hat{C}_{i,n} - C_{i,n})^2 | D] \\ &= \text{Var}(C_{i,n}|D) + (\text{E}[C_{i,n}|D] - \hat{C}_{i,n})^2. \end{aligned} \tag{4}$$

Här betecknar $D = \{C_{i,k} : i + k \leq n + 1\}$ observerad data. Den första termen $\mathrm{Var}(C_{i,n}|D)$ anger processrisk och $(\mathrm{E}[C_{i,n}|D] - \hat{C}_{i,n})^2$ skattningsrisk.

Under antaganden (CL1), (CL2) och (CL3) kan medelkvadratfelet för $\hat{R}_i^{CL}$ estimeras enligt

$$\mathrm{mse}(\hat{R}_i^{CL}) = \hat{C}_{i,n}^2 \sum_{k=n+1-i}^{n-1} \frac{\hat{\sigma}_k^2}{\hat{f}_k^2} \left(\frac{1}{\hat{C}_{i,k}} + \frac{1}{\sum_{j=1}^{n-k} \hat{C}_{j,k}}\right). \tag{5}$$

## 2.3 Bornhuetter-Ferguson

Som nämns i (Mack, 2008) är skadereserven för Chain ladder metoden starkt beroende av nuvarande skadebelopp $C_{i,n+1-i}$. Det kan mycket väl hända att nuvarande belopp är 0. Då kommer reserven för det skadeåret skattas som 0, vilket kan leda till ett felaktigt resultat. Den bristen hos Chain Ladder undviks med Bornhuetter-Ferguson då reserven är oberoende av tidigare skadebelopp, man antar att den totala skadekostnaden är känd vid beräkningstidpunkten och kan skattas med hjälp av ett exponeringsmått. I praktiken brukar man använda premieinkomsten som exponeringsmått eftersom den kan spegla risken i affären bättre. En annan fördel med Bornhuetter-Ferguson är att den även tar hänsyn till svansutvecklingen medan Chain Ladder endast betraktar utvecklingen fram till ett givet utvecklingsår, metoden tillåter inte längre utvecklingsår än skadeår. Detta är mycket användbart för produkter som har lång utbetalningsperiod.

Låt $U_i$ beteckna den totala skadekostnaden för skadeår $i$, Bornhuetter-Ferguson reserven skattas enligt (Mack, 2008)

$$\hat{R}_i^{BF} = \hat{U}_i(1 - \hat{z}_{n+1-i}) \tag{6}$$

där $\hat{U}_i = v_i \hat{q}_i$ med $q_i = U_i/v_i$ av skadeår $i$ och $v_i$ är premie. $\hat{z}_k$ är den

skattade andelen av det totala skadebeloppet som är utbetalt efter utvecklingsår $k$. Man kan tolka det som att reserven för Bornhuetter-Ferguson räknas som den totala skadekostnaden multiplicerat med andelen som är ej utbetalda skadekostnader. Vi ser här att reserven i Bornhuetter-Ferguson är starkt beroende av valet av ultimo.

Det finns olika sätt att skatta $\hat{z}_k$, ett sätt är att skatta den med hjälp av utvecklingsfaktorer $\hat{f}_k$

$$\hat{z}_n = \hat{f}_\infty^{-1}, \hat{z}_{n-1} = (\hat{f}_n \hat{f}_\infty)^{-1}, ..., \hat{z}_1 = (\hat{f}_1 ... \hat{f}_n \hat{f}_\infty)^{-1} \tag{7}$$

där $\hat{f}_\infty$ är svansfaktorn, (Mack, 2008). Men detta val av $\hat{z}_k$ strider mot Bornhuetter-Ferguson antagande att reserven är oberoende av tidigare skadebelopp. Därför kommer vi inte att använda den metod för skattning av $z_k$, metoden som används i denna studie kommer att presenteras i avsnitt 2.3.2.

### 2.3.1 Antanganden i Bornhuetter-Ferguson metoden

Även för Bornhuetter-Ferguson så finns några antaganden som måste vara uppfyllda. För $1 \leq i \leq n$ och $1 \leq k \leq n+1$ har vi följande tre antaganden. (Mack, 2008)

**(BF1)** Alla inkrement $S_{i,k}$ är oberoende.

Alla inkrementella utbetalningarna $S_{i,k}$ av samma skadeår är oberoende, även själva skadeåren antas vara oberoende.

**(BF2)** Det existerar okända parametrar $x_i, y_k$ med $\mathrm{E}[S_{i,k}] = x_i y_k$ och $y_1 + y_2 + ... + y_{n+1} = 1$.

Observera att $x_i y_i = (x_i a)(y_k/a)$ för något $a > 0$, det vill säga $x_i$ och $y_i$ är bara unika upp till en konstant faktor. Därför kan man anta att $y_1 + y_2 + ... + y_{n+1} = 1$. Detta ger att $\mathrm{E}[U_i] = \mathrm{E}[S_{i,1} + ... + S_{i,n+1}] = \mathrm{E}[x_i y_i + ... + x_i y_{n+1}] = x_i$.

**(BF3)** Det existerar okända proportionalitetskonstanter $s_k^2$ sådana att $\mathrm{Var}(S_{i,k}) = x_i s_k^2$.

### 2.3.2 Parameterskattningar i Bornhuetter-Ferguson metoden

Låt $\hat{U}_i$ vara det skattade totala skadebeloppet för skadeår $i$ som kan fås fram genom att multiplicera premien med en skattad förlustkvot. Antag att vi har estimerat $\hat{U}_i$. Enligt formel (6) skattas Bornhuetter-Ferguson reserven för skadeår $i$ genom att multiplicera ultimo $\hat{U}_i$ med andelen ännu ej utbetalt $1 - \hat{z}_k$, där $\hat{z}_k = \hat{y}_1 + ... + \hat{y}_k$. Vi kallar $\hat{y}_k$ för den skattade skadekostnadsprocenten för utvecklingsår $k$ och beräknas enligt

$$\hat{y}_k = \frac{\sum_{i=1}^{n+1-k} S_{i,k}}{\sum_{i=1}^{n+1-k} \hat{U}_i}, \quad 1 \leq k \leq n \tag{8}$$

där $S_{i,k}$ är den inkrementella skadeutbetalningar under utvecklingsår $k$ för skador inträffade år $i$.

Som vi nämnt ovan, i Bornhuetter-Ferguson metoden inkluderar man hela svansen av skadeutvecklingen, det vill säga att vi även behöver skatta svansen vid skattningar av parametrarna. I (Mack, 2008) föreslås att man kan skatta $y_{n+1}$ genom att använda svanskvoten från liknande portföljer där skadeerfarenhet av senare utvecklingsår än år $n$ är tillgänglig. Om man inte har tillgång till sådan information så kan man estimera $y_{n+1}$ genom extrapolering från $\hat{y}_1, ..., \hat{y}_n$ med $\hat{y}_1 + ... + \hat{y}_n + \hat{y}_{n+1} = 1$.

Vi har nu allt vi behöver för att skatta reserven, den skattade reserven för Bornhuetter-Ferguson kan skrivas som

$$\begin{aligned} \hat{R}_i^{BF} &= \hat{U}_i(\hat{y}_{n+2-i} + ... + \hat{y}_{n+1}) \\ &= \hat{U}_i(1 - \hat{z}_{n+1-i}) \end{aligned} \tag{9}$$

med $\hat{z}_k = \hat{y}_1 + ... + \hat{y}_k$.

Enligt (Mack, 2008) skattas den okända proportionalitetskonstanten $s_k$ för utvecklingsår $k$ som

$$\hat{s}_k^2 = \frac{1}{n-k} \sum_{i=1}^{n+1-k} (S_{i,k} - \hat{U}_i \hat{y}_k)^2 / \hat{U}_i, \quad 1 \leq k \leq n-1. \tag{10}$$

Med ovanstående formel kan vi endast skatta $s_k$ fram till $k = n-1$. För skattningen av $s_n^2$ behöver vi använda extrapolering och $\hat{s}_{n+1}^2$ kan vi få fram genom att interpolera en regression av $\hat{s}_k^2$ mot $|\hat{y}_k|$ vid punkten $|\hat{y}_{n+1}|$ i enlighet med (Mack, 2008).

### 2.3.3 Medelkvadratfelet i Bornhuetter-Ferguson metoden

I (Mack, 2008) definieras medelkvadratfelet av $\hat{R}_i^{BF}$ för skadeår $i$ som

$$\text{mse}(\hat{R}_i^{BF}) = \text{E}[(\hat{R}_i^{BF} - R_i)^2 | S_{i,1}, ..., S_{i,n+1-i}]. \tag{11}$$

Enligt antagandet för Bornhuetter-Ferguson är $R_i = S_{i,n+2-i} + ... + S_{i,n+1}$ oberoende av $S_{i,1}, ..., S_{i,n+1-i}$ och även $\hat{R}_i^{BF}$ kan anses vara oberoende av $S_{i,1}, ..., S_{i,n+1-i}$.

Vi kan skriva om (11) som

$$\begin{aligned} \text{mse}(\hat{R}_i^{BF}) &= \text{E}[(\hat{R}_i^{BF} - R_i)^2] \\ &= \text{Var}(\hat{R}_i^{BF} - R_i) + (\text{E}[\hat{R}_i^{BF}] - \text{E}[R_i])^2 \\ &= \text{Var}(\hat{R}_i^{BF}) + \text{Var}(R_i). \end{aligned} \tag{12}$$

Vi ser från uttrycket ovan att medelkvadratfelet är summan av skattningsrisk $\text{Var}(\hat{R}_i^{BF})$ och processrisk $\text{Var}(R_i)$.

Vidare kan vi under antagandet för Bornhuetter-Ferguson metoden skatta medelkvadratfelet enligt

$$\text{mse}(\hat{R}_i^{BF}) = \hat{U}_i(\hat{s}_{n+2-i}^2 + ... + \hat{s}_{n+1}^2) +$$
$$(\hat{U}_i^2 + (\text{s. e.}(\hat{U}_i))^2)(\text{s. e.}(\hat{z}_{n+1-i}))^2 + (\text{s. e.}(\hat{U}_i))^2)(1 - \hat{z}_{n+1-i}^2) \tag{13}$$

där $\text{s. e.}(\hat{U}_i)$ beräknas med följande formel

$$\text{s. e.}(\hat{U}_i)^2 = \frac{v_i}{n-1} \sum_{j=1}^{n} v_j (\frac{\hat{U}_i}{v_j} - \hat{q})^2, \quad \text{med} \quad \hat{q} = \frac{\sum_{j=1}^{n} \hat{U}_j}{\sum_{j=1}^{n} v_j}. \tag{14}$$

Men den formeln kan vi endast använda om $\hat{U}_i$ antas vara okorrelerade enligt (Mack, 2008) sidan 96. För fallet då det existerar positiv korrelation mellan $\hat{U}_i$, bör termen $n-1$ enligt (Mack, 2008) ersätts med $n - \sqrt{n}$ för en konstant korrelationskoefficient $\hat{\rho}_{ij} = \frac{1}{\sqrt{n}}$, eller med $n - \sqrt{2n}$ för en minskande korrelationskoefficient $\hat{\rho}_{ij} = \frac{1}{1+|i-j|}$. Detta innebär att skattningar blir mer osäkra om det finns någon korrelation mellan $\hat{U}_i$. Mer exakt formel för fallet då det existerar korrelation mellan $\hat{U}_i$ är

$$n - \sum_{i,j} \rho_{ij}^U \sqrt{\frac{v_i}{v_+} \frac{v_j}{v_+}}$$

med $v_+ = \sum_{j=1}^{n} v_i$.

För att få ut medelkvadratsfelet för $\hat{R}_i^{BF}$, behöver vi även ha variansen för den skattade andelen utbetalda skadekostnader (Mack, 2008)

$$\text{Var}(\hat{z}_k) = \min(\text{Var}(\hat{y}_1) + ... + \text{Var}(\hat{y}_k), \quad \text{Var}(\hat{y}_{k+1}) + ... + \text{Var}(\hat{y}_{n+1})). \tag{15}$$

Eftersom $\hat{y}_k \approx \sum_{i=1}^{n+1-k} S_{i,k} / \sum_{i=1}^{n+1-k} x_i$ (Enligt antagandet är $E[U_i] = x_i$),

följer det att

$$\mathrm{Var}(\hat{y}_k) \approx \mathrm{Var}(\sum_{i=1}^{n+1-k} S_{i,k} / \sum_{i=1}^{n+1-k} x_i) = \frac{s_k^2}{\sum_{i=1}^{n+1-k} x_i} \quad , 1 \leq k \leq n. \tag{16}$$

Formel (16) kan vidare skrivas som

$$(\mathrm{s.\,e.}(\hat{y}_k))^2 = \frac{\hat{s}_k^2}{\sum_{i=1}^{n+1-k} \hat{U}_i} \quad , 1 \leq k \leq n. \tag{17}$$

Med hjälp av ovanstående formeln kan man endast räkna variansen för $\hat{y}_k$ där $k = 1, ..., n$. Variansen för $\hat{y}_{n+1}$ kan beräknas på olika sätt. I denna uppsats väljer vi att använda den metod som nämns i (Mack, 1993), det vill säga

$$(\mathrm{s.\,e.}(\hat{y}_{n+1}))^2 = \min(\frac{(\mathrm{s.\,e.}(\hat{y}_n))^2}{(\mathrm{s.\,e.}(\hat{y}_{n-1}))^2}, \quad \min((\mathrm{s.\,e.}(\hat{y}_{n-1}))^2, (\mathrm{s.\,e.}(\hat{y}_n))^2)). \tag{18}$$

Slutligen får vi

$$\begin{aligned}(\mathrm{s.\,e.}(\hat{z}_k))^2 = \min((&\mathrm{s.\,e.}(\hat{y}_1))^2 + ... + (\mathrm{s.\,e.}(\hat{y}_k))^2, \\ &, (\mathrm{s.\,e.}(\hat{y}_{k+1}))^2 + ... + (\mathrm{s.\,e.}(\hat{y}_{n+1}))^2).\end{aligned} \tag{19}$$

# 3 Modellanpassning på riktig data

I detta avsnitt gör vi reservskattningar med Chain Ladder och Bornhuetter-Ferguson på data från ett försäkringsbolag, resultaten kommer sedan jämföras

med varandra. Notera att datat som vi använder här är en långsvansad försäkringsprodukt, samtliga utbetalningar är inte helt slutreglerade efter 15 utvecklingsår, men eftersom Chain Ladder metoden inte tillåter fler utvecklingsår än skadeår, så väljer vi i detta arbete att använda data för 15 skadeår och 15 utvecklingsår vid båda skattningsmetoderna.

## 3.1 Chain Ladder

### 3.1.1 Kontroll av Chain Ladder antaganden

Som vi nämnt i Avsnitt 2.2.1, för att Chain Ladder metoden ska fungera bra, krävs att vissa antaganden måste vara uppfyllda.

Första antagandet är oberoende mellan skadeutbetalningarna under olika skadeår, detta är svårt att bevisa. I praktiken kan inflationseffekter eller liknande faktorer skapa beroende mellan betalningar under olika skadeår. I (Mack, 1993) föreslås att man kan undersöka detta modellantagande genom att göra ett test för skadeår influenser, mer om testet kan läsas i (Mack, 1993). I denna uppsats väljer vi att bortse ifrån sådana faktorer och antar att detta antagande är uppfyllt.

Det andra antagandet är att utbetalningsmönstret är detsamma för alla skadeår. Vi undersöker detta genom att plotta var och $C_{.,k}\hat{f}_k$ mot $C_{.,k+1}$. Antagandet anses vara uppfyllt om det visar en rät linje på samtliga plottar. I Figur 1 redovisas det ackumulerade utbetalningsbeloppet för utvecklingsår 1-15.

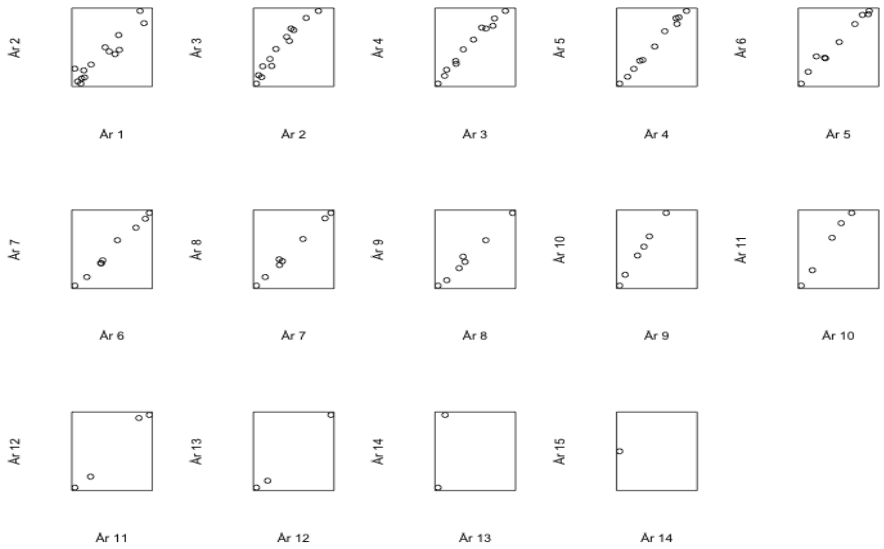



Figure 1: $C_{.,k}\hat{f}_k$ mot $C_{.,k+1}$ för utvecklingsår 1-15 baserad på riktig data.

Om vi bortser ifrån de senare skadeåren där vi har få observationer, kan vi ser att modellantagande ser bra ut för de andra åren och vi anser därmed att antagandet är uppfyllt.

Enligt Avsnitt 2.2.1 (CL3), ska det finnas $\sigma_k$ och den är oberoende av skadeår $i$, så att

$$\mathrm{Var}(C_{i,k+1}|C_{i1}, ..., C_{ik}) = C_{ik}\sigma_k^2$$

där $1 \leq i \leq 15$ och $1 \leq k \leq 14$. Vi undersöka detta genom att först räkna ut standardiserade residualer

$$\frac{C_{i,k+1} - \hat{f}_k C_{i,k}}{\hat{\sigma}_k \sqrt{C_{i,k}}}$$

där $i = 1, ..., 15 - k$. Sedan plottar vi de residualerna mot $i$. Antagandet är uppfyllt om residualerna inte visar några trender.

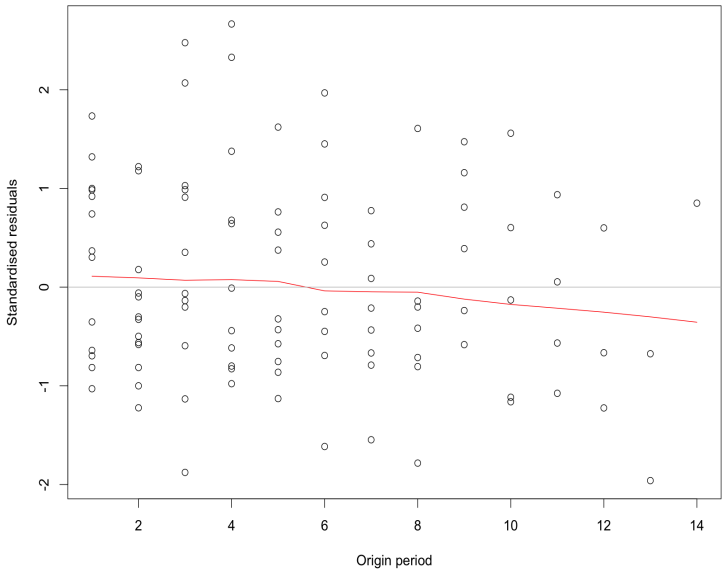



Figure 2: Standardiserad residual för Chain Ladder baserad på riktig data.

Från Figur 2 ser vi att de standardiserade residualerna är ganska jämnt spridda och vi anser att även detta antagande är approximativt uppfyllt.

### 3.1.2 Reservsättning med Chain Ladder-metoden

Vi går vidare med att skatta skadereserverna. Vi börjar med att räkna ut skattningar för utvecklingsfaktorer $\hat{f}_k$ med formel (1). Figuren nedan visar utvecklingsfaktorer $\hat{f}_k$ för $k = 1, ..., 14$ med Chain Ladder metoden.

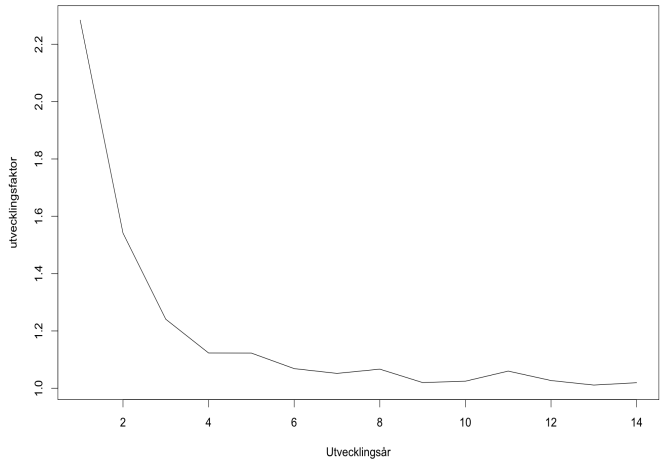



Figure 3: Skattade utvecklingsfaktorer $\hat{f}_k$ med Chain Ladder baserad på riktig data.

Vi ser från Figur 3 att de skattade utvecklingsfaktorerna har en avtagande trend, det innebär att större delen av utbetalningarna sker under de tidigaste utvecklingsåren för den försäkringsprodukten. Utvecklingsfaktorerna verkar fluktuera kring 1 från och med utvecklingsår 6.

Nästa steg är att skatta den totala skadekostnaden, detta gör vi genom att använda formel (2). Sist skattas reserven enligt formel (3). Figur 16 i appendix visar resultatet.

## 3.2 Bornhuetter-Ferguson

### 3.2.1 Kontroll av Bornhuetter-Ferguson antaganden

Även för Bornhuetter-Ferguson finns tre antaganden som måste vara uppfyllda. Det första antagande i Bornhuetter-Ferguson är att alla inkrementella utbetalningarna $S_{i,k}$ från samma skadeår är oberoende, även själva skadeåren antas vara oberoende. Detta antagande är väldigt svårt att bevisa matematiskt och i praktiken är det väldigt sällan så. Vi kommer att anta att alla

$S_{i,k}$ är oberoende och därmed även okorrelerade.

Det andra och tredje antagandet säger att det existerar okända parametrar $x_i$, $y_k$ samt $s_k^2$, där $E[S_{i,k}] = x_i y_k$ och $Var(S_{i,k}) = x_i s_k^2$ för $y_1 + ... + y_{n+1} = 1$. Dessa två antaganden undersöker vi genom att först räkna ut standardiserade residualer med följande formel

$$\frac{S_{i,k} - x_i \hat{y}_k}{\sqrt{x_i \hat{s}_k^2}}$$

där $i = 1, ..., 15$ och $k = 1, ..., 15$. Sedan plottar vi de standardiserade residualerna mot skadeår $i$. Antagandet är uppfyllt om residualerna ser slumpmässiga ut.

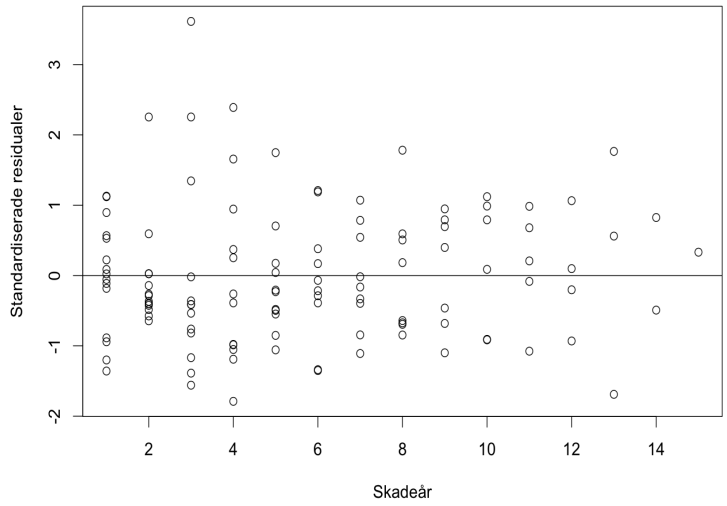



Figure 4: Standardiserade residualer för Bornhuetter-Ferguson baserad på riktig data.

Figur 4 redovisar de standardiserade residualerna för Bornhuetter-Ferguson metoden baserad på riktig data. Vi ser från figuren att residualerna inte

verkar ha några särskilda trender och ser slumpmässiga ut, därmed säger vi att det andra och tredje antagandet för Bornhuetter-Ferguson metoden är approximativt uppfyllda.

### 3.2.2 Reservsättning med Bornhuetter-Ferguson metoden

Vi börjar med att skatta den totala skadekostnaden $U_i$ för skadeår $i$. I detta arbete skattar vi den totala skadekostnaden genom att multiplicera premien med en skattad skadekvot, det vill säga $\hat{U}_i = v_i \cdot \hat{q}_i$. $\hat{q}_i$ är den skattade skadekvoten för skadeår $i$ som beräknas enligt

$$\hat{q}_i = \hat{r}_i \hat{m} \quad \text{med} \quad \hat{m} = \hat{m}_1 + ... + \hat{m}_{n+1} \tag{20}$$

där $\hat{r}_i$ är den skattade On-level premie faktorn för skadeår $i$, $\hat{m}_k$ är den genomsnittliga inkrementella förlustkvoten vid utvecklingsår $k$ och $\hat{m}_{n+1}$ är svanskvoten.

I (Mack, 2006) definieras den genomsnittliga inkrementella förlustkvoten för utvecklingsår $k$ enligt följande formel

$$\hat{m}_k = \frac{\sum_{i=1}^{n+1-k} S_{i,k}}{\sum_{i=1}^{n+1-k} v_i}. \tag{21}$$

I praktiken kan premienivåerna ändras över tid på grund av inflationen, för att få en bättre skattning av reserven behöver vi räkna bort inflationseffekten. Detta görs genom att multiplicera premien för varje skadeår med en On-level premie faktor, där den skattade On-level premie faktorn för skadeår $i$ definieras som

$$\hat{r}_i = \frac{C_{i,n+1-i}/v_i}{\sum_{k=1}^{n+1-k} \hat{m}_k}. \tag{22}$$

Vi använder sedan den On-level premie faktor i formel (21), då får vi den justerade genomsnittliga inkrementella förlustkvoten som visas i Figur 5 (blå linje). Vi ser att förlustkvoten har ganska starka svängningar efter utvecklingsår 5, detta kan bero på att vi har begränsat antal observationer för de sista utvecklingsåren. För att undvika oönskad volatilitet behöver den justerade genomsnittliga inkrementella förlustkvoten utjämnas och dessa värden bör avta mot 0 i enlighet med (Mack, 2006) Avsnitt 3. Vi väljer därför att utjämna värdena för utvecklingår 5 till 15 med ett linjärt polynom och sätter svanskvoten $\hat{m}_{n+1}$ till 0. Den utjämnade versionen av den justerade genomnisnittliga inkrementella förlustkvoten kallar vi för $\hat{m}_k^{adj*}$. Figur 5 visar den skattade inkrementella förlustkvoten utan utjämning samt med utjämning för utvecklingsår $k = 1, ..., 15$.

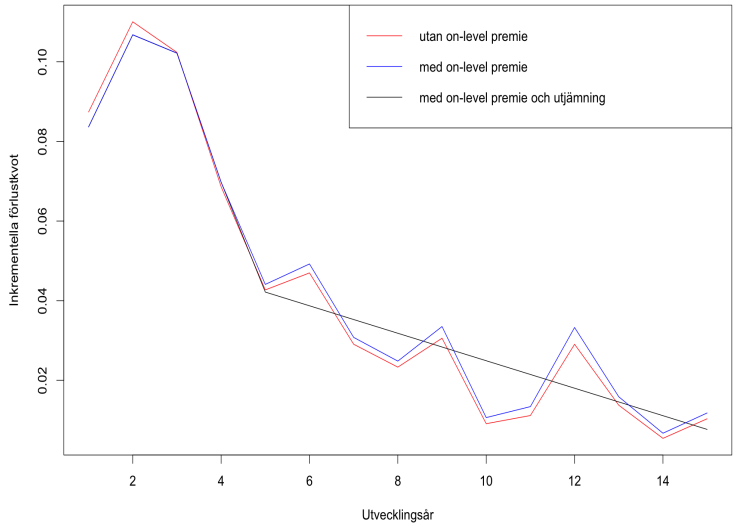



Figure 5: Inkrementella förlustkvoten för $k = 1, ..., 15$ baserad på riktig data.

Den skattade skadekvoten $\hat{q}_i$ beräknas nu med följande formel

$$\hat{q}_i = r_i \hat{m}^{adj*} \quad \text{med} \quad \hat{m}^{adj*} = \hat{m}_1^{adj*} + ... + \hat{m}_{n+1}^{adj*}. \tag{23}$$

Vidare estimerar vi den totala skadekostnaden $\hat{U}_i$ för Bornhuetter-Ferguson metoden enligt

$$\hat{U}_i = \hat{q}_i v_i. \tag{24}$$

Nu har vi estimerat den totala skadekostnaden, nästa steg är att skatta reserven med Bornhuetter-Ferguson. Vi börjar med att skatta skadekostnadsprocenten $y_k$ för $k = 1, ..., 15$ med formel (8). I detta arbete har vi använt ett datamaterial från en sjuk- och olycksfallsprodukt där samtliga skadeutbetalningar inte är helt slutreglerade efter 15 utvecklingsår, därför kan en svansskattning behövas.

Som vi nämnt i Avsnitt 2.3.2 så kan man skatta svanskvoten $y_{n+1}$ genom att få den från en liknande portfölj där skadeerfarenhet av senare utvecklingsår än år $n$ är tillgänglig. Men vi har tyvärr inte tillgång till sådan information, därför väljer vi istället att använda oss av metoden som står skriven i (Mack, 2008) som bygger på utjämning och extrapolation.

Figur 6 visar den skattade skadekostnadsprocenten utan utjämning (Röda linje). Vi ser att skattningarna är mer volatila i andra halvan av utvecklingsåren, därför väljer vi att göra en linjär utjämning för de utvecklingsåren och beteckna den utjämnade versionen av $\hat{y}_k$ med $\hat{y}_k^*$. Vidare genom extrapolation med villkoren $\hat{y}_1^* + ... + \hat{y}_{n+1}^* = 1$ kan vi komma fram till svanskvoten $\hat{y}_{n+1}^*$. Figur 6 visar $\hat{y}_k$ och $\hat{y}_k^*$.

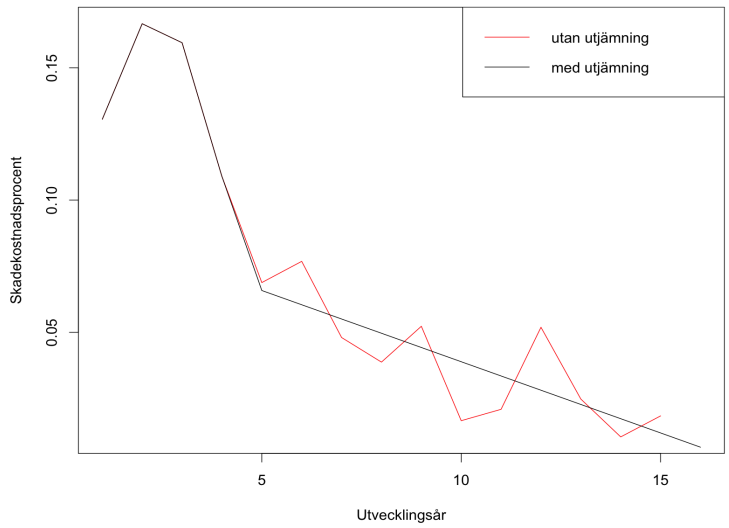



Figure 6: Skattade skadekostnadsprocent med Bornhuetter-Ferguson baserad på riktig data.

Slutligen skattas reserven med Bornhuetter-Ferguson enligt formel (9). Reservskattningar med Bornhuetter-Ferguson metoden jämfört med reservskattningar med Chain Ladder metoden baserad på riktig data visas i Figur 16 (Se appendix).

## 3.3 Medelkvadratfel för Chain Ladder-metoden och Bornhuetter-Ferguson metoden

För att undersöka osäkerheten i skattningar för båda metoderna vill vi studera medelkvadratfelet av reserverna. Felet beräknas enligt Avsnitt 2.2.3 och Avsnitt 2.3.3. Tar vi kvadratroten ur medelkvadratfelet så får vi standardfelet. Figuren nedan visar standardfelet för Chain Ladder och Bornhuetter-Ferguson.

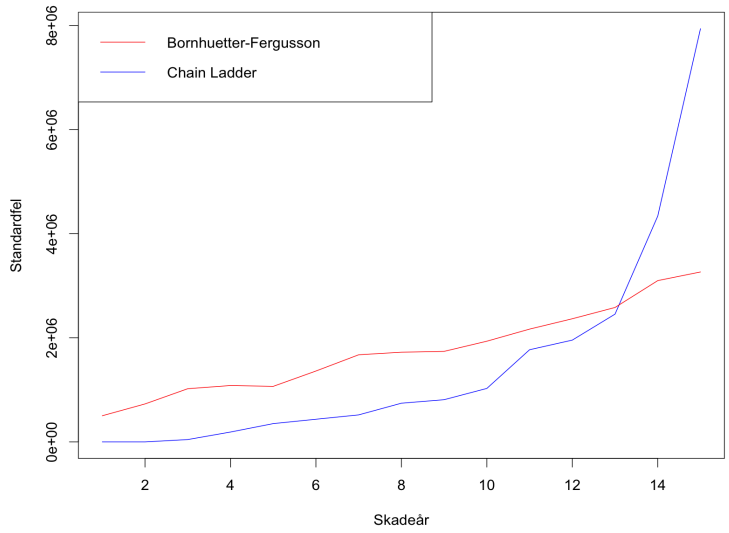



Figure 7: Standardfel för Chain Ladder och Bornhuetter-Ferguson baserad på riktig data.

Vi ser från Figur 7 att standardfelet ökar med skadeåren för båda metoderna. Detta innebär att osäkerheten i reservskattning ökar med ökande skadeår för båda metoderna, eftersom ju större skadeår desto mindre data har vi. Om vi jämför standardfelet för dessa två metoder, så har Bornhuetter-Ferguson större fel för skadeår 1-13, det vill säga osäkerheten i reservskattningar är större för den metoden under de åren. Däremot så har Bornhuetter-Ferguson metoden mycket lägre standardfel än Chain Ladder metoden för de senare skadeåren, vilket tyder på att Chain Ladder metoden ger sämre skattningar för de sista skadeåren, då dessa skattningar är väldigt osäkra. Utifrån ovanstående analys skulle det kunna antyda att Bornhuetter-Ferguson fungerar bättre än Chain Ladder, eftersom storleken på standardfel i generella termer är mer acceptabel, men vi kan inte dra någon direkt slutsats då vi inte har den faktiska reserven att jämföra med.

Vi kommer senare undersöka metoderna baserat på simulerat data som dels har kortare svans och dels har längre svans. Där kommer vi ha möjlighet att jämföra den skattade reserven med de faktiska simulerade framtida skadeut-

betalningarna.

# 4 Modellanpassning på simulerad data

## 4.1 Datasimulering

Vi undersökte hur Chain Ladder metoden och Bornhuetter-Ferguson metoden passar för en försäkringsprodukt inom sjuk- och olycksfall. Nu vill vi även se hur pass bra de två metoderna är om vi använder metoderna på andra typer av data. Vi väljer därför att simulera data för att sedan prediktera reserven med dessa två metoder. I detta arbete kommer vi att simulera två dataset som har olika längd på utbetalningsperioderna. För att sedan kunna undersöka variationen i reservskattningarna kommer varje dataset simuleras 10 000 gånger.

Vi simulerar data för skadeår $i$ med $i = 1, ..., 10$ enligt följande

1. Först simulerar vi antalet skador och beteckna den som $N_i$ för skadeår $i$. Vi antar att $N_i$ är Poissonfördelad med parametrar $w_i\lambda_i$, där $w_i$ motsvarar antalet aktiva kontrakt och $\lambda_i$ är skadefrekvens. I vårt fall väljer vi att ha (5503, 6060, 8780, 8597, 5276, 10918, 9600, 5576, 9647, 12116) aktiva kontrakt för respektive skadeår och skadefrekvensen är 0.2. Här har vi slumpat ut antalet aktiva kontrakt från en likformig fördelning över (5000, 12500).

2. Vidare låter vi $M$ beteckna antal utbetalningar för varje skada och antar att den är Poissionfördelad med väntevärde 1 och varians 1.

3. Givet antal utbetalningar väljer vi att låta $T$ som är tider mellan utbetalningar vara gammafördelade. För det första datasetet kommer vi att ha väntevärde 1 och varians 0.25, det andra datasetet har väntevärde 3 och varians 1.8.

4. Vi låter $Y$ beteckna betalningsstorlekar. Givet antal utbetalningar, antar att betalningsstorlekar är gammafördelade med väntevärde 1 och varians 0.5.

5. Låt $v_i$ beteckna premier för skadeår $i$. Premien för varje skadeår beräknas som den totala skadekostnaden multiplicerat med ett slumptal från en likformig fördelning mellan 0.9 och 1.3.

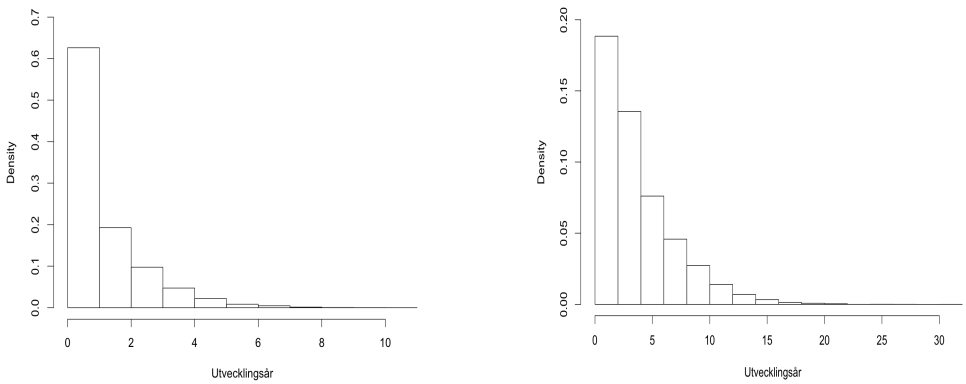
Här antar vi att alla $Y$ är oberoende och har samma fördelning, de är även oberoende av $N$.

Låt oss nu kalla det simulerade datat med förväntad fördröjningstid 1 som försäkringstyp 1 och datat med förväntad fördröjningstid 3 som försäkringstyp 2, båda försäkringstyperna simuleras 10 000 gånger. Notera att för de två datasetet har vi att allt är samma förutom $T$.

För att undersöka hur Chain Ladder metoden och Bornhuetter-Ferguson metoden passar för dessa två typer av data så väljer vi slumpmässigt ut ett datamaterial från varje försäkringstyp. Vi använder sedan båda metoderna på de utvalda datamaterialen och studerar även medelkvadratfelet för de olika metoderna. Sist kommer vi att skatta reserven med båda metoderna på alla 10 000 simulerade datamaterial för båda försäkringstyper. De skattade reserverna kommer sedan jämföras med de faktiska simulerade framtida utbetalningarna.

## 4.2 Dataanalys

Till att börja med vill vi undersöka hur betalningsfördröjningen ser ut för respektive försäkringstyp baserat på de slumpmässiga utvalda datamaterialen. Figur 8 visar tätheten för utbetalningstider för försäkringstyp 1 och försäkringstyp 2 baserat på de slumpmässiga utvalda datamaterialen.




Försäkringstyp 1 Försäkringstyp 2

Figure 8: Täthet för utbetalningstid baserad på simulerade data.

Från figuren ser vi att de flesta utbetalningar sker på de första utbetalningsåren för båda försäkringstyper. För försäkringstyp 1 är 99.9% utbetalningar slututbetalda inom 10 år, alla utbetalningar är slutbetalda vid utbetalningsår 11. För försäkringstyp 2 är 92.5% av utbetalningar slutreglerade inom 10 år och samtliga utbetalningar är slutreglerade vid utbetalningsår 31. Man kan säga att för försäkringstyp 2 är tiden mellan att skadorna rapporteras och att samtliga skadeutbetalningar är slutreglerade mycket längre än för försäkringstyp 1.

## 4.3 Kontroll av antaganden i Chain Ladder-metoden och Bornhuetter-Ferguson metoden

Som vi nämnt tidigare har Chain Ladder metoden tre antaganden. Antagande 1 är att varje skadeår är oberoende av varandra. Då vi har antagit att varje skadeutbetalningar är oberoende av antalet skador för alla skadeår, därför är detta antagande uppfyllt.

Figur 17 och 18 i appendix visar det ackumulerade utbetalningsbeloppet $C_{.,k}$ mot $C_{.,k+1}$ för respektive försäkringstyp. Vi bortser från de åren med få

observationer, för de övriga utvecklingsåren ser vi att för försäkringstyp 1 ser linjäriteten bra ut. När det gäller försäkringstyp 2 så verkar linjäriteten inte hålla för utvecklingsåren 1 mot år 2. För resterande åren håller linjäriteten väl. Eftersom det bara är för utvecklingsår 1 mot år 2 som visar sämre linjäritet så väljer vi att anse att antagande 2 ändå är approximativ uppfyllt för båda försäkingstyperna.

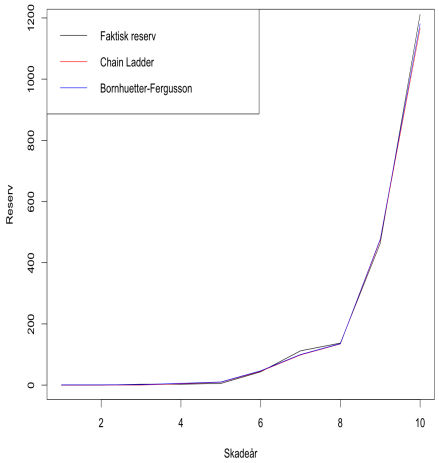
I Figur 19 (Se appendix) visas de standardiserade residualerna för Chain Ladder metoden för försäkringstyp 1 och försäkringstyp 2. Residualerna verkar inte visa några trender för båda försäkringstyper, därmed säger vi att även antagande 3 för Chain Ladder metoden är uppfyllt.

Vidare tittar vi på Bornhuetter-Ferguson antaganden. Då vi har antagit att alla skadeutbetalningar av samma skadeår är oberoende, därför är antagande 1 för Bornhuetter-Ferguson uppfyllt för båda försäkringstyperna.

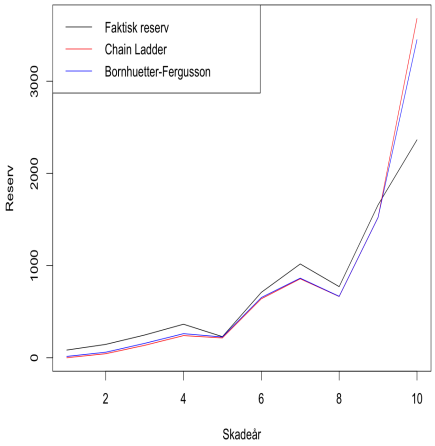
Figur 20 i appendix visar de standardiserade residualerna för Bornhuetter-Ferguson metoden för respektive försäkringstyp. Från figuren ser vi att residualerna är jämnt spridda, detta tyder på att antagande 2 och antagande 3 är uppfyllda för båda typerna.

## 4.4 Reservsättning med Chain Ladder-metoden och Bornhuetter-Ferguson metoden

I detta avsnitt skattar vi reserver med Chain Ladder och Bornhuetter-Ferguson för försäkringstyp 1 och försäkringstyp 2 baserat på de slumpmässiga utvalda datamaterialen. Resultatet ses i Figur 9.




Försäkringstyp 1




Försäkringstyp 2

Figure 9: Reservskattning med Chain Ladder och Bornhuetter-Ferguson jämfört med den faktiska reserven för försäkringstyp 1 och försäkringstyp 2 baserad på simulerad data.

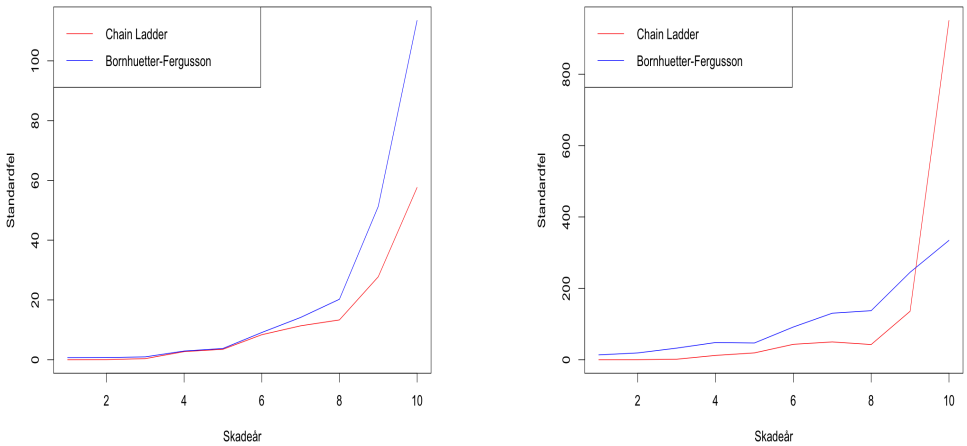
Figur 9 visar jämförelsen mellan den faktiska reserven och den skattade reserven för respektive försäkringstyp. Figuren till vänster visar reserven för försäkringstyp 1 med kort utbetalningsperiod, figuren till höger visar för försäkringstyp 2 där utbetalningsperioden är lång. Vi ser att båda metoderna ger bra reservskattningar för försäkringstyp 1, avvikelsen från den faktiska reserven är liten för båda metoderna. För försäkringstyp 2 verkar Chain Ladder och Bornhuetter-Ferguson ge något underskattade reserver för skadeår 1-8 men överskattar reserver för de sista två skadeåren, avvikelsen mellan reservskattningar med de två metoderna och den faktiska reserven är större för denna försäkringstyp. Notera att detta bara gäller för ett slumpmässigt utvalt datamaterial från varje försäkringstyp, för övriga datamaterialen kan resultatet variera. I Figur 21 (Se appendix) finns reservskattningar baserat på ett annat slumpmässigt utvald datamaterial. Vi ser att reservskattningar avviker mer från den faktiska skadeutbetalningar för de datamaterialen och även skillnaden mellan dessa två metoders avvikelse från den faktiska reserven är större.

Det är dock svårt att dra några slutsatser bara genom att jämföra skat-

tningar med faktiska reserver grafiskt, vi väljer därför gå vidare att titta på deras medelkvadratfel.

## 4.5 Medelkvadratfel för Chain Ladder och Bornhuetter-Ferguson

Medelkvadratfelet av reserven beräknas enligt avsnitt 2.2.3 och avsnitt 2.3.3. Standardfelet får vi genom att ta kvadratroten ur medelkvadratfelet. I Figur 10 redovisas standardfelet för Chain Ladder metoden och Bornhuetter-Ferguson metoden för försäkringstyp 1 samt försäkringstyp 2 baserat på de slumpmässiga utvalda datamaterialen.




Försäkringstyp 1 Försäkringstyp 2

Figure 10: Standardfel för försäkringstyp 1 och försäkringstyp 2 baserad på simulerade data.

Det visar sig att för de första skadeåren är standardfel för Bornhuetter-Ferguson reserven och Chain Ladder reserven väldigt nära varandra för båda försäkringstyper. För senare skadeår har Bornhuetter-Ferguson större osäkerhet i reservskattningar för försäkringstyp 1. Däremot så har den mindre standardfel jämfört med Chain Ladder metoden för det sista skadeåret för försäkringstyp

2. Om vi jämför standardfelet mellan dessa två försäkringstyper så ser vi att storleken på felet är större för försäkringstyp 2 än för försäkringstyp 1, detta tyder på att reservskattningar för data med längre utbetalningsperiod är mer osäkra än data med kort utbetalningsperiod, vilket känns rimligt då utvecklingen är svårare att skatta för långsvansade försäkringsprodukter än för kortsvansade produkter.

Ovanstående analys antyder att Chain Ladder metoden kan fungera bättre på data där utbetalningsperioden är kort. Dels för att den är enklare att implementera, dels för att storleken på standardfelet för Chain Ladder är mindre jämförd med Bornhuetter-Ferguson. För data med längre utbetalningsperiod verkar Bornhuetter-Ferguson passa bättre eftersom osäkerheten i reservskattningar är mindre för den metoden än för Chain Ladder, speciellt för de senare skadeåren.

## 4.6 Avvikelse från faktisk reserv för olika försäkringstyper

För att utvärdera vilken metod som passar bättre för simulerad data så vill vi nu jämföra reservskattningar för de 10 000 simulerade datamaterialen med den faktiska reserven. Det vill säga att vi skattar reserven 10 000 gånger för varje metod och försäkringstyp, sedan tittar vi på hur mycket våra reservskattningar avviker från den faktiska reserven för de olika försäkringstyperna.

För att kunna jämföra resultaten kommer vi att analysera

$$\frac{R_i - \hat{R}_i}{R_i} \quad i = 1, 2, ..., 10000 \tag{25}$$

där $R_i$ är den totala faktiska simulerade framtida utbetalningar för simulering $i$ och $\hat{R}_i$ är den skattade totala reserven.

Vi ser från ovanstående formeln att den kommer ge negativt resultat om $\hat{R}_i$ är större än $R_i$ och positivt resultat ifall $\hat{R}_i$ är mindre än $R_i$. Om $\hat{R}_i$ är lika med $R_i$, ger den 0 i resultat.

Vi börjar med att skatta reserven 10 000 gånger med Chain Ladder och

Bornhuetter-Ferguson för försäkringstyp 1 som har kort utbetalningsperiod, sedan använder vi formel (25). Resultatet presenteras i Figur 11 nedan. Den röda vertikala linjen i figuren motsvarar medelvärdet för 10 000 Chain Ladder reserver, blåa vertikala linjen motsvarar medelvärdet för 10 000 Bornhuetter-Ferguson reserver.

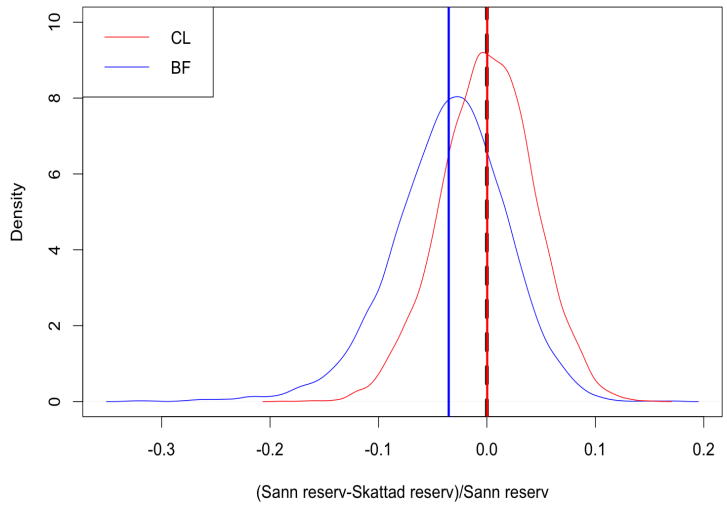



Figure 11: Avvikelse från faktisk reserv för försäkringstyp 1.

Vi ser att den skattade reserven med Chain Ladder metoden är mer centrerad kring 0 än Bornhuetter-Ferguson. Medelvärdet för Chain Ladder reserven är väldigt nära 0, vilket antyder att skattaren är approximativt väntevärdesriktig. Bornhuetter-Ferguson metoden ger större spridning än Chain Ladder och tenderar att avvika mer från den faktiska reserven. Vi ser även att fördelningen för Bornhuetter-Ferguson lutar sig mer åt vänster, medelvärdet av Bornhuetter-Ferguson reserver är -0.035, detta tyder på att en större andel av utfallen är negativa än positiva, alltså är reserven med Bornhuetter-Ferguson metoden något överskattad.

Vidare vill vi titta närmare på hur de skattade reserven avviker de faktiska utbetalningarna för varje skadeår. Vi använder följande formel

$$\frac{R_{i,k} - \hat{R}_{i,k}}{R_i} \quad i = 1, 2, ..., 10000 \quad \text{och} \quad k = 1, ..., 10 \tag{26}$$

där $R_{i,k}$ motsvarar den faktiska utbetalningen för skadeår $k$ och simulering $i$. Jämförelsen enligt (26) för försäkringstyp 1 ses i Figur 12.

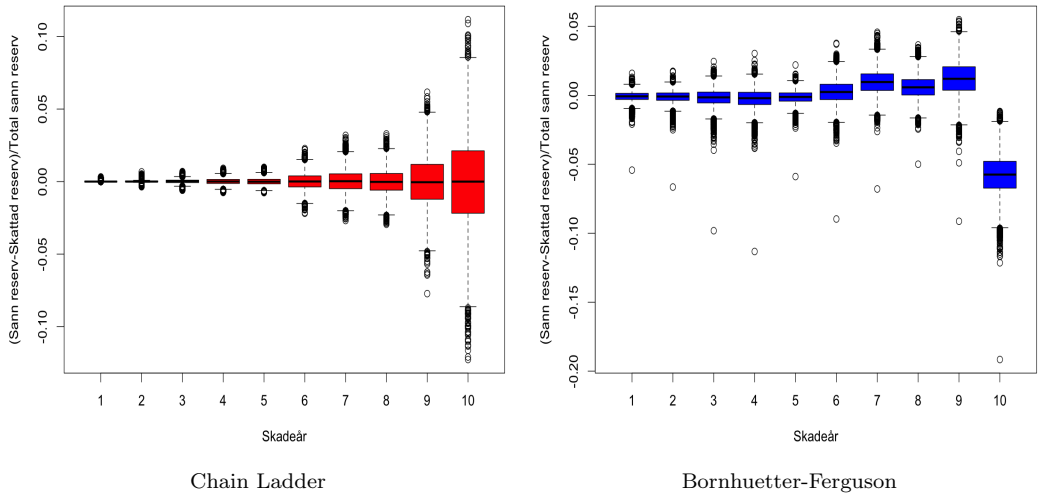



Figure 12: Reserven enligt (26) för försäkringstyp 1.

I Figur 12 ser vi att medianen för Chain Ladder reserven är väldigt nära 0 för samtliga skadeår. Vi ser tydligt att spridningen ökar med skadeår, den skattade reserven för skadeår 10 avviker mest från de faktiska utbetalningarna.

För Bornhuetter-Ferguson reserven ser vi att medianen ligger nära 0 för skadeår 1-9. För skadeår 10 är alla reservskattningarna överskattade. Detta antyder att det är reserven för skadeår 10 som orsakar att den totala Bornhuetter-Ferguson reserven blir överskattad.

Notera att i Figur 12 har vi olika skalor på y-axeln. För att kunna jämföra Chain Ladder reserven med Bornhuetter-Ferguson reserven för olika skadeår

gör vi även en plot där de har samma skala på y-axeln. Figur 22 i appendix visar resultatet för försäkringstyp 1.

Från Figur 22 ser vi att för skadeår 1-9 verkar resultatet inte skilja sig mycket mellan dessa två metoder. Den största skillnaden ligger i det sista skadeåret, där är variationen för Chain Ladder reserven mycket större än för Bornhuetter-Ferguson reserven. Detta tyder på att Chain Ladder reserven för skadeår 10 skiljer sig mer från den faktiska utbetalningen, men att denna osäkerhet är centrerad kring 0 till skillnad från Bornhuetter-Ferguson.

Nu estimerar vi reserven för försäkringstyp 2 där vi har data för 10 skadeår och utbetalningsperioden för denna försäkringstyp är lång, sedan använder vi formel (25) för att kunna jämföra med den faktiska reserven. Resultatet redovisas i Figur 13.

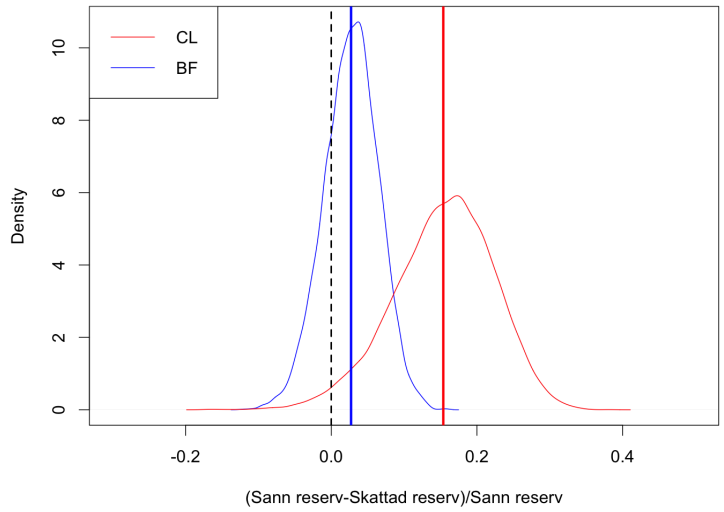
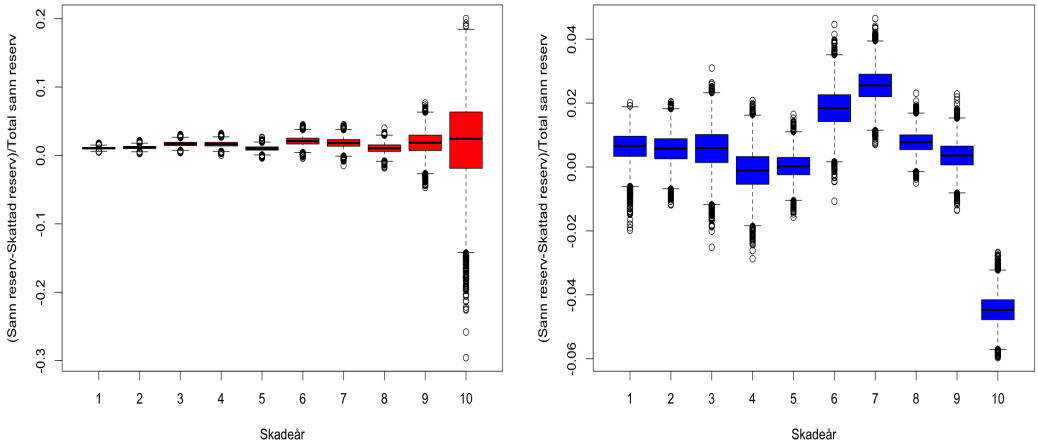



Figure 13: Avvikelse från faktisk reserve för försäkringstyp 2.

I Figur 13 ser vi att reserven för Chain Ladder metoden har större variation än för Bornhuetter-Ferguson. Fördelningen för Chain Ladder reserven lutar

sig mycket åt höger, nästan hela fördelningen hamnar på den högra sidan om 0 och medelvärdet för reserven är 0.1538284. Detta innebär att de flesta Chain Ladder reserverna är underskattade. För Bornhuetter-Ferguson verkar fördelningen luta sig också något åt höger men den ser mycket bättre ut än Chain Ladder. Medelvärdet för reserven är 0.0271976, detta skulle kunna antyda på att Bornhuetter-Ferguson reserven är mer likt den faktiska utbetalningen.

Med användning av formel (26), kan vi undersöka hur den skattade reserven avviker från den faktiska utbetalningen för varje skadeår. Figur 14 visar resultatet för försäkringstyp 2.




Chain Ladder

Bornhuetter-Ferguson

Figure 14: Reserven enligt (26) för försäkringstyp 2.

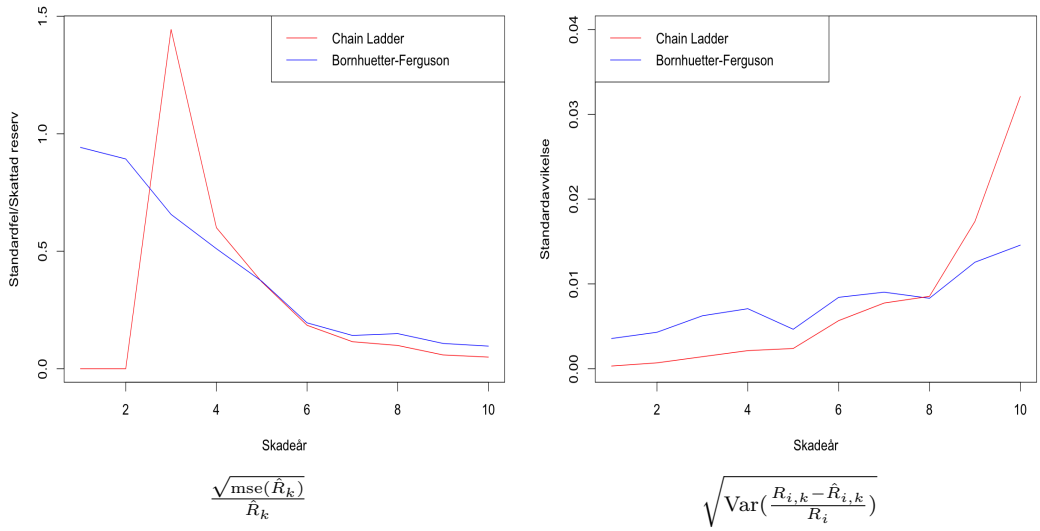
För försäkringstyp 2 visar Chain Ladder reserven liknande beteende som för försäkringstyp 1, spridningen blir större ju större $k$ blir. Skattningarna i Bornhuetter-Ferguson metoden är mer underskattade för skadeår 6 samt 7 och något överskattad för skadeår 10.

Vi tittar närmare då vi har samman skala på y-axel(Se Figur 23 i appendix).

Från Figur 23 ser vi att för skadeår 1-9 verkar skattningar inte skilja sig mycket mellan dessa två metoder. För skadeår 10 är variationen i skattningar större för Chain Ladder än för Bornhuetter-Ferguson även om Bornhuetter-Ferguson har fler systematiska avvikelser från 0. En del skattningar i Chain Ladder metoden är underskattade och en del är överskattade. Medan för Bornhuetter-Ferguson är alla skattningar överskattade för skadeår 10.

Om vi jämför Figur 11 med Figur 13 ser vi att Chain Ladder reserven ger sämre resultat på data med lång utbetalningsperiod, reserven är ganska mycket underskattad. För Bornhuetter-Ferguson metoden är spridningen större på data där utbetalningsperioden är kort och tenderar att överskatta reserven. Baserar på de 10 000 simulerade datamaterialen kan vi säga att Chain Ladder metoden är att föredra för försäkringstyp 1 medan för försäkringstyp 2 verkar Bornhuetter-Ferguson fungera bättre.

Sist vill vi även undersöka om medelkvadratfelet från Mack är ett bra mått på reservosäkerheten, detta görs genom att jämföra standardfelet av reserven med standardavvikelsen av (26) baserat på 10 000 simulerade datamaterial. På grund av tidbrist kommer vi i denna uppsats inte beräkna medelkvadratfelet för alla 10 000 simulerade data. Istället jämför vi standardfelet för ett datamaterial med standardavvikelsen för 10 000 simulerade datamaterial, det vill säga vi jämför hur nära $\frac{\sqrt{\text{mse}(\hat{R}_k)}}{\hat{R}_k}$ och $\sqrt{\text{Var}(\frac{R_{i,k}-\hat{R}_{i,k}}{R_i})}$ är, där $\hat{R}_k$ och $\text{mse}(\hat{R}_k)$ är den skattade reserven för skadeår $k$ samt medelkvadratfelet för den skattade reserven baserat på ett slumpmässigt utvalt datamaterial. $\sqrt{\text{Var}(\frac{R_{i,k}-\hat{R}_{i,k}}{R_i})}$ är standardavvikelsen av formel (26). Figur 15 och Figur 24 (Se appendix) visar jämförelsen.




$\frac{\sqrt{\text{mse}(\hat{R}_k)}}{\hat{R}_k}$ $\sqrt{\text{Var}(\frac{R_{i,k}-\hat{R}_{i,k}}{R_i})}$

Figure 15: Jämförelse mellan standardfelet av reserven och standardavvikelsen av (26) för försäkringstyp 1.

Vi ser från figurerna att standardfelet av reserven baserat på ett datamaterial är mycket större än standardavvikelsen av (26) baserat på 10 000 simulerade datamaterial för båda försäkringstyperna. Även om jämförelsen endast är baserat på ett utvalt datamaterial skulle detta kunna antyda att de teoretiska medelkvadratfelen är konservativa skattningar av reservosäkerhet för Chain Ladder och Bornhuetter-Ferguson reserver.

# 5 Slutsatser och diskussion

I detta arbete tittade vi på två reservsättningsmetoder: Chain Ladder och Bornhuetter-Ferguson. Först skattade vi reserver med hjälp av dessa två metoder på en sjuk- och olycksfallsprodukt. För att utvärdera metoden valde vi att titta på standardfel för skattningarna.

Vi såg att den skattade reserven med Chain Ladder och Bornhuetter-Ferguson är väldigt lika. När det gäller standardfelet, så visade sig att Bornhuetter-Ferguson reserven ger större osäkerhet på skattningen vid de tidiga skadeåren, men mycket mindre osäkerhet för de senare skadeåren. För Chain Ladder

reserven är osäkerheten väldigt stor för de senare skadeåren, exempelvis för skadeår 15 är Chain Ladder reserven runt 40 000 000 och standardfelet är runt 8 000 000 som är ungefär $\frac{1}{5}$ av reserven. Detta skulle kunna antyda att för denna försäkringsprodukt fungerar Bornhuetter-Ferguson bättre än Chain Ladder, eftersom storleken på standardfel för Bornhuetter-Ferguson är mer acceptabel än för Chain Ladder.

Observera att vi drog en liknande slutsats för försäkringstyp 2 där utbetalningsperioden är lång, detta bekräftar ytterligare att Bornhuetter-Ferguson är mer lämplig för försäkringsprodukten som har längre svans.

Då syftet även är att undersöka hur dessa två metoder passar för andra typer av data så valde vi att simulera egen data där vi själva kan bestämma hur data skulle se ut. Vi simulerade två dataset 10 000 gånger, den ena med kort utbetalningsperiod och den andra med lång utbetalningsperiod. För att undersöka närmare hur de olika metoderna fungerar på de simulerade datamaterialen, valde vi ut ett datamaterial från varje datatyp och tillämpade båda metoder på de två datamaterialen. När vi jämförde den skattade reserven med den faktiska simulerade framtida skadeutbetalningar såg vi att för försäkringstyp 1 är reserven för båda metoderna väldigt lika den faktiska. Båda Chain Ladder och Bornhuetter-Ferguson gav bra skattningar. För försäkringstyp 2 där utbetalningsperiod är betydlig längre än försäkringstyp 1 gav båda metoderna sämre skattning för de senare skadeåren. Chain Ladder reserven var mer överskattad än Bornhuetter-Ferguson reserven.

För att utvärdera vilken metod som gav mest säkra resultat tittade vi även på standardfelet av reserven för respektive försäkringstyp. Det visade sig att för försäkringstyp 1 är Bornhuetter-Ferguson reserven mer osäker än Chain Ladder reserven för de sista skadeåren. Däremot har Bornhuetter-Ferguson reservskattningar mindre osäkerhet för det sista skadeåret jämfört med Chain Ladder för försäkringstyp 2.

Sist men inte minst skattade vi reserven på de 10 000 simulerade data och jämförde de skattade reserverna med den faktiska simulerade framtida skadeutbetalningar för respektive försäkringstyp. För försäkringstyp 1 är Chain Ladder reserven väldigt lik den faktiska reserven. Medan för Bornhuetter-Ferguson metoden fick vi att den totala reserven är överskattad, det visade sig att det är reserven för skadeår 10 som i huvudsak orsakade detta.

För försäkringstyp 2 såg vi en helt annan bild. De flesta simuleringarna gav underskattade resultat när vi implementerade Chain Ladder metoden. Variationen är mycket större för Chain Ladder reserven än för Bornhuetter-Ferguson reserven för sista skadeåret, detta innebär att skattningar avviker mycket från den faktiska reserven för det året. Skattningar i Bornhuetter-Ferguson metoden är också något underskattade, men den såg mycket bättre ut än Chain Ladder, reserven med Bornhuetter-Ferguson metoden är mer lik den faktiska utbetalningen.

Till slut gjorde vi även en jämförelse mellan standardfelet av reserven baserat på ett slumpmässigt utvalt datamaterial och standardavvikelsen av formel (26) baserat på 10 000 simulerade datamaterial. Det visade sig att de teoretiska medelkvadratfelen är konservativa skattningar av reservosäkerheten för Chain Ladder och Bornhuetter-Ferguson reserver .

Av alla analyser vi gjorde kan vi dra slutsatsen att ju äldre skadeåren är, desto bättre resultat och mindre standardfel får vi med Chain Ladder metoden. Detta tyder på att Chain Ladder metoden fungerar bäst när vi har mer observerad data, vilket är väldigt rimligt då skattningar i Chain Ladder metoden är baserade på observerade ackumulerade utbetalningar. Chain Ladder metoden passar också bra för försäkringsprodukten med korta utbetalningsperiod, den ger bra skattning och mindre osäkerhet för sådana produkter. För långsvansade produkter verkar Bornhuetter-Ferguson metoden fungera bättre, detta kan bero på att vid reservskattningar tar den med hela svansen och metoden predikterar framtida utbetalningar med hjälp av premie som verkar ge oss bättre skattning än den observerade ackumulerade utbetalningar för produkten där utbetalningsperioden är lång.

# 6 Vidareutveckling av arbetet

I modelleringen såg vi att svansskattning spelade väldigt stort roll för Bornhuetter-Ferguson metoden. Metoden använder utjämning och extrapolering vid parameterskattningar för att skatta svansen. I detta fall hade det varit intressant att studera alternativa utjämningsmetoder såsom exponentiell utjämning och splineutjämning. Vid användning av olika utjämningsmetoder i param-

eterskattningen har man möjlighet att undersöka hur de olika metoderna påverkar resultatet.

På grund av brist på tid så beräknade vi bara medelkvadratfelet för riktig data och för de slumpmässiga utvalda datamaterialen. För alla 10 000 simulerade data skattade vi reserverna med olika metoder och jämförde resultatet med den faktiska utbetalningen. Det hade varit intressant att även undersöka medelkvadratfelet för alla 10 000 simulerade datamaterialen för att se osäkerheten i reservskattningar.

# 7 Referenser

Mack, T. (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. *Astin Bulletin 23(2): 213-225.*

Mack, T. (1993): Measuring the variability of chain ladder reserve estimates. *Casualty Actuarial Society Forum.*

Mack, T. (2006): Parameter Estimation of Bornhuetter/Ferguson. *Casualty Actuarial Society Forum.*

Mack, T. (2008): The Prediction Error of Bornhuetter-Ferguson. *Casualty Actuarial Society Forum.*

Wüthrich, M. V. & Merz, M. (2006): Stochastic Claims Reserving Methods in Non-Life Insurance. *Department of Mathematics, ETH Zürich & Faculty of Economics, University Tübingen.*

Wüthrich, M. V. (2015): Non-Life Insurance: Mathematics & Statistics. *Department of Mathematics, ETH Zürich.*

# 8 Appendix

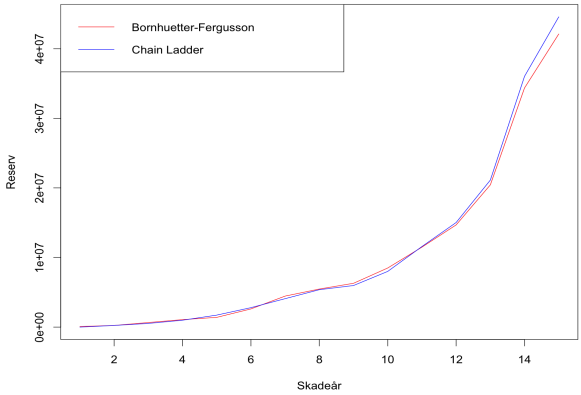



Figure 16: Reservskattning med Chain Ladder och Bornhuetter-Ferguson baserad på riktig data.

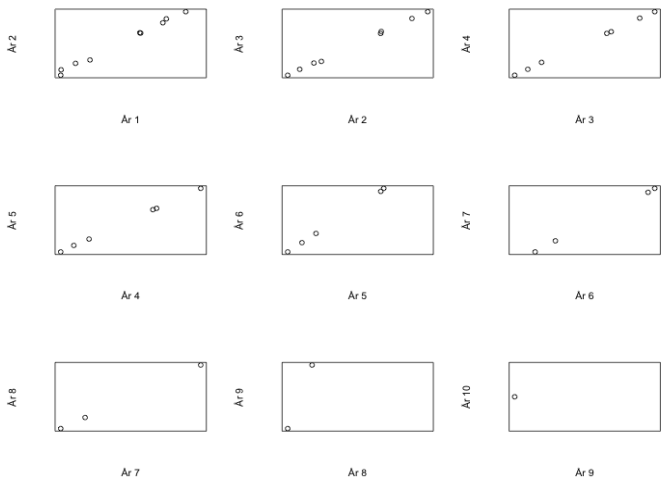



Figure 17: $C_{.,k}\hat{f}_k$ mot $C_{.,k+1}$ för utvecklingsår 1-10 för försäkringstyp 1.

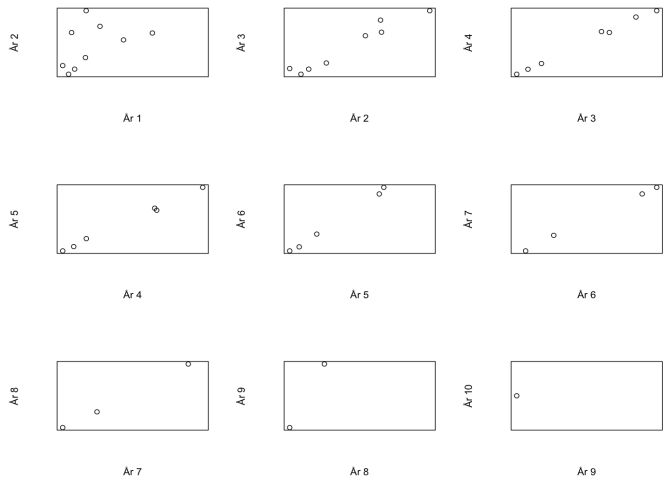
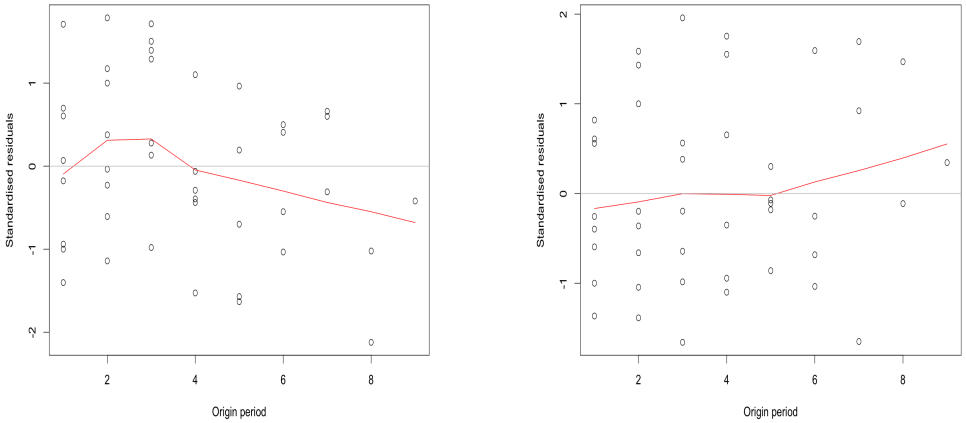
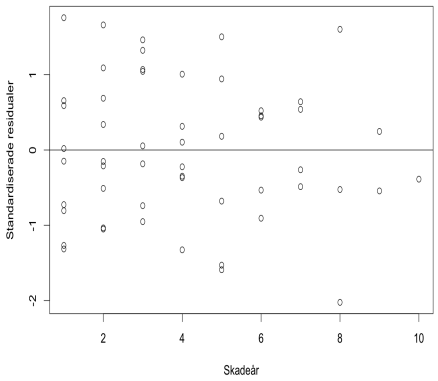



Figure 18: $C_{.,k}\hat{f}_k$ mot $C_{.,k+1}$ för utvecklingsår 1-10 för försäkringstyp 2.

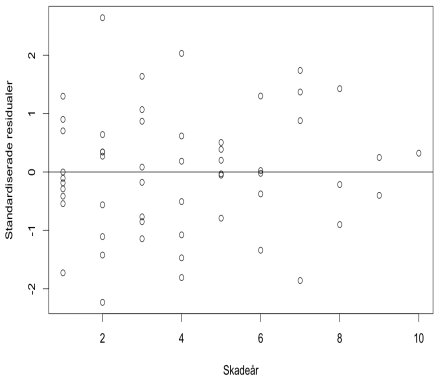



Försäkringstyp 1 Försäkringstyp 2

Figure 19: Standardiserade residualer för Chain Ladder för försäkringstyp 1 och försäkringstyp 2.

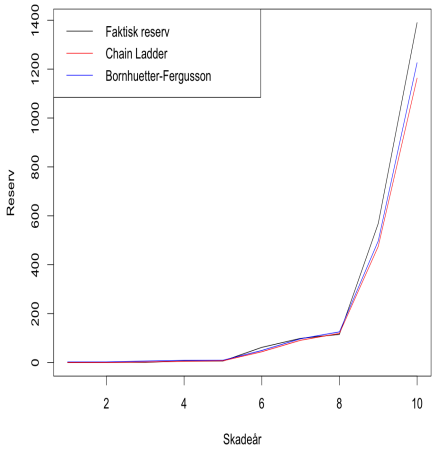



Försäkringstyp 1

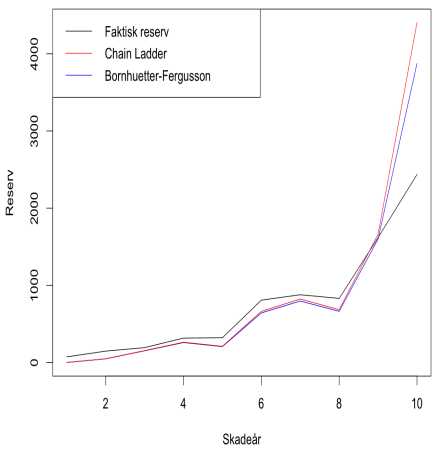



Försäkringstyp 2

Figure 20: Standardiserad residualer för Bornhuetter-Ferguson för försäkringstyp 1 och försäkringtyp 2.

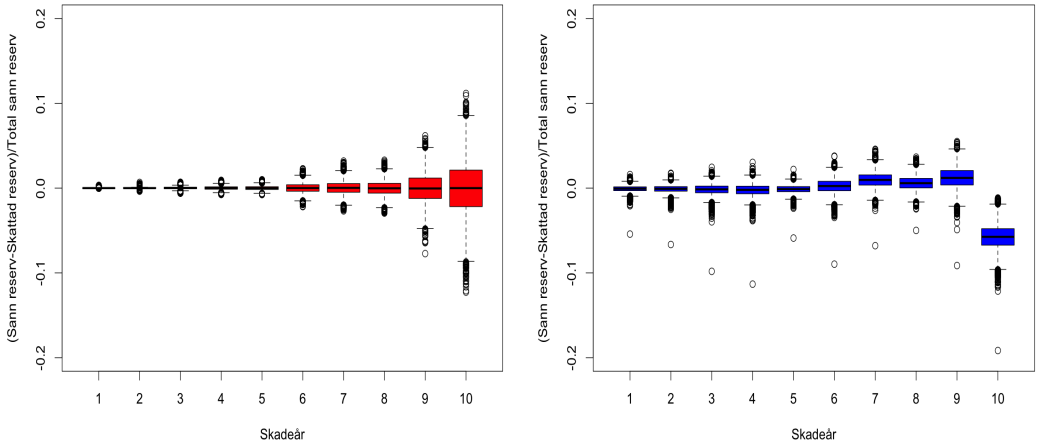



Försäkringstyp 1



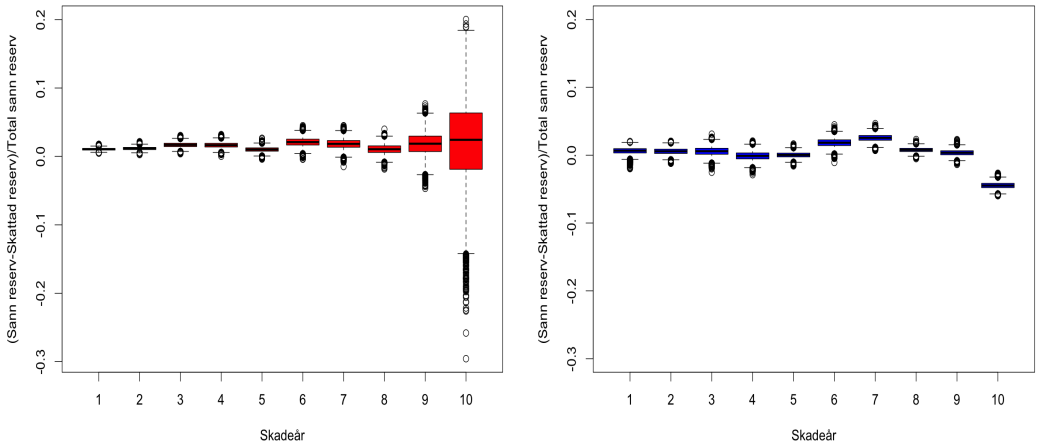

Försäkringstyp 2

Figure 21: Reservskattning med Chain Ladder och Bornhuetter-Ferguson jämfört med den faktiska reserven för försäkringstyp 1 och försäkringstyp 2 baserad på simulerad data.



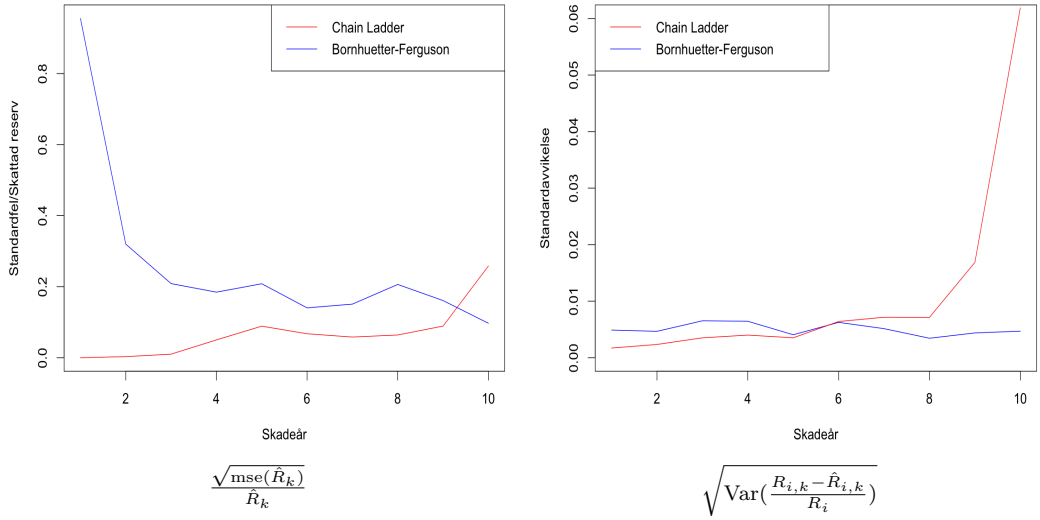

Chain Ladder Bornhuetter-Fergusson

Figure 22: Reserven enligt (26) för försäkringstyp 1 med samma skala på y-axel.




Chain Ladder Bornhuetter-Fergusson

Figure 23: Reserven enligt (26) för försäkringstyp 2 med samma skala på y-axel.




$\frac{\sqrt{\mathrm{mse}(\hat{R}_k)}}{\hat{R}_k}$ $\sqrt{\mathrm{Var}(\frac{R_{i,k}-\hat{R}_{i,k}}{R_i})}$

Figure 24: Jämförelse mellan standardfelet av reserven och standardavvikelsen av (26) för försäkringstyp 2.