

Matematisk statistik
Stockholms universitet

Inflationens inverkan på den parvisa korrelationen mellan skadetrianglar

Elin Roos

Examensarbete 2011:3

Postadress:

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm
Sverige

Internet:

<http://www.math.su.se/matstat>



Matematisk statistik
Stockholms universitet
Examensarbete 2011:3,
<http://www.math.su.se/matstat>

Inflationens inverkan på den parvisa korrelationen mellan skadetrianglar

Elin Roos*

September 2011

Sammanfattning

För att beräkna den sammanlagda osäkerheten av avsättningen för oreglerade skador för ett skadeförsäkringsbolag behöver vi veta hur de olika produkterna korrelerar med varandra. I detta arbete begränsas bolagets portfölj till att innehålla två produkter.

Vi kontrollerar att trianglarna uppfyller grundantagandena för Chain ladder, sedan görs en skattning av avsättningen för oreglerade skador och prediktionens skattade medelkvadratfel per produkt. Den totala prediktionens skattade medelkvadratfel för skadeavsättningen för bolagets portfölj beräknas med Brauns metod, genom att korrelationen mellan de båda produkterna skattas. Vi inflationsjusterar våra trianglar och skattar återigen korrelationen mellan produkterna. Vi kan därmed undersöka huruvida korrelationen mellan produkterna minskar med inflationsjusteringen.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige. E-post: elin.roos@folksam.se. Handledare: Ola Hössjer.

Abstract

To calculate the total prediction error of claims reserve for an nonlife insurance company we need to know how the different products correlate with each other. In this work the portfolio is limited to contain only two products.

We check that our triangles satisfy the basic assumptions of underlying the Chain Ladder method. Then we calculate an estimate of the prediction mean squared error and the claims reserve for each product. The mean squared error of the total prediction for the whole portfolio will be calculated using Braun's method, by estimating the correlation between the products. We adjust the triangles with the inflation and then re-estimate the correlation between the products. We can investigate whether the correlation between the two products decreases with the inflation adjustment.

Innehållsförteckning

1. Förord	4
2. Inledning	5
2.1 Problem	5
2.2 Allmänt om beräkning av skadeavsättning.....	5
3. Teori	8
3.1 Thomas Macks Chain Ladder.....	8
3.2 Kontroll av triangelarnas uppfyllnad av Macks grundantaganden.....	9
3.3 Osäkerheten i skattningen	10
3.4 Braun	13
3.4 Inflationsjustering.....	16
4. Genomförande och resultat	17
4.1 Data	17
4.2 Utfall av kontroll av triangelarna	17
4.2.1 Antagande 1	17
4.2.2 Antagande 2.....	19
4.2.3 Visuell kontroll av antagande 1 och 3.....	19
4.3 Avsättning och osäkerhet för en triangel.....	21
4.4 Avsättning och osäkerhet för två trianglar	23
4.5 Utfall efter inflationsjustering	26
5. Slutsatser	28
6. Referenslista	28
7. Appendix	29

1 Förord

Denna rapport utgör ett examensarbete om 30 högskolepoäng och leder till en magisterexamen i matematisk statistik vid Matematiska institutionen på Stockholms universitet.

Jag vill tacka min handledare professor Ola Hössjer för vägledning och goda råd under arbetets gång.

Jag vill även tacka min handledare Jesper Andersson på Folksam och mina övriga kollegor på aktuarieavdelningen.

2 Inledning

2.1 Problem

I ett sakförsäkringsbolag bedriver man oftast verksamhet inom olika försäkringsprodukter. För att beräkna försäkringsbolagets totala avsättning för oreglerade skador adderas de respektive produkternas avsättningar. Detta ger oss ett punkttestimat av dess väntevärde. Då detta punkttestimat sannolikt aldrig kommer att bli det sammanlagda utbetalda beloppet i framtiden ger ett prediktionsintervall för skattningen oss ytterligare information om vår förväntade framtida kostnad.

I detta arbete begränsas bolagets portfölj till att endast innehålla två olika produkter. Vi börjar med att undersöka huruvida Macks modell går att tillämpa på skadetrianglarna för respektive produkt. Avsättningen för oreglerade skador beräknas för båda produkterna och prediktions skattade medelkvadratfel. Först antar vi att de båda produkterna är oberoende och beräknar då den totala prediktions skattade medelkvadratfel. I nästa steg så använder vi oss av Brauns metod och skattar korrelationen mellan produkterna och beräknar det nya skattade medelkvadratfelet. Beroende på hur våra produkter korrelerar med varandra ökar eller minskar osäkerheten för vårt bolag.

Syftet med detta arbete är att kvantifiera korrelationen mellan ett antal utvalda produkter samt att undersöka om korrelationen kan förklaras av inflationen.

2.2 Allmänt om beräkning av skadeavsättning

Ett försäkringsavtal som ingås mellan en kund och ett försäkringsbolag ger upphov till ett betalflöde. För bolaget börjar betalflödet med inbetalningar av premie, vanligtvis sker detta med en engångsinbetalning vid avtalets början eller månadsvis via autogiro. I de avtalsperioder där kunden inte drabbas av en skada tar betalflödet slut med premieinbetalningen. I de fall där kunden drabbas av en skada anmäls denna till bolaget. Bolaget handlägger skadan, betalar ut ersättning samt avslutar skadan. Skadan kan även återöppnas då försäkringstagaren inkommer med nya krav på ersättning. Då vi här endast skattar risken i skadeavsättningen begränsar oss till att enbart använda det betalflöde som härrör från skadehändelserna.

Vid försäkringsperiodens slut som vanligtvis är ett år eller kvartal beräknar aktuarien en avsättning för de skador som har inträffat men inte betalats ut. Ett exempel på hur denna avsättning beräknas är följande:

Vi aggregerar data från skadesystemet till skadetrianglar innehållande utbetalt belopp. Raderna beskriver året då skadan inträffade och kolumnerna är utvecklingsåren, dvs det antal år som förflutit sedan skadan inträffade.

Låt $B_{i,k}$ vara för skadeår i , $1 \leq i \leq n$, det utbetalda beloppet under utvecklingsår k , $1 \leq k \leq n$. Triangeln nedan innehåller utbetalningarna för kända data, dvs då $i + k \leq n + 1$.

Skadeår i	Utvecklingsår k					
	1	2	3	..	n-1	n
1	B _{1,1}	B _{1,2}	B _{1,3}	..	B _{1,n-1}	B _{1,n}
2	B _{2,1}	B _{2,2}	B _{2,3}	..	B _{2,n-1}	
3	B _{3,1}	B _{3,2}	B _{3,3}	..		
	:	:	:			
n-1	B _{n-1,1}	B _{n-1,2}				
n	B _{n,1}					

$C_{i,k}$ är det ackumulerade utbetalda beloppet för skadeår i till och med utvecklingsår k, dvs.

$$C_{i,k} = \sum_{j=1}^k B_{i,j} .$$

Den icke skuggade triangeln nedan innehåller de ackumulerade utbetalningarna för kända data $i + k \leq n + 1$.

Efter vår aggregering av data vill vi skatta den slutgiltiga skadekostnaden för respektive skadeår. Vi gör ett förenklat antagande om att inga betalningar kommer att ske efter år n. Det innebär att vi inte behöver göra någon svansuppskattning. Därmed har vi redan den slutgiltiga

skadekostnaden för skadeår 1 och behöver endast beräkna $\hat{C}_{i,n}$, för $2 \leq i \leq n$. En vanlig modell för denna skattning är Macks modell, även känd som Chain-ladder. Med Macks modell skattar vi den okända delen av triangeln då $i + k > n + 1$ (som nedan är skuggad) och därmed även den slutgiltiga skadekostnaden.

Skadeår i	Utvecklingsår k					
	1	2	3	..	n-1	n
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$..	$C_{1,n-1}$	$C_{1,n}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$..	$C_{2,n-1}$	$\hat{C}_{2,n}$
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$..	$\hat{C}_{3,n-1}$	$\hat{C}_{3,n}$
	:	:	:	..	:	:
n-1	$C_{n-1,1}$	$C_{n-1,2}$	$\hat{C}_{n-1,3}$..	$\hat{C}_{n-1,n-1}$	$\hat{C}_{n-1,n}$
n	$C_{n,1}$	$\hat{C}_{n,2}$	$\hat{C}_{n,3}$..	$\hat{C}_{n,n-1}$	$\hat{C}_{n,n}$

Då blir vår skattning av skadeavsättningen för år i:

$$\hat{R}_i = \hat{C}_{i,n} - C_{i,n+1-i}$$

och den sammanlagda skadeavsättningen för samtliga år skattas till:

$$\hat{R} = \sum_{i=2}^n \hat{R}_i = \sum_{i=2}^n \hat{C}_{i,n} - \sum_{i=2}^n C_{i,n+1-i} .$$

Detta är en väldigt förenklad metod, i praktiken är det betydligt mer komplicerat. Aktuarien kan använda sig av många verktyg dels genom att fylla skadetrianglarna med andra uppgifter än utbetalt belopp såsom antal skador, skadereglerarnas uppskattade kostnad för både frekvensskador och storskador och göra en skattning av ultimokostnaden respektive ultimo antalet skador. Vi kan även använda oss av medelskada och skadefrekvens som stöd i beräkning av avsättningen.

Beräkning av avsättning sker som oftast nedbrutet på skademoment och adderas sedan ihop till en sammanlagd avsättning för produkten. Detta beror på att olika skademoment har olika betalmönster, dels kan produkten både innehålla långsvansade och kortsvansade skademoment men även säsongsbetingade skademoment. De olika skademomentens totala kostnad kan utvecklas olika med tiden och det kan innebära att betalflödet för den aggregerade triangeln förändras även om de enskilda momentens mönster är intakta.

Skadebetalningarna är i regel positiva betalningar till kund eller leverantör. Det kan dock förekomma negativa betalningar i form av regresser. Vanligt förekommande regresser är trafikegendom, som man kan välja att följa upp separat. Det finns andra mer slumpvis återkommande skadehändelser som kan föranleda regresser såsom kostnader för översvämningsskador, där försäkringsbolaget kan kräva ersättning av kommuner om dessa inte underhållit vattenavrinningen. Då kraftiga översvämningssomrar är slumpvis återkommande är dessa inbetalningar sällan något som ska appliceras på de efterföljande skadeåren utan man bör rensa för detta i modellen.

Det finns även stora enskilda skadehändelser som kan påverka utbetalningsflödet kraftigt. Ett exempel är stormen Gudrun som inträffade 8-10 januari 2005. 2007 så uppskattades den totala kostnaden för de fyra stora sakförsäkringsbolagen till 3965 mkr¹ vilket svarar för drygt 20 %² av den totala uppskattade skadekostnaden för hem och villa, företag och fastighet samt motorfordon. Då en sådan stor händelse inträffar tidigt på året får vi ett annat utbetalningsflöde för det aktuella skadeåret.

Det finns även andra förändringar av skadekostnader som vi behöver ta i beaktande. Dessa kan komma genom egna villkorsförändringar men även samhällsförändringar såsom trafikförsäkringsreformen och ändrade sjukskrivningsregler. Dessa förändringar innebär att vi får nya kostnader eller att vissa kostnader försvinner. Det sannolika är att de nya kostnaderna inte följer samma betalmönster som de tidigare, därmed kan vi inte applicera det historiska betalmönstret som vår modell genererar. Vi behöver därför göra en egen skattning över de kommande diagonalerna eller för skadeåret efter införandet.

Vi gör även ett förenklande antagande att alla betalningar sker inom n år i detta arbete. För kortsvansade moment sker utbetalningarna inom 5-10 år. För de långsvansade momenten förekommer det betalningar 20 år efter skadans inträffande. Väljer vi ett lågt n kommer vi att bortse från en mängd utbetalningar.

Med ovanstående kommentarer kan man undra vad nyttan av den modell som arbetet beskriver är om vi ändå inte tillämpar den i verkligheten. En statisk modell är bra att använda sig av dels som kontrollverktyg för våra egna manuellt korrigerade modeller men även i kommunikationen med revisioner. När det gäller den kunskapen vi får av korrelationen ger det oss en bättre förståelse över den risk försäkringsbolaget är utsatt för.

¹ Statens offentliga utredningar (2007): Sverige inför klimatförändringar – hot och möjligheter

² Finansinspektionens årsblankett blad H per 2006-12-31

3 Teori

3.1 Thomas Macks Chain Ladder

Chain ladder är den populäraste reservsättningsmetoden, detta på grund av att den är enkel, fördelningsfri och väldigt intuitiv. Den ger en känsla av att metoden inte bygger på några som helst antaganden men det stämmer inte, eftersom den grundar sig på följande tre grundantaganden:

Antagande 1:

$$E[C_{i,k} | C_{i,1}, C_{i,2}, \dots, C_{i,k-1}] = C_{i,k-1} * f_k \quad (3.1.1)$$

Det förväntade ackumulerade utbetalda beloppet för nästkommande år kan skattas genom att multiplicera utbetalt belopp hittills med en skattning av utvecklingsfaktorn f_k . Det okända värdet $C_{i,k}$ har en slumpmässig fördelning med väntevärde $C_{i,k-1} * f_k$, där f_k är skattad av all tillgänglig data genom

$$\hat{f}_k = \frac{\sum_{i=1}^{n+1-k} C_{i,k}}{\sum_{i=1}^{n+1-k} C_{i,k-1}} = \sum_{i=1}^{n+1-k} \frac{C_{i,k-1}}{\sum_{i=1}^{n+1-k} C_{i,k-1}} * \frac{C_{i,k}}{C_{i,k-1}} \text{ för } 2 \leq k \leq n \quad (3.1.2)$$

Vår skattning är därmed ett viktat medelvärde av de skattade utvecklingsfaktorerna för varje skadeår och inte den mest intuitiva ansatsen att ta ett oviktat medelvärde. Det innebär att skattningen tar hänsyn till volymen för respektive skadeår.

Med en rekursiv metod kan vi härleda väntevärdet för den slutgiltiga skadekostnaden för respektive skadeår med:

$$E[C_{i,n} | C_{i,1}, \dots, C_{i,k-1}] = C_{i,n+1-i} * f_{n+2-i} * \dots * f_n.$$

Då antagande 1 är betingat på $C_{i,k-1}$ är denna en konstant, vilket leder till att vi enkelt kan skriva om ekvation 2.1.1 till följande:

$$E\left[\frac{C_{i,k}}{C_{i,k-1}} | C_{i,1}, C_{i,2}, \dots, C_{i,k-1}\right] = f_k \quad (3.1.3)$$

Detta innebär att de skattade utvecklingsfaktorerna kan antas vara väntevärdesriktiga.

Antagande 2

Vektorerna $\{C_{i,1}, C_{i,2}, \dots, C_{i,n}\}$ och $\{C_{j,1}, C_{j,2}, \dots, C_{j,n}\}$ är oberoende om $i \neq j$.

Antagande 3

Den framtida variansen erhålls genom:

$$\text{Var}[C_{i,k} | C_{i,1}, C_{i,2}, \dots, C_{i,k-1}] = C_{i,k-1} * \sigma_k^2 \quad (3.1.4)$$

där vår skattning av σ_k^2 blir följande givet att ovanstående tre antagande är uppfyllda:

$$\hat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=2}^{n+1-k} C_{i,k-1} \left(\frac{C_{i,k}}{C_{i,k-1}} - \hat{f}_k \right)^2 \quad \text{för } 2 \leq k \leq n-1. \quad (3.1.5)$$

3.2 Kontroll av trianglarnas uppfyllnad av Macks grundantagande

Det finns flera olika sätt att kontrollera om våra trianglar uppfyller Macks grundantagande. Några exempel på detta är:

Kontroll av antagande 1

I Mack (1994) appendix G beskrivs ett förfarande där det testas om skattningarna

$$F_{i,k} = \frac{C_{i,k}}{C_{i,k-1}}$$

av utvecklingsfaktorerna är okorrelerade för olika utvecklingsår k. Detta är en

naturlig hypotes givet ekvation (3.1.3). Testet går till så att alla faktorer $F_{i,k}$ per utvecklingsår rankas i storleksordning. Sedan jämförs rankningen för $F_{i,k}$ med faktor $F_{i,k-1}$ för

$1 \leq i \leq n+1-k$. Är skillnaden i ranking för stor innebär detta att en hög faktor följs av en låg eller vice versa. För testet används Spearmans rank korrelationstest då vi har samma väntevärde men olika varians för våra faktorer. Mack antar att Spearmans teststatistika T för hela triangeln är normalfördelad och väljer att kontrollera trianglarna på konfidensgraden 50 %. Han motiverar det hela med att det dels är en approximation och att han vill kunna upptäcka korrelationer på delar av triangeln.

Konfidensintervallet är väldigt snävt och 50 % av fallen kommer att förkastas felaktigt. Vi kommer därför att använda oss av en högre konfidensgrad. Vi har tillgång till åtta skadetrianglar där vi kan tänka oss att en av trianglarna förkastas i genomsnitt.

Låt N vara antalet trianglar som förkastas, och

$$N_i = \begin{cases} 1 \\ 0 \end{cases} \text{ beroende på om testet för triangel i förkastas (1) eller ej (0).}$$

Då fås

$$E(N) = \sum_{i=1}^8 E(N_i) = 8 * \alpha$$

Där α väljs så att $E(N)=1$, dvs

$$\alpha = \frac{E(N)}{8} = \frac{1}{8}$$

Då får vi $1 - \frac{\alpha}{2} = 1 - 0,0625 = 0,9375$ vilket svarar mot kvantilen $z \approx 1,53$ för standardnormalfördelningen och en signifikansnivå på 12,5 % per triangel.

Kontroll av antagande 2

I Mack (1994) appendix H finns ett test som kontrollerar om skadeåren är oberoende. Den främsta orsaken till att trianglar inte uppfyller detta antagande är när det sker en kalenderårsvis förändring. Några exempel på detta är stora förändringar av skaderegleringen eller yttre faktorer som inflation, domstolsbeslut och förändringar från försäkringskassan.

En kalenderårsvis förändring påverkar både den aktuella diagonalen men även faktorerna för den aktuella diagonalen och den efterföljande. Testet går till så att alla faktorer per utvecklingsår delas upp i små respektive stora faktorer, medianen tas bort. Antalet små faktorer på diagonal $A_j = \{F_{j,2}, F_{j-1,3}, \dots, F_{1,j+1}\}$ är L_j som antas vara binomialfördelat enligt $\text{Bin}(r, 1/2)$, där r är antalet klassade värden på diagonalen. Samma antagande gäller för antalet stora faktorer S_j . För att mäta eventuell avvikelse skapas en ny variabel $Z_j = \min(L_j, S_j)$.

Om Z_j är signifikant mindre än $\frac{L_j + S_j}{2}$ så är triangeln påverkad av en kalenderårsvis trend.

Testet utförs på hela triangeln där teststatistikan $Z = \sum_{j=2}^{n-1} Z_j$ approximeras med normalfördelningen och kontrolleras med 5 % signifikans.

Visuell kontroll av antagande 1 och 3

Det är även möjligt att göra en visuell kontroll genom att plotta $C_{i,k+1}$ mot $C_{i,k}$ samt residualerna per respektive utvecklingsår. I plotten över $C_{i,k+1}$ mot $C_{i,k}$ läggs även den räta linjen $y = \hat{f}_k * x + 0$ in för att visuellt avgöra om antagande 1 är rimligt, dvs om man kan se att våra mätvärden är slumpvist fördelade längs den räta linjen. För residualerna $(C_{i,k+1} - C_{i,k}) / \sqrt{C_{i,k}}$ sker plottning mot $C_{i,k}$ och här söker vi en slumpvis fördelning av mätvärdena kring noll som inte beror på värdet av $C_{i,k}$ för att kontrollera antagande 3.

3.3 Osäkerheten i skattningen

Vår skattning av skadekostnaden för skadeår i är $\hat{C}_{i,n} = C_{i,n+1-i} * \hat{f}_{n+2-i} * \dots * \hat{f}_n$ och då blir skadeavsättningen $\hat{R}_i = \hat{C}_{i,n} - C_{i,n+1-i}$. Den totala skadekostnaden för hela triangeln blir då $\sum_{i=1}^n \hat{C}_{i,n}$ och den totala skadeavsättningen $\hat{R} = \sum_{i=2}^n \hat{R}_i = \sum_{i=2}^n (C_{i,n} - C_{i,n+1-i})$. \hat{R} kan ses som en prediktion av skadeavsättningen R som erhålls genom att i definitionen av \hat{R} ersätta alla $\hat{C}_{i,n}$ med de sanna men okända värdena $C_{i,n}$.

Vi övergår nu till beräkning av osäkerheten och börjar med att beräkna den för ett skadeår:

Ett skadeår

Osäkerheten beräknas genom medelkvadratfelet:

$$mse(\hat{C}_{i,n}) = E(C_{i,n} - \hat{C}_{i,n})^2 | T_n$$

Genom förhållandet mellan avsättningen och vår slutgiltiga kostnad som är $\hat{R}_i = \hat{C}_{i,n} - C_{i,n+1-i}$ får vi att:

$$mse(\hat{R}_i) = E(R_i - \hat{R}_i)^2 | T_n = E(C_{i,n} - C_{i,n+1-i} - \hat{C}_{i,n} + C_{i,n+1-i})^2 | T_n = E(C_{i,n} - \hat{C}_{i,n})^2 | T_n = mse(\hat{C}_{i,n})$$

och motsvarande skattning av medelkvadratfelet är:

$$s.e.(\hat{R}_i) = \sqrt{mse(\hat{R}_i)} = \sqrt{mse(\hat{C}_{i,n})} = s.e.(\hat{C}_{i,n})$$

där T_k är mängden $\{C_{i,j} | 1 \leq i \leq n, 1 \leq j \leq k, i + j \leq n + 1\}$ för $1 \leq k \leq n$ av alla kända värden i vår triangel upp till utvecklingsår k . För att förenkla notationen framöver så kommer betingningen att vara underförstådd. Det innebär att alla uttryck såsom $E(C_{i,k} | T_{n+1-i})$ respektive $Var(C_{i,k} | T_{n+1-i})$ betecknas med $E(C_{i,k})$ och $Var(C_{i,k})$. Detsamma gäller för våra skattningar.

I Mack (1994) presenteras en formel för skattningsfelet av den slutgiltiga skadekostnaden per respektive skadeår. I Mack (1999) visas att man kan skriva om det till en mer lättprogrammerad rekursiv formeln. Den senare varianten är den som Braun (2004) har använt sig av och det kommer även vi att göra. I Braun (2004) kan man tydligt följa hur medelkvadratfelet approximeras med summan av ett stokastiskt fel och ett skattningsfel.

$$mse(\hat{R}_i) = mse(\hat{C}_{i,n}) \approx Var(C_{i,n} | T_{n+1-i}) + Var(\hat{C}_{i,n} | T_{n+1-i}) \quad (3.3.1)$$

Braun härleder en rekursiv formel för respektive term i högerledet och när han lägger ihop dem får han följande skattning:

$$mse(\hat{C}_{i,k}) = f_k^2 mse(\hat{C}_{i,k-1}) + C_{i,k-1}^2 \left(\frac{\sigma_k^2}{\sum_{j=1}^{n+1-k} C_{j,k-1}} + \frac{\sigma_k^2}{C_{i,k-1}} \right) \quad (3.3.2)$$

där startvärdet i vår rekursiva skattning är:

$$mse(\hat{C}_{i,n+1-i}) = 0$$

Indexet $k = n + 1 - i \Leftrightarrow k + i = n + 1$ betecknar den sista diagonalen med kända värden i vår triangel. Diagonalen och de tidigare diagonalerna är konstanter och har därmed medelkvadratfelet = 0.

Vi skattar σ_k genom ekvation (3.1.5) för $2 \leq k \leq n - 1$. För $k=n$ får vi hitta en annan skattning då detta utvecklingsår endast innehåller en observation. Om $f_n = 1$ och vi inte tror att det kommer några fler betalningar så kan vi sätta $\sigma_n = 1$. I annat fall interpolerar vi vilket leder till att vår skattning blir:

$$\hat{\sigma}_n = \min\left(\frac{\hat{\sigma}_{n-1}^4}{\hat{\sigma}_{n-2}^2}, \min(\hat{\sigma}_{n-2}^2, \hat{\sigma}_{n-1}^2)\right)$$

Hela triangeln

Då vi är intresserade av avsättningen för hela triangeln behöver vi även dess osäkerhet. Vi har i inledningen begränsat oss till att alla betalningar sker till och med år n . Skadekostnaden för skadeår 1 är därmed redan känd och den totala avsättningen beräknas genom $R = R_2 + \dots + R_n$.

Skattningen av den slutgiltiga skadekostnaden för respektive år, $C_{i,n}$ för $2 \leq i \leq n$ är ej oberoende eftersom att de alla beror på skattningar av samma faktorer f_k , $2 \leq k \leq n$. Därför kan vi inte addera varianserna som vi gör med avsättningen. Vi vill beräkna en skattning av vårt totala prediktionsfel, $mse(\sum_{i=2}^n \hat{C}_{i,n})$.

På liknande sätt som för ett skadeår fås skattningen genom följande rekursiva formel:

$$mse\left(\sum_{i=n+2-k}^n \hat{C}_{i,k}\right) = mse\left(\sum_{i=n+3-k}^n \hat{C}_{i,k-1}\right) * f_k^2 + (\hat{C}_{\geq,k-1})^2 * \left(\frac{\hat{\sigma}_k^2}{\hat{C}_{\geq,k-1}} + \frac{\hat{\sigma}_k^2}{\hat{C}_{<,k-1}}\right) \quad (3.3.3)$$

där:

$$\hat{C}_{<,k-1} = \sum_{i=1}^{n+1-k} \hat{C}_{i,k-1}$$

$$\hat{C}_{\geq,k-1} = \sum_{i=n+2-k}^n \hat{C}_{i,k-1}$$

startvärdet är:

$$mse(\hat{C}_{n+1-k,k}) = 0.$$

Summationen av medelkvadratfelet börjar på $i = n + 2 - k$. Detta beror på att vi vill börja summera över de första okända värdena. De kända värdena i triangeln begränsas av

$i + k \leq n + 1 \Leftrightarrow i \leq n + 1 - k$ vilket ger det första okända värdet för skadeåret

$i = n + 2 - k$ givet respektive utvecklingsår. Rekursionen ger oss det totala medelkvadratfelet för hela triangeln.

Prediktionsintervall

För att beräkna ett prediktionsintervall för vår skattning av skadeavsättningen för skadeår i så måste vi börja med att anta en fördelningsfunktion för R_i . En möjlighet är att anta en normalfördelning. Dock är fördelningen vanligtvis skevare än den symmetriska normalfördelningen, detta speciellt om $\sqrt{Var(R_i)} > 50\% * R_i$. I dessa fall passar lognormalfördelning med parametrarna μ_i och σ_i bättre, dvs

$$\exp(\mu_i + \sigma_i^2 / 2) = E(R_i)$$

$$\exp(2\mu_i + \sigma_i^2)(\exp(\sigma_i^2) - 1) = \text{Var}(R_i)$$

Detta leder till skattningarna:

$$\hat{\sigma}_i^2 = \ln(1 + (s.e(\hat{R}_i))^2 / \hat{R}_i)$$

$$\hat{\mu}_i = \ln(\hat{R}_i) - \hat{\sigma}_i^2 / 2$$

Prediktionsintervallet för vår skadeavsättning för år i blir då:

$$\hat{R}_i \cdot \exp(\pm \lambda * \hat{\sigma}_i - \hat{\sigma}_i^2 / 2)$$

där λ är en kvantil för en standard normalfördelning svarande mot den valda konfidensgraden.

Prediktionsintervallet för hela avsättningen beräknas exakt som ovan för det enskilda skadeåret med \hat{R} istället för \hat{R}_i .

3.4 Braun

Modell för korrelationen mellan två skadetrianglar

Med Brauns modell beräknar vi korrelationen mellan två trianglar. Vi antar inte att korrelationen ska vara en förutbestämd konstant över alla årgångar utan den beräknas per utvecklingsår. Här nedan följer de antaganden vi har för de båda trianglar med de ackumulerade betalningarna $C_{i,k}$ och $D_{i,k}$ som vi vill undersöka:

Triangel 1	Triangel 2
$\hat{f}_k = \frac{\sum_{i=1}^{n+1-k} C_{i,k}}{\sum_{i=1}^{n+1-k} C_{i,k-1}}$	$\hat{g}_k = \frac{\sum_{i=1}^{n+1-k} D_{i,k}}{\sum_{i=1}^{n+1-k} D_{i,k-1}}$
$\hat{\sigma}_k^2 = \frac{1}{n-k} \sum_{i=2}^{n+1-k} C_{i,k-1} \left(\frac{C_{i,k}}{C_{i,k-1}} - \hat{f}_k \right)^2$	$\hat{\tau}_k^2 = \frac{1}{n-k} \sum_{i=2}^{n+1-k} D_{i,k-1} \left(\frac{D_{i,k}}{D_{i,k-1}} - \hat{g}_k \right)^2$
$F_{i,k} = \frac{C_{i,k}}{C_{i,k-1}}$	$G_{i,k} = \frac{D_{i,k}}{D_{i,k-1}}$
$\text{Var}(F_{i,k} T_{i,k-1}) = \frac{\hat{\sigma}_k^2}{C_{i,k-1}}$	$\text{Var}(G_{i,k} T_{i,k-1}) = \frac{\hat{\tau}_k^2}{D_{i,k-1}}$

Vi antar att korrelationen är ett enkelt parvis samband mellan $C_{i,k}$ och $D_{i,k}$. Vi generaliserar variansekvationerna ovan till:

$$\text{Cov}(F_{i,k}, G_{i,k} | T_{k-1}) = \frac{\rho_k}{\sqrt{C_{i,k-1} D_{i,k-1}}} \quad (3.4.1)$$

där T_{k-1} är mängden $\{C_{i,j}, D_{i,j} | 1 \leq i \leq n, 1 \leq j \leq k-1, i+j \leq n+1\}$. För att följa antagande två utvecklar vi oberoendet mellan skadeår till att omfatta båda portföljerna. Vi får därför att:

$$\text{Cov}(F_{i,k}, G_{j,k} | T_{k-1}) = 0 \text{ då } i \neq j.$$

På samma sätt som vi skattade σ_k och τ_k skattar vi ρ_k . Den enda skillnaden blir nämnaren, $1/(n-k-1+w_k^2)$ istället för $1/(n-k)$. Detta beror på att vi vill ha en väntevärdesriktig skattning

$$\hat{\rho}_k = \frac{1}{n-k-1+w_k^2} \sum_{i=1}^{n+1-k} \sqrt{C_{i,k-1} * D_{i,k-1}} * (F_{i,k} - \hat{f}_k)(G_{i,k} - \hat{g}_k) \quad (3.4.2)$$

där

$$w_k^2 := \frac{\left(\sum_{i=1}^{n+1-k} \sqrt{C_{i,k-1} D_{i,k-1}} \right)^2}{C_{<,k-1} D_{<,k-1}}.$$

Med Cauchy Schwarz olikhet kan vi visa att $0 \leq w_i \leq 1$.

Genom att använda definitionen av korrelationskoefficienten får vi fram att korrelationen mellan $F_{i,k}$ och $G_{i,k}$ ges av $\frac{\rho_k}{\sigma_k \tau_k}$.

Estimering av prediktionsfelet för de båda trianglarna för ett skadeår

Vi börjar som tidigare med att beräkna prediktionsfelet för det enskilda skadeåret.

Medelkvadratfelet för $\hat{C}_{i,n} + \hat{D}_{i,n}$ approximeras med det stokastiska felet och skattningsfelet:

$$\text{mse}(\hat{C}_{i,n} + \hat{D}_{i,n}) \approx \text{Var}(C_{i,n} + D_{i,n} | T_n) + \text{Var}(\hat{C}_{i,n} + \hat{D}_{i,n} | T_n) \quad (3.4.3)$$

där det stokastiska felet delas upp i:

$$\text{Var}(C_{i,n} + D_{i,n}) = \text{Var}(C_{i,n}) + 2\text{Cov}(C_{i,n}, D_{i,n}) + \text{Var}(D_{i,n}) \quad (3.4.4)$$

och skattningsfelet i:

$$\text{Var}(\hat{C}_{i,n} + \hat{D}_{i,n}) = \text{Var}(\hat{C}_{i,n}) + 2\text{Cov}(\hat{C}_{i,n}, \hat{D}_{i,n}) + \text{Var}(\hat{D}_{i,n}) \quad (3.4.5)$$

Genom formel (3.4.3), (3.4.4), (3.4.5) och (3.3.1) får vi fram följande skattning av medelkvadratfelet för ett skadeår:

$$\begin{aligned}
\hat{mse}(\hat{C}_{i,n} + \hat{D}_{i,n}) &\approx \hat{Var}(C_{i,n}) + \hat{Var}(D_{i,n}) + 2 \hat{Cov}(C_{i,n}, D_{i,n}) + \hat{Var}(\hat{C}_{i,n}) + \hat{Var}(\hat{D}_{i,n}) + 2 \hat{Cov}(\hat{C}_{i,n}, \hat{D}_{i,n}) \\
&= \hat{Var}(C_{i,n}) + \hat{Var}(\hat{C}_{i,n}) + \hat{Var}(D_{i,n}) + \hat{Var}(\hat{D}_{i,n}) + 2(\hat{Cov}(C_{i,n}, D_{i,n}) + \hat{Cov}(\hat{C}_{i,n}, \hat{D}_{i,n})) \\
&\approx \hat{mse}(\hat{C}_{i,n}) + \hat{mse}(\hat{D}_{i,n}) + 2(\hat{Cov}(C_{i,n}, D_{i,n}) + \hat{Cov}(\hat{C}_{i,n}, \hat{D}_{i,n}))
\end{aligned}$$

där de första termerna är medelkvadratfelet för de enskilda trianglarna som skattas genom (3.3.3). De respektive kovarianserna skattas med nedanstående rekursiva formler:

$$\begin{aligned}
\hat{Cov}(C_{i,k}, D_{i,k}) &= \hat{Cov}(C_{i,k-1}, D_{i,k-1}) \hat{f}_k \hat{g}_k + \sqrt{\hat{C}_{i,k-1} \hat{D}_{i,k-1}} \hat{\rho}_k \\
\hat{Cov}(C_{i,k}, D_{i,k}) &= \hat{Cov}(\hat{C}_{i,k-1}, \hat{D}_{i,k-1}) \hat{f}_k \hat{g}_k + \frac{\hat{C}_{i,k-1} \hat{D}_{i,k-1}}{\hat{C}_{<,k-1} \hat{D}_{<,k-1}} \hat{\rho}_k \sum_{j=1}^{n+1-k} \sqrt{\hat{C}_{j,k-1} \hat{D}_{j,k-1}}
\end{aligned}$$

med startvärden:

$$\begin{aligned}
\hat{Cov}(C_{i,n+1-i}, D_{i,n+1-i}) &= 0. \\
\hat{Cov}(\hat{C}_{i,n+1-i}, \hat{D}_{i,n+1-i}) &= 0.
\end{aligned}$$

Estimering av prediktionsfelet för de båda trianglarna

Tillvägagångssättet för hela triangeln är detsamma som för ett skadeår. Här får vi att approximationen av medelkvadratfelet blir:

$$\begin{aligned}
\hat{mse}\left(\sum_{i,j=2}^n (\hat{C}_{i,n} + \hat{D}_{j,n})\right) &\approx \hat{Var}\left(\sum_{i=2}^n C_{i,n}\right) + \hat{Var}\left(\sum_{j=2}^n D_{j,n}\right) + 2 \sum_{i,j=2}^n \hat{Cov}(C_{i,n}, D_{j,n}) \\
&+ \hat{Var}\left(\sum_{i=2}^n \hat{C}_{i,n}\right) + \hat{Var}\left(\sum_{j=2}^n \hat{D}_{j,n}\right) + 2 \sum_{1 \leq i,j \leq n} \hat{Cov}(\hat{C}_{i,n}, \hat{D}_{j,n} | T_{n+1-\min(i,j)})
\end{aligned}$$

Dess skattning blir då:

$$\hat{mse}\left(\sum_{i,j=2}^n (\hat{C}_{i,n} + \hat{D}_{j,n})\right) \approx \hat{mse}\left(\sum_{i=2}^n C_{i,n}\right) + \hat{mse}\left(\sum_{i=2}^n D_{i,n}\right) + 2\left(\hat{Cov}\left(\sum_{i=2}^n C_{i,n}, \sum_{i=2}^n D_{i,n}\right) + \sum_{1 \leq i,j \leq n} \hat{Cov}(\hat{C}_{i,n}, \hat{D}_{j,n})\right)$$

där de första två termerna hämtas ifrån rekursionen för totala medelkvadratfelet för en triangel (3.3.3) och kovarianserna skattas genom dessa två nedanstående ekvationer:

$$\hat{Cov}\left(\sum_{i=n+2-k}^n C_{i,k}, \sum_{i=n+2-k}^n D_{i,k}\right) = \hat{Cov}\left(\sum_{i=n+3-k}^n C_{i,k-1}, \sum_{i=n+3-k}^n D_{i,k-1}\right) \hat{f}_k \hat{g}_k + \hat{\rho}_k \sum_{i=n+2-k}^n \sqrt{\hat{C}_{i,k-1} \hat{D}_{i,k-1}}$$

startvärde $k=2$ och

$$\sum_{i,j=n+2-k}^n \hat{Cov}(\hat{C}_{i,k}, \hat{D}_{j,k}) = \sum_{i,j=n+3-k}^n \hat{Cov}(\hat{C}_{i,k-1}, \hat{D}_{j,k-1}) \hat{f}_k \hat{g}_k + \hat{C}_{\geq,k-1} \hat{D}_{\geq,k-1} \hat{\rho}_k \frac{\sum_{i=1}^{n+1-k} \sqrt{\hat{C}_{i,k-1} \hat{D}_{i,k-1}}}{\hat{C}_{<,k-1} \hat{D}_{<,k-1}}$$

med startvärde $k=n+1-\min(i,j)$.

Utifrån våra skattningar av kovarianserna kan vi beräkna den totala korrelationen mellan de båda trianglarna genom ekvationen:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\rho(X, Y)\sqrt{\text{Var}(X)\text{Var}(Y)}$$

$$\rho(X, Y) = \frac{\text{Var}(X + Y) - \text{Var}(X) - \text{Var}(Y)}{2\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

Där X och Y svarar mot prediktionsfelet för respektive triangel. Genom att ersätta $\text{Var}(X+Y)$ med vårt skattade medelkvadratfel för båda trianglarna och $\text{Var}(X)$ och $\text{Var}(Y)$ med de enskilda trianglarnas skattade medelkvadratfel kan vi beräkna korrelationen mellan trianglarna. Detta görs på både enskilda skadeår och för hela trianglarna.

3.5 Inflationsjustering

Enligt Braun och tesen i detta arbete borde korrelationen minska om man inflationsjusterar trianglarna.

Inflationsjustering innebär att vi räknar om triangeln till penningvärdet för det senast kända året, n. Vi utgår ifrån de inkrementella trianglarna. Den sista kända diagonalen är den där

$i + k = n + 1$. De skadeutbetalningar som har gjorts under år n är $\sum_{i=1}^n B_{i, n+1-i}$.

Vår indextabell:

År	Index
1	a_1
2	a_2
:	:
:	:
n-2	a_{n-2}
n-1	a_{n-1}
n	a_n

De inkrementella inflationsjusterade betalningarna betecknas med $L_{i,k}$,

där: $L_{i,k} = B_{i,k} * a_n / a_{i+k-1}$.

För att få våra ackumulerade betalningar adderas triangeln upp som i inledningen.

4 Genomförande och resultat

4.1 Data

Vi har använt oss av data från flera produkter per produktgren. Vi har valt produkter som är ganska stabila i storlek från år till år. Följande är produkterna:

Hem och Villa

Hemförsäkring
Fritidshusförsäkring
Villahemförsäkring

Motor

Trafik person
Trafik egendom

Sjuk och olycksfall

Individuell vuxenolycksfall
Kollektiv barnförsäkring
Individuell barnförsäkring

För att kunna presentera det datamaterial som har använts så har vi räknat om trianglarna till en egen valuta. Det innebär att alla faktorer och relationer är riktiga, dock blir inte den totala avsättningen och prediktionsfelet det.

Vi använder oss av de totala skadetrianglarna för produkterna. Vi får därför göra ett förenklande antagande om att trianglarna är homogena.

4.2 Utfall av kontroll av trianglar

4.2.1 Antagande 1

Kontroll av okorrelerade utvecklingsfaktorer. För de aktuella skadetrianglar får vi fram följande teststorhet av Spearmans test-statistika när vi använder oss av signifikansnivån 12.5 % per triangel. Analysen har gjorts på trianglar med 8, 10 respektive 14 års historik. De skuggadefälten indikerar att teststorheten ligger utanför det kritiska området, så att nollhypotesen, okorrelerade faktorer, inte förkastas.

Produkt	8 år	10 år	14 år
Villahem	0,396	0,471	0,236
Fritidshus	0,154	0,144	0,091
Hem	0,030	0,077	0,370
Trafik Person	0,006	0,517	0,370
Trafik Egendom	0,156	0,121	0,007
Ind. Barn	-0,128	0,075	0,127
Ind. Vuxen	-0,354	0,004	-0,158
Koll. Barn	0,038	0,106	0,179

De kritiska områdena utgörs av de värden som ligger utanför för 8 år (-0,395 0,395), för 10 år utanför (-0,289 0,289) samt för 14 år utanför (-0,188 0,188). Vi får alltså förkasta hypotesen för två respektive tre av trianglarna med 10 respektive 14 skadeår. För trianglarna med 8 år förkastas endast en triangel.

Vi gör en liten djupdykning i den skadetriangel där nollhypotesen förkastades för samtliga år, villaförsäkring. Villa utgör en väldigt kortsvansad affär. Med det menar man att tiden från skadans inträffande till att kunden har fått all ersättning är kort. Redan tre år efter skadeåret så har över 95 % av utbetalningarna redan lämnat bolaget. Den höga teststorheten genereras av de två första åren medan de kommande skadeåren är slumpmässigt fördelade som de borde vara. Det är alltså inte någon slump som har genererat denna höga teststorhet.

Genom att inflationsjustera data undersöker vi om det bli någon skillnad.

Antal år	E(T)			
	Raka data	Inflations-justerat	Min 12,5 % signifikans	Max 12,5 % signifikans
8 år	0,396	0,396	-0,395	0,395
10 år	0,471	0,471	-0,289	0,289
14 år	0,236	0,244	-0,188	0,188

Förändringen av teststorheternas värden är endast marginell.

Ytterligare en ansats till att få fram en lämplig triangel som man kan applicera Macks modell på görs genom att dela upp triangeln i de största skademomenten. Villa kan grovt delas upp i 1/3 Brand, 1/3 Vatten och 1/3 övriga skadekostnader. Testet genomförs igen med de tre uppdelningarna.

För triangeln med 10 år fås följande resultat:

Skademoment	E(T)			
	Raka data	Inflations-justerat	Min 12,5 % signifikans	Max 12,5 % signifikans
Brand	0,161	0,067	-0,289	0,289
Vatten	0,077	0,047	-0,289	0,289
Övrigt	0,296	0,319	-0,289	0,289

Med denna uppdelning hamnar både Brand och Vatten utanför det kritiska området. När det gäller övrigt så består den av många mindre skademoment där en omfördelning av kostnadsmassan mellan de olika momenten är förklaringen till att den blir signifikant. De inflationsjusterade trianglarna för brand och vatten ligger även inom Macks snävare 50 % konfidensintervall.

Som uppmärksammas i inledningen är det svårt att använda sig av trianglar på totalnivå då man har olika betalflöden inom olika delmoment vars andel av den totala skadekostnaden kan variera över tiden. Därmed finns det en risk för att den sammanslagna triangeln indikerar på korrelation mellan de nästkommande utvecklingsfaktorerna när den egentliga orsaken är förändring i totala skademassan mellan de olika skademomenten.

Trots att både trafik person och villa blev signifikant skilda ifrån nollhypotesen, att ingen korrelation mellan efterföljande faktorer finns, så väljer vi att gå vidare med trianglarna på 10 år. Vi skulle nå ett bättre resultat i testet om vi momentuppdelade skadetriangeln, men målet

för detta arbete är att kontrollera korrelationen mellan de olika trianglarna så därmed är vi acceptera med att 6 av 8 trianglar föll väl ut i testet.

4.2.2 Antagande 2

Vid kontroll av oberoende mellan skadeår eller kalenderårsvis oberoende för triangeln med 10 år finner jag följande:

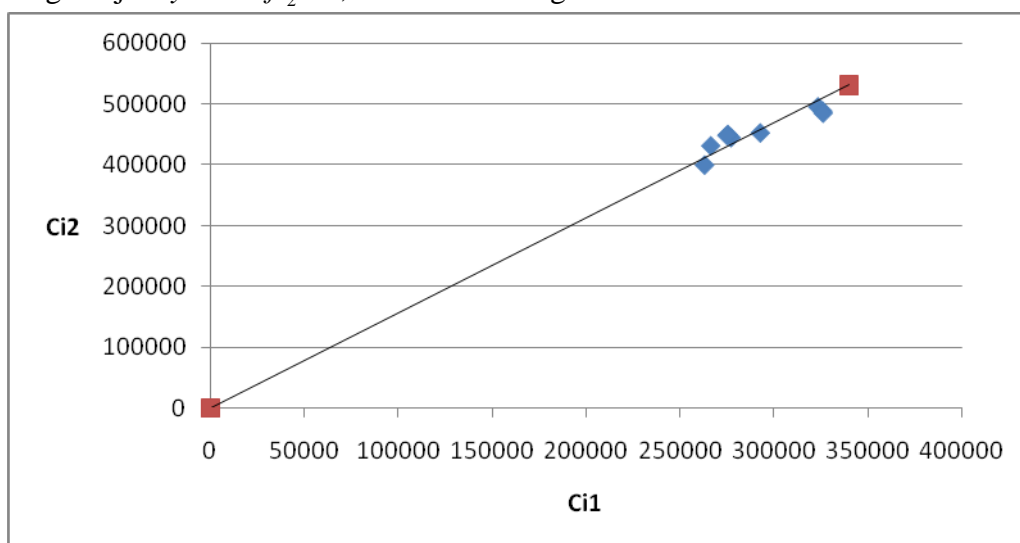
Produkt	Z	Min 5% signifikans	Max 5% signifikans
Villa	13,0	8,8	16,5
Fritidshus	14,0	8,9	16,6
Hem	10,0	9,0	16,6
Trafik Person	9,0	8,8	16,2
Trafik Egendom	13,0	8,9	16,2
Ind. Barn	11,0	9,0	16,9
Ind. Vuxen	11,0	8,8	16,2
Koll. Barn	10,0	8,9	16,2

För alla trianglar ligger teststorheten utanför det kritiska området. Vi kan därmed inte förkasta att skadeåren är oberoende.

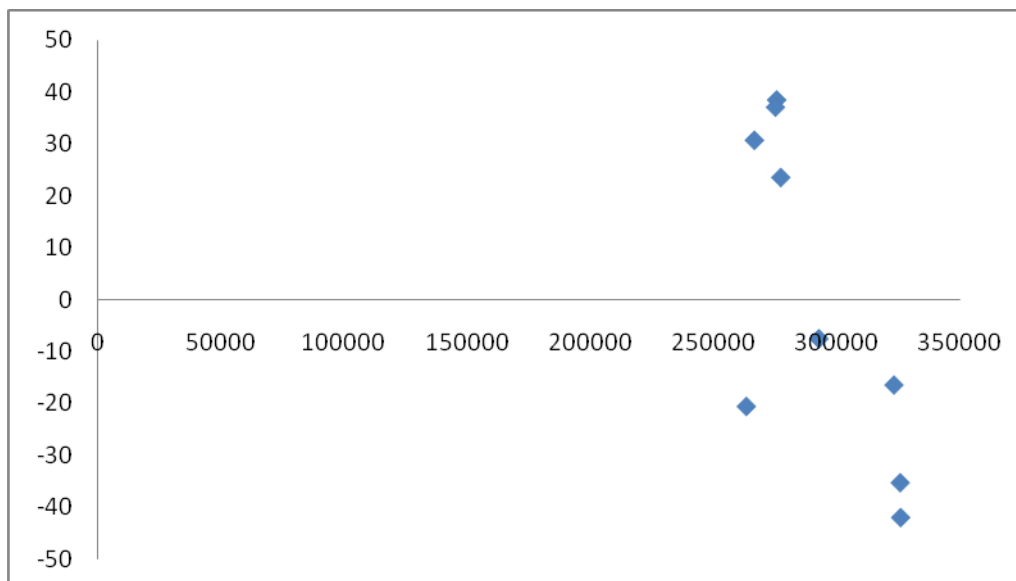
4.2.3 Visuell kontroll av antagande 1 och 3

Av utrymmesskäl har vi valt att inte kontrollera alla utvecklingsfaktor och residualer för alla trianglar. En kontroll av detta skulle innebära 16 plottar för vardera av våra 8 trianglar och därmed 128 plottar totalt. Nedan följer ett exempel för trafik egendom:

Vi börjar med att plotta det ackumulerade utbetalda beloppet till och med år två mot utbetalningarna för år ett. Enligt antagande 1 förväntar vi oss att se en slumpmässig fördelning längs linjen $y = 0 + \hat{f}_2 * x$, vilket vi också gör.



Då de standardiserade residualerna för utvecklingsår två, $(C_{i,2} - C_{i,1}) / \sqrt{C_{i,1}}$ plottas mot $C_{i,1}$ får vi följande utfall:

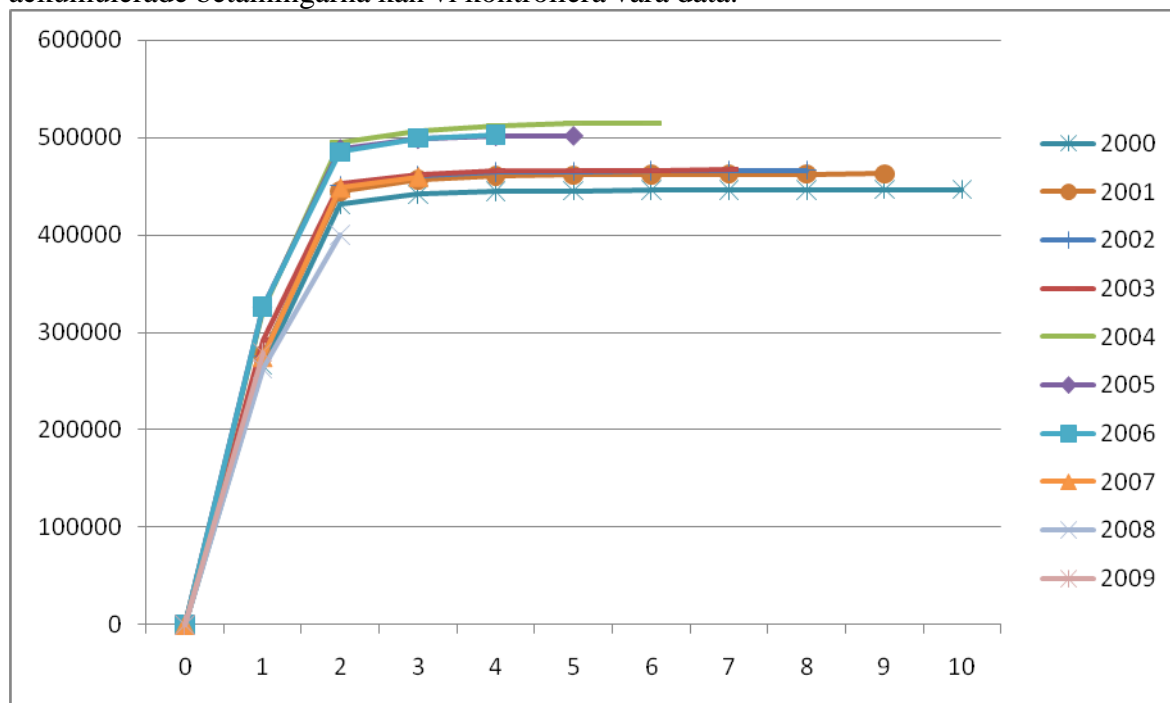


även här ser det relativt slumpmässigt ut. Det som försvårar analysen är att de ackumulerade betalningarna för år ett är relativt stabila. Med en större spridning av dem hade det varit lättare att urskilja avvikande mönster.

Vi väljer att gå vidare med våra trianglar trots att de tidigare testerna inte fullt visat att Chain-ladder metoden är tillämpbar. Det är ett för kraftigt antagande att anta att de totala trianglarna per produkt är homogena. Det finns även en tendens till att faktorerna är korrelerade för vissa produkter. Vi har dock anpassat storleken på trianglarna för att minimera detta och få så bra underlag som möjligt.

4.3 Avsättning och osäkerhet för en triangel

Vi börjar med en skadetriangel för trafik egendom. Genom att göra en graf av de ackumulerade betalningarna kan vi kontrollera våra data.

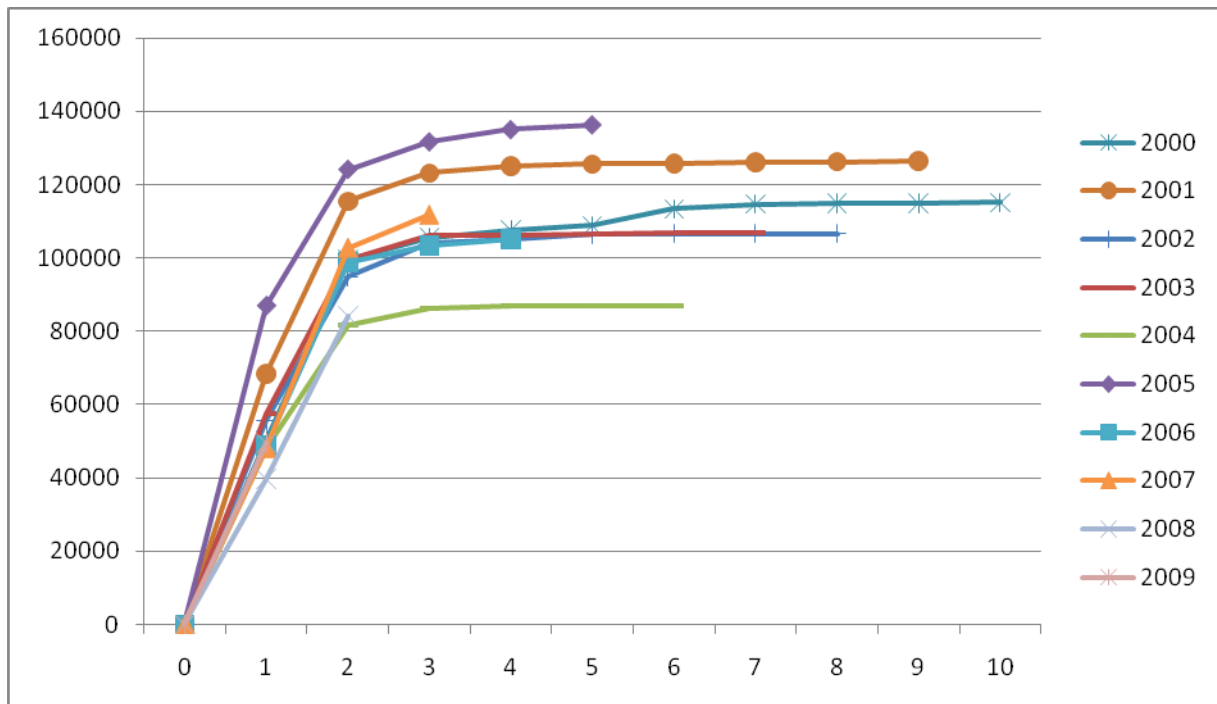


Y-axeln betecknar det ackumulerade utbetalda beloppet, x-axeln utvecklingsåren och graferna skadeåren. Alla grafer har en liknande form och vi kan tydligt se att de planar av efter utvecklingsår fyra. Detta kan vi även se i tabellen över de skattade utvecklingsfaktorer nedan där t.ex. faktorn för utvecklingsår fem är skattad till 1,002 och för utvecklingsår åtta är den 1,000. Därmed uppfyller trafik egendom vårt antagande om att alla utbetalningar ska ske inom 10 år.

ÅR	Skade- avsättning (tkr)	Medelfel	Utvecklingsår	\hat{f}_k	$\hat{\sigma}_k^2$
2000	0	0	2	1,561	1015,4
2001	36	72	3	1,024	4,7
2002	148	101	4	1,008	1,8
2003	267	127	5	1,002	1,2
2004	599	214	6	1,001	0,3
2005	1 170	463	7	1,001	0,0
2006	2 219	958	8	1,000	0,0
2007	5 656	1 336	9	1,000	0,0
2008	14 568	1 922	10	1,000	0,0
2009	172 776	18 478			
Totalt	197 440	18 742			

Faktorerna avtar stabilt över tiden som de bör. Ett 90 % prediktionsintervall för totala avsättningen under antagande av lognormal fördelning beräknas till (168 200, 229 693).

För vår andra triangel som är fritidshus, där $D_{i,k}$ betecknar det ackumulerade utbetalda beloppet, får vi fram följande:



Här har vi en betydligt större spridning mellan skadeårgångarna. Detta beror på att fritidshus är en mindre produkt än trafik egendom. Utfallet blir därmed mer slumpmässigt. Vi kan ändå se att de olika skadeårgångarna följer samma mönster. Efter utvecklingsår sex är faktorerna så pass små att ett antagande om att betalningarna sker inom 10 år är rimligt.

ÅR	Skade- avsättning	Medelfel	Utvecklingsår	\hat{g}_k	$\hat{\tau}_k$
2000	0	0	2	1,791	3500,0
2001	193	183	3	1,067	26,9
2002	278	255	4	1,014	7,4
2003	380	288	5	1,008	3,0
2004	625	630	6	1,009	36,8
2005	2 266	2 661	7	1,004	3,2
2006	2 580	2 370	8	1,001	0,1
2007	4 326	2 669	9	1,001	0,3
2008	9 157	2 870	10	1,002	0,1
2009	48 822	15 582			
Totalt	68 626	17 140			

Här beräknas vårt 90 % prediktionsintervall för lognormal till (44 424, 99 788).

Vi delar upp produkterna mellan två separata försäkringsbolag där bolagen endast bedriver verksamhet i en av produkterna. Försäkringsbolag A med trafik egendom behöver avsätta 197,4 mkr i avsättning för oregerade skador och ytterligare $229,7 - 197,4 = 32,3$ mkr i extra kapital för att kunna täcka kostnaden i 19 av 20 fall. Försäkringsbolag B med fritidshus behöver avsätta 68,6 mkr i avsättning för oregerade skador och ytterligare 31,1 mkr i extra kapital.

Prediktionsintervall för båda trianglarna vid oberoende.

Bolag A och B genomgår en sammanslagning och därmed vill vi beräkna prediktionsintervallet för hela bolagets nya portfölj, dvs. båda produkterna. Vi börjar med att anta att de är oberoende. För att beräkna prediktionsintervallet för oberoende trianglar så gör vi följande:

De totala avsättningarna är för Fritidshus respektive Trafik egendom:

$$\hat{R}_1 = \sum_{i=2}^{10} \hat{C}_{i,10} - \sum_{i=2}^{10} \hat{C}_{i,n+1-i} = 68626$$
$$\hat{R}_2 = \sum_{i=2}^{10} \hat{D}_{i,10} - \sum_{i=2}^{10} \hat{D}_{i,n+1-i} = 197440.$$

R är den sammanlagda avsättningen och vi har att:

$$\hat{R} = \sum_{i=1}^2 \hat{R}_i = 68626 + 197440 = 266065.$$

För variansen får vi följande:

$$Var\left(\sum_{i=1}^n (\hat{R}_i - R_i)\right) = \sum_{i=1}^n Var(\hat{R}_i - R_i) + 2 \sum_{i < j} Cov(\hat{R}_i - R_i, \hat{R}_j - R_j)$$

Om avsättningarna är oberoende så blir kovariansen 0. Det innebär att vi får

$$mse\left(\sum_{i=1}^2 \hat{R}_i\right) = \sum_{i=1}^2 mse(\hat{R}_i) = 17140^2 + 18742^2 = 25398^2$$

Då blir ett 90 % prediktionsintervall för lognormal (226 454, 309 782) för våra båda trianglar. Detta innebär att avsättningen för oreglerade skador är densamma men genom att gå ihop minskar de det ytterligare kapitalkravet från 31,1+32,3=63,4 till 309,8-266,1=43,7 mkr.

4.4 Avsättning och osäkerhet för två trianglar

Brauns exempel:

I Brauns kompendium avslutar han med ett exempel med data från Reinsurance Association of America. För vår programmerade modell så får vi samma värden som Braun presenterar i tabellerna på skadeavsättning, medelfel och korrelation. I den löpande texten anger han ett 90 % prediktionsintervall med antagen lognormal fördelning för avsättningen till (7 459 480, 9 157 228). Med samma antaganden vi får fram prediktionsintervallet (7 409 266, 9 082 103), där beräkning är avstämd med exempel från Mack (1994). Både den undre och övre gränsen av Brauns intervall är högre än det intervall som vi har beräknat.

Återkopplar vi till den beräkning av prediktionsintervallet som presenterades i kapitel 3.2 ges de yttre gränserna för intervallet av:

$$\hat{R} * \exp(\pm \lambda * \sigma - \sigma^2 / 2)$$

Vi antar att de yttre gränserna av intervallet samt skattningen av avsättning och medelfelet är korrekta och beräknar λ , kvantilen i en standard normalfördelning svarande mot den valda

konfidensgraden. Vi får då fram 1,54 för den lägre gränsen i prediktionsintervallet och 1,78 för den högre. För ett 90 prediktionsintervall är $\lambda = 1,64$ för både den övre och undre gränsen. Därmed har Braun utgått ifrån en annan skattning av skadeavsättningen för beräkning av prediktionsintervallet.

Två trianglar:

Vi fortsätter med fritidshus och trafik egendom. Istället för att anta att trianglarna är oberoende beräknar vi dess korrelation genom Brauns modell och får fram följande:

ÅR	Fritidshus		Trafik Egendom		Braun	
	\hat{g}_k	$\hat{\tau}_k$	\hat{f}_k	$\hat{\sigma}_k$	w_k^2	$\hat{\rho}_k / (\hat{\sigma}_k \hat{\tau}_k)$
2	1,791	3 500,0	1,561	1 015,4	0,989	0,231
3	1,067	26,9	1,024	4,7	0,996	-0,170
4	1,014	7,4	1,008	1,8	0,995	-0,500
5	1,008	3,0	1,002	1,2	0,994	-0,500
6	1,009	36,8	1,001	0,3	0,994	0,423
7	1,004	3,2	1,001	0,0	0,999	-0,521
8	1,001	0,1	1,000	0,0	0,999	0,770
9	1,001	0,3	1,000	0,0	1,000	-1,000
10	1,002	0,1	1,000	0,0	1,000	

Vi ser att w_k^2 är väldigt nära 1, så en god approximation av första delen av ekvation (2.3.8) är

$$\text{således } \frac{1}{n-k-1+w_k^2} \approx \frac{1}{n-k}.$$

Skattningen av korrelationen $\hat{\rho}_k / (\hat{\sigma}_k \hat{\tau}_k)$ mellan utvecklingsåren varierar kraftigt mellan åren. Vi har en positiv korrelation för år 2,6,8 och resterande korrelationer är negativa. Detta beror dels på ett litet antal observationer för de senare åren att skatta korrelationen med men i detta fall beror det främst på att vi jämför två stycken väldigt kortsvarsade affärer. För trafik egendom kvarstår ca 1 % av betalningarna efter år 3 och för Fritidshus är det ca 4 %. Det innebär att det inte är intressant att titta på korrelationen för de senare utvecklingsåren, slumpen har för stor inverkan.

Om vi istället jämför två långsvansade affärer, där utbetalningsflödet är betydligt längre, får vi en betydligt stabilare korrelation över utvecklingsåren.

Nedan följer medelfelet och korrelationen för fritidshus och trafik egendom både enskilt och sammanslaget.

Skadeår	Fritidshus		Trafik egendom		Totalt		
	Avsättning	Medelfel	Avsättning	Medelfel	Avsättning	Medelfel	Korrelation
2000	0	0	0	0	0	0	0,00
2001	193	183	36	72	229	196	0,00
2002	278	255	148	101	425	217	-0,55
2003	380	288	267	127	646	294	-0,17
2004	625	630	599	214	1 224	576	-0,41
2005	2 266	2 661	1 170	463	3 437	2 834	0,30
2006	2 580	2 370	2 219	958	4 799	2 580	0,03
2007	4 326	2 669	5 656	1 336	9 982	2 839	-0,12
2008	9 157	2 870	14 568	1 922	23 725	3 227	-0,14
2009	48 822	15 582	172 776	18 478	221 597	26 695	0,22
Totalt	68 626	17 140	197 440	18 742	266 065	27 780	0,20

Vi ser att korrelationen för hela trianglarna blir 0,20. Den beror till största delen på den korrelation som vi såg för utvecklingsår två, vilken endast påverkar skadeår 2009. Hade vi jämfört två långsvansade trianglarna hade vi även här fått en stabilare korrelation över skadeåren.

Vi får vår skattning av standardavvikelsen till att bli 27 780 mot 25 398 när vi i kapitel 4.3 antog oberoende. Därmed ökar osäkerheten något. Vi väljer även att titta på standardavvikelsen då vi antar att trianglarna är fullt positivt respektive fullt negativ och får därmed följande utfall:

Korrelation	Medelfelet
$\rho=1$	35 883
$\rho=0,20$	27 780
$\rho=0$	25 398
$\rho=-1$	1 602

Att trianglarna är positivt korrelerade innebär en ökad risk för oss. Om vi avsätter för lite pengar för fritidshus finns det en risk att vi har gjort detsamma för trafik egendom.

Om vi återkopplar till vårt tidigare prediktionsintervall så får vi följande ytterligare kapitalkrav för produkterna vid olika korrelation. Vi ser att

90 % prediktionsintervall

Korrelation	Undre	Övre	Ytterligare kapitalkrav
1,00	211 429	328 839	62 774
0,20	222 971	314 065	47 999
0,00	226 454	309 782	43 717
-1,00	263 438	268 709	2 644

Det ytterligare kapitalkravet går från 3 mkr vid fullständig negativ korrelation till 63 mkr vid fullständig positiv korrelation.

Ett prediktionsintervall med en positiv korrelation är bredare än om det hade varit okorrelerat, och smalast är de prediktionsintervall där trianglarna är negativt korrelerade.

Alla trianglar:

Vid en parvis jämförelse av alla våra skadetrianglar får vi följande korrelationsmatris:

	Hem	Villa- hem	Fritids- hus	Trafik Person	Trafik Egendom	Ind Vux	Ind Barn	Koll Barn
Hem	100%							
Villahem	63%	100%						
Fritidshus	28%	65%	100%					
Trafik Person	-1%	-2%	7%	100%				
Trafik Egendom	-24%	1%	20%	5%	100%			
Ind Vux	-24%	-14%	15%	-17%	16%	100%		
Ind Barn	-55%	-45%	-36%	-19%	6%	44%	100%	
Koll Barn	-46%	-44%	-50%	1%	-8%	8%	68%	100%

De skuggade fälten har en positiv korrelationen och de övriga har en negativ korrelation.

Korrelationsmatrisens utseende är inte helt oväntat. I det stora så är de produkter som liknar varandra positivt korrelerade och olika produkttyper är negativ korrelerade. Hem, villa, fritidshus och trafik egendom är alla kortsvansad till största del egendomsaffär. De är alla positivt korrelerade med varandra förutom hem och trafikegendom. För personolycksfallsförsäkringarna kollektiv barn, individuell vuxen och individuell barn är alla dessa positivt korrelerade.

Det som är oväntat är att trafik person enbart är svagt positivt korrelerad med trafik egendom och kollektiv barnförsäkring. På grund av de långsvansade personskadorna hade vi trott att den skulle ha en positiv korrelation med olycksfallsförsäkringarna. En möjlig orsak till detta är att när vi endast tittar på trianglar med 10 skadeår så fångar vi inte upp det.

4.5 Utfall efter inflationsjustering

Vad händer med korrelationen när trianglarna inflationsjusteras? Enligt vår tes så borde korrelationen minska då data inflationsjusteras.

Trianglarna inflationsjusteras med index från SCB. Vi har justerat med KPI, löneindex och prisbasbeloppet.

Tabellen nedan innehåller differensen av korrelationerna, dvs. $|\rho_{\text{inflationjusterat}}| - |\rho_{\text{raka_data}}|$, där $|\cdot|$ anger absolutbeloppet:

	Hem	Villa- hem	Fritids- hus	Trafik Person	Trafik Egendom	Ind Vux	Ind Barn	Koll Barn
Hem								
Villahem	1,6%							
Fritidshus	-1,1%	-1,8%						
Trafik Person	-0,1%	0,0%	0,7%					
Trafik Egendom	2,3%	4,4%	-2,6%	0,4%				
Ind Vux	0,8%	1,0%	0,7%	-0,2%	1,4%			
Ind Barn	0,8%	2,8%	2,4%	-0,3%	-2,0%	-3,2%		
Koll Barn	-0,4%	1,2%	2,3%	1,4%	1,4%	-4,0%	-0,2%	

Det innebär att om korrelationen går mot noll då data inflationsjusteras så kommer vi få ett negativt värde i tabellen. De negativa värdena är skuggade i tabellen, det är 11 av 28 stycken (39 %) av våra parvisa trianglar. Vi kan därmed inte se någon effekt på korrelationen med den genomförda inflationsjusteringen och förändringen i korrelationen är ytterst marginell på maximalt 4,4 procentenheter.

5 Slutsatser

Målet med arbetet var att kvantifiera korrelationen mellan olika produkter samt att undersöka om den konstaterade korrelationen kan förklaras av inflationen.

Den korrelationsmatris som vi fick fram var ganska väntad. Kortsvansad affär såsom hem och villa samt trafik egendom var i de flesta fall positivt korrelerade med varandra, likaså för den långsvansade olycksfallsaffären. Det som avvek var trafik person. Vi hade trott att den skulle vara mer positivt korrelerad med olycksfallsaffären. Det är dock möjligt att vår begränsade triangel på 10 år inte fångar upp den effekten.

Efter inflationsjusteringen av trianglarna minskar tyvärr inte korrelationen, den är i stort sett oförändrad. Vi får därmed förkasta hypotesen att SCBs index förklarar korrelationen mellan trianglarna.

Ett sätt att gå vidare är att söka efter något annat index som kan passa utvecklingen bättre eller att genom separationsmetoden skatta inflationen i trianglarna.

6 Referenser

Braun, C (2004): The prediction error of the chain ladder method applied to correlated run-off triangles. ASTIN Bulletin, Vol.34, No. 2, 399-423

Dahl, P. (2003): Introduction to Reserving

Mack, T (1993): Distribution-free calculation of the standard error of chain ladder reserve estimates. ASTIN Bulletin, Vol. 23, No. 2, 213-225.

Mack, T. (1994): Measuring the variability of chain ladder reserve estimates. Casualty Actuarial Society, Spring, Vol 1, 101-182.

Mack, T (1999): The standard error of chain ladder reserve estimates: Recursive calculation and inclusion of a tail factor. ASTIN Bulletin, Vol. 29, No. 2, 361-366

Statens offentliga utredningar (2007): Sverige inför klimatförändringar – hot och möjligheter, Slutbetänkande av klimat- och sårbarhetsutredningen, 2007:60, 495-496

7 Appendix

Skadetrianglarna

Akkumulerad utbetalningstriangel för fritidshus.

	1	2	3	4	5	6	7	8	9	10
2000	50 086	99 686	105 550	107 466	108 822	113 400	114 684	114 923	114 923	115 098
2001	68 287	115 487	123 172	125 120	125 584	125 764	126 122	126 212	126 472	
2002	55 680	94 902	104 030	104 944	106 467	106 663	106 663	106 663		
2003	57 429	99 423	106 088	106 251	106 644	106 861	106 861			
2004	48 072	81 682	86 350	86 889	87 039	86 874				
2005	87 015	124 173	131 688	135 031	136 332					
2006	48 821	98 865	103 394	105 108						
2007	48 041	102 672	111 660							
2008	39 688	84 211								
2009	49 526									

Akkumulerad utbetalningstriangel för trafik egendom.

	1	2	3	4	5	6	7	8	9	10
2000	266 305	431 645	441 548	444 810	445 336	446 115	446 282	446 435	446 582	446 617
2001	277 092	445 024	457 042	460 603	461 409	462 154	462 413	462 567	462 637	
2002	275 397	450 182	460 679	464 289	465 169	466 039	466 228	466 270		
2003	292 609	452 773	462 340	465 616	466 268	466 439	466 915			
2004	323 236	495 311	506 107	512 111	514 708	514 889				
2005	325 802	488 525	498 355	501 322	501 780					
2006	326 005	484 997	499 275	502 704						
2007	274 885	448 632	459 799							
2008	263 018	400 087								
2009	279 544									