Matematisk statistik
Stockholms universitet

Variansjämförelse av excess-of-loss-kontrakt med och utan aggregerat självbehåll

Sabina Jusupovic

Examensarbete 2003:9

Postadress:

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm
Sverige

Internet:

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

Matematisk statistik
Stockholms universitet
Examensarbete **2003:9**,
http://www.math.su.se/matstat

Variansjämförelse av excess-of-loss-kontrakt med och utan aggregerat självbehåll

Sabina Jusupovic*

November 2003

Abstract

The aim of this work is to compare the variance of the loss (claim amount) caused by two different reinsurance contracts, given that they have the same expected value. The reinsurance contracts considered are ordinary excess of loss and excess of loss with aggregate retention.

A parametric distribution is fitted to observed claim amounts from the Swedish insurance company Trygg-Hansa. The distributions taken into consideration are Pareto and lognormal and parameter estimation is made by the maximum likelihood method and the method of moments.

Sammanfattning

Den här rapporten är resultatet av en studie som jag utfört vid Trygg-Hansa. Målet har varit att jämföra effekten av olika typer av återförsäkringslösningar inom skadeförsäkring. Jämförelsen går ut på att undersöka hur skadekostnadens varians påverkas av valet av återförsäkringskontrakt, givet att de har samma väntevärde. De två kontrakt som analyseras är ett vanligt excess of loss kontrakt och ett excess of loss kontrakt med aggregerat självbehåll.

För att kunna göra jämförelsen bestämdes en parametrisk fördelning för skadebeloppen inom återförsäkring, vilken sedan användes för att beräkna den förväntade kostnaden och variansen. Fördelningar som betraktas är Pareto- och lognormalfördelning. Parameterskattningar görs med ML- och momentmetoden. Uppgiften är utförd utifrån önskemål från Trygg-Hansa och de står för skaderegisterdata.

*E-post: sabina2005@hotmail.com. Handledare: Esbjörn Ohlsson.

Innehållsförteckning

Förord

Detta arbete utgör ett 20 poängs examensarbete vid Stockholms Universitet, i matematisk statistik. Studien har genomförts på försäkringsbolaget Trygg-Hansa.
Jag skulle vilja tacka alla som hjälpt mig under arbetet med denna uppsats. Förs och främst vill jag tacka min handledare på Trygg-Hansa, Erik Hevreng, samt Roland Svensk och Anders Lindstöm, aktuarier på Trygg-Hansa. Jag har haft stor glädje och nytta av att samarbeta med Er. Ett speciellt tack vill jag säga till min handledare på Stockholms Universitet, Esbjörn Ohlsson för allt stöd och för all värdefull hjälp.

Företagsfakta

Trygg-Hansa är ett av de ledande skadeförsäkringsbolagen i Sverige med 1.8 miljoner kunder. Företaget finns idag på 34 orter och har ca 1600 anställda. Bolaget erbjuder ett heltäckande sortiment av sakförsäkringar till privatpersoner och företag. Trygg-Hansa startades 1828 men då under namnet ”Städernas Allmänna Brandstodsbolag”.
Fram till 1970-talet var företaget med om ett 30-tal fusioner och namnet ändrades ett flertal gånger. 1971 fick bolaget det namnet som fortfarande används. Sedan oktober 1999 ingår Trygg-Hansa i den danska försäkringskoncernen Codan, vilken i sin tur ägs av Royal & Sun Alliance, som är ett av världens största försäkringsbolag.

1 Introduktion

Trygg-Hansa behöver kunna uppskatta hur många skador, över en viss nivå, som kommer att inträffa under ett år, samt deras kostnad. Genom att beräkna den förväntade kostnaden väljer bolaget om det ska behöva köpa återförsäkring samt vilken typ av återförsäkringsskydd som passar bäst för dem. Syfte med detta arbete är att ge vägledning vid utvärdering av olika återförsäkringsalternativ genom att jämföra effekten av olika typer av återförsäkringslösningar.

Uppgiften är att jämföra *ett vanligt excess of loss kontrakt* med *ett excess of loss kontrakt med aggregerat självbehåll* ur varianssynpunkt. Jämförelsen går ut på att undersöka hur den totala skadekostnadens varians påverkas av utformningen av de olika kontrakten vid samma väntevärde.

Genom att utnyttja information från skador som inträffade under tiden 1995-2001 försökte vi göra en bedömning om skadekostnaden för återförsäkringsbolaget. Datamaterialet som vi använde var observationer av skadekostnad som inträffade under perioden 1995-2001. Varje observation i datamaterialet är den skadekostnad som återförsäkringsbolaget betalar.

Data som använts kan delas upp enligt

- *Ansvar*
- *Trafik*
- *Olycksfall*
- *Egendom*

Där *ansvar* är ansvarsdelen från företagsförsäkring och ansvarsmomentet i boendeförsäkringar, *olycksfall* är individuellt tecknade sjuk & olycksfallsförsäkringar och *trafikförsäkring* är trafikmomentet i motorförsäkring. Slutligen är *egendom* all övrig civil- (hem, villa, villahem, fritidshus, fastighet ...), motorkasko och företagsförsäkring.

Uppgiften ska koncentreras på *egendom*. Skadegrenarna civil, företag och kasko kommer att analyseras var för sig. För analysen används SAS samt Microsoft Excel. Rapporten är skriven i Microsoft Word.

2 Återförsäkring

2.1 Allmänt om återförsäkring

Detta avsnitt bygger på Gustafsson [6].

Försäkringsbranschen har drabbats hårt av rekordmånga naturkatastrofer under 1990-talet. Förluster som har orsakats av sådana skador överstiger ofta försäkringsbolagens kapacitet. För att förhindra detta samt att utjämna resultatet väljer ofta försäkringsbolag att ”sälja bort” bitar av stora risker man ekonomiskt bedömer sig inte vilja eller kunna ta till en annan försäkringsgivare. Reassurans eller återförsäkring är i princip försäkring i ytterligare ett led och går ut på att ett försäkringsbolag, som i detta sammanhang kallas cedent, överför en större eller mindre del av en försäkrad risk, eller en samling försäkrade risker, på en eller flera återförsäkrare mot en återförsäkringspremie. Från den ursprunglige försäkringstagarens synpunkt sker ingen förändring, hans försäkringsgivare står för hela det ansvar som försäkringsavtalet innebär.

Återförsäkringen uppstod ursprungligen ur transportförsäkringen någon gång på 1300-talet och ordet återförsäkra dyker då för första gången upp i Florens. Återförsäkring är till sin karaktär starkt internationell. Genom återförsäkringsförfarandet sprids de stora riskerna över hela värden. Nordamerika och Västeuropa dominerar den globala återförsäkringsmarknaden.

Återförsäkringen har två huvuduppgifter:

1. *Att skydda det enskilda bolaget mot katastrofskador*
2. *Att utjämna det enskilda bolagets årliga affärsresultat*

Definitionsmässigt kan återförsäkringen indelas i dels *obligatorisk* och *fakultativ*, dels *proportionell* och *icke proportionell* återförsäkring.

Med *obligatoriskt* menas återförsäkring som avtalats på förhand med bindande verkan för båda parter medan fakultativ innebär att varje risk eller försäkring återförsäkras individuellt och ingen av parterna är i förväg bunden av något avtal. I modern återförsäkring tillämpas den fakultativa typen endast för mycket stora och komplicerade risker.

Med *proportionell* återförsäkring menas att återförsäkraren tar ansvaret för en viss andel av den ursprungliga försäkringen, och erhåller en motsvarande del av premien som cedenten i sin tur erhåller från försäkringstagaren.

I en *icke-proportionell* återförsäkring ansvarar återförsäkraren för skadebelopp som överstiger en avtalad skadegräns, vilken kan bestämmas per risk, per skadehändelse eller för en viss period. Fördelen med en icke-proportionell återförsäkring är att cedenten lättare kan förutsäga resultaten, eftersom han vet att han inte kommer att betala för skador som överstiger en viss gräns. Man kan kombinera olika typer av återförsäkring beroende på behovet och de omständigheter som förekommer.

Ett självbehåll innebär att återförsäkringsbolaget inte ersätter skador vilkas kostnader understiger en viss gräns, den så kallade excesspunkten.

Den *proportionella* återförsäkringen kan indelas i:

1. *Kvotåterförsäkring*
2. *Excedentåterförsäkring*

Kvotåterförsäkring innebär att alla försäkringar delats enligt en given kvot t ex 10 % självbehåll och 90 % återförsäkring. Metoden innebär att även den minsta skada, som bolaget mycket väl skulle kunna ta helt för egen del, återförsäkras.

Excedentåterförsäkring är numera den vanligaste formen av proportionell obligatorisk återförsäkring. Sedan cedenten fastställt sitt självbehåll (excesspunkt) på den individuella risken återförsäkras det överskjutande beloppet – excedenten - automatiskt under excedentkontraktet inom ramen för dettas totala kapacitet. Procentuellt kan självbehållet därigenom variera från 100 % på de mindre riskerna som alltså inte blir föremål för återförsäkring ned till en eller ett par procent på de största riskerna.

Den *icke-proportionella* återförsäkringen kan indelas i

1. *Excess of loss*
2. *Stop loss*

Gemensamt för dem båda är att endast de skador som överskrider en viss gräns betalas av återförsäkrarna. Småskador betalas i sin helhet av försäkringsbolaget.
Excess of loss är den vanliga formen av icke-proportionell återförsäkring och kan vara såväl *obligatorisk* som *fakultativ*. *Excess of loss*-återförsäkring innebär att återförsäkrarens betalningsansvar inträder först när en individuell skada överstiger en på förhand avtalad gräns eller när det sammanlagda ersättningsbeloppet för skador som orsakats av en och samma skadehändelse överstiger en sådan gräns. Gränsen för återförsäkrarens inträde benämns *excesspunkt* och omfattningen av skyddet som oftast är beloppsbegränsat kallas för *cover*. En excess of loss återförsäkring byggs ofta upp på så sätt att det totala skyddet spjälkas upp i ett antal *layers* som staplas ovanpå varandra upp till en nivå som cedenten anser som betryggande skydd för varje tänkbar skadehändelse. Skadebelopp som överstiger skyddets övre gräns brukar kallas *spill over* och faller direkt tillbaka på cedentens självbehåll.

Stop loss-återförsäkring arbetar med det ackumulerade skadeutfallet under en period. Den är alltid obligatorisk och följaktligen kontraktsbunden. Återförsäkraren kommer in i bilden då summan av alla skador passerar ett bestämt belopp under ett år, i övrigt fungerar det som excess of loss. *Stopp loss-återförsäkring* är lämpligt inom områden där skadeutfallet växlar mycket mellan olika år. Exempel på detta är stormförsäkringar. Val av återförsäkringsmetod påverkas av en mängd faktorer. I den här rapporten kommer vi endast att betrakta *excess of loss* återförsäkring.

2.2 Kortfattad projektbeskrivning

2.2.1 Inflation

Eftersom datamaterialet spänner över en 7-årsperid under vilken penningvärdet ändrats behöver man korrigera skadekostnaderna med ett lämpligt prisindex till aktuella dagsvärden.

Tidigare års skadekostnader räknas upp med KPI så att alla data är i penningvärdet som gällde år 2002.

2.2.2 Excess of loss med och utan aggregat

Trygg-Hansa bestämmer sig för ett självbehåll c *Mkr* i varje skada, och tecknar skydd med excesspunkt c_i *MSK* och obegränsad limit. *Excess-of-loss-avtal* fungerar så att återförsäkringsbolaget för varje skada betalar den del av skadekostnaden som överstiger c_i. Det ersättningsbelopp som är av intresse är den delen som återförsäkringsbolaget betalar. På förslag av uppdragsgivaren kommer vi bara att räkna med självbehåll i intervallet 1.00 Mkr till 5.00 Mkr. Datamaterialet från Trygg-Hansa innehåller bara de skador som är större än eller lika med $c = 1.00$ Mkr.

Låt Z_i beteckna hela kostnaden för en skada. Dessa skadebelopp betraktas som utfall av oberoende och likafördelade stokastiska variabler. Den del som överskrider självbehållet c kallar vi storskadekostnaden och betecknar X_i.Om ersättningsbelopp uppgår till Z_i kronor betalar återförsäkringsbolaget beloppet

$$X_i = \begin{Bmatrix} 0 & om \quad Z_i <= C_i \\ Z_i - C_i & om \quad Z_i > C_i \end{Bmatrix}$$

I praktiken finns alltid en övre gräns (limit) för återförsäkringsskyddet. Denna sätts normalt så högt att risken för ”spillover” (skadebeloppen som överskrider limiten) är väldigt liten. För enkelhets skull räknar vi i detta arbete utan någon limit. Återförsäkringsskyddet är oftast uppdelat på flera återförsäkrare (som täcker olika s.k. ”layers”). Vi kan bortse från detta eftersom vi betraktar problemet ur Trygg-Hansas synvinkel.

Nästa steg är att försöka finna en fördelningsfunktion som passar bra till detta modifierade datamaterial. Fördelningen som anpassats gäller i intervallet $[0, \infty)$.

Vi ska jämföra två olika återförsäkringskontrakt vad gäller deras varians vid givet väntevärde. Den första typen av återförsäkringskontrakt kallas för *excess of loss med aggregerad* återförsäkring medan den andra är vanlig *excess of loss* återförsäkring. *Excess of loss med* *aggregerad* återförsäkring innebär att man betraktar det totala skadebeloppet, T, över excesspunkten. Återförsäkraren betalar bara den delen, S, som överskrider aggregatet a, dvs.

$$S = \begin{cases} 0 & om \quad T < a \\ T - a & annars \end{cases}$$

För att kunna jämföra kontrakten kommer vi att välja a, givet excesspunkterna C_1 och C_2, så att deras väntevärde blir detsamma.

Vi förklarar detta i ett exempel: Först bestäms ett självbehåll C_1 för kontrakt ett. Därefter beräknas väntevärdet för den totala skadekostnaden. Sedan görs samma sak för kontrakt två fast med ett nytt självbehåll C_2, som måste vara är större än C_1. Eftersom deras väntevärde ska vara samma så måste man med hjälp av numerisk lösning bestämma ett aggregat så att detta krav blir uppfyllt.

Om vi exempelvis har ett datamaterial som ser ut på följande sätt

År	Skadekostnad före återförsäkring
1	3 000 000
2	6 000 000
2	2 000 000
2	2 500 000
3	1 000 000
3	2 500 000
4	2 000 000
4	4 000 000
5	1 500 000
5	3 000 000
5	2 000 000

Trygg-Hansa bestämmer att de ska stå för det första $C_1 = 1\ 000\ 000$ resp. $C_2 = 1\ 500\ 000$ kronor i varje skada, och tecknar skydd med excesspunkt C_1 resp. C_2.

År	Skadekostnad som återförsäkringsbolag står för med C_1	År	Skadekostnad som återförsäkringsbolag står för med C_2
1	2 000 000	1	1 500 000
2	5 000 000	2	4 500 000
2	1 000 000	2	500 000
2	1 500 000	2	1 000 000
3	1 500 000	3	1 000 000
4	1 000 000	4	500 000
4	3 000 000	4	2 500 000
5	500 000		
5	2 000 000	5	1 500 000
5	1 000 000	5	500 000

Väntevärde för den totala skadekostnaden i detta fall är:

Väntevärde1: = 3 700 000 Väntevärde2: = 2 700 000

Ett aggregat bestäms så att Väntevärde1 = Väntevärde2.

I detta fall blir aggregatet =1 000 000. Efter detta får man gå tillbaka till kontrakt ett och införa det aggregerade värde som vi har fått.

I så fall betalar återförsäkringsbolaget följande:

År	Skadekostnad som återförsäkringsbolag står för med C_1+aggregat
1	1 000 000
2	4 000 000
2	1 000 000
2	1 500 000
3	500 000
4	3 000 000
5	1 500 000
5	1 000 000

Därefter beräknar man variansen för båda kontrakten. Varianspåverkan kommer att studeras vid dessa olika typer av kontrakt.

3 Statistisk modell

3.1 Parameterskattningar

Parametrarna till de olika fördelningarna skattas dels med maximum likelihood metoden (ML-metoden) och dels med momentmetoden. ML-metoden skattar parametrarna genom att likelihoodfunktionerna maximeras. Momentmetoden skattar parametrarna så att de första momenten för fördelningen överensstämmer med de skattade momenten för försäkringstypen.

3.2 Val av fördelning för antalet skador

En Poissonfördelning uppträder då händelser inträffar slumpmässigt i tiden, d.v.s. när händelserna är oberoende av varandra och kan inträffa när som helst. Dessutom förutsätts det att händelserna inträffar med en konstant frekvens så att λ händelser inträffar i genomsnitt per tidsenhet.

Poissonfördelningen är allmänt använd inom försäkringsbranschen, inte minst inom återförsäkring. Vi antar att antalet skador, N, under en viss tidsperiod är Poissonfördelat. Skadefrekvensen, l, är medelvärdet av antalet skador som inträffat per år under 1995-2001 och som är över C_1=1.00 Mkr.

3.3 Val av fördelning för skadekostnad

Målet är att hitta en modell som passar data och som kan användas för att beräkna hur stor skadekostnad återförsäkringsbolaget förväntas ha under ett år. Den totala kostnaden är summan av dessa skadekostnader.

Karakteristiskt för en fördelning för skadeförsäkringsdata inom återförsäkring är att den är skev och tjocksvansad. De fördelningar som är lämpade att använda för detta ändamål och som används mest för att modellera storleken på skadebeloppen inom återförsäkringen är Pareto och lognormalfördelning. Man kan säga att Pareto och lognormalfördelning är

”farliga” fördelningar, där mycket stora skador är möjliga. Anpassningen av dessa två fördelningar till den empiriska fördelningen har jämförts för varje skadegren.

Not: Varför anpassa en parametrisk fördelning och inte använda den empiriska fördelningen? Nackdelen med den empiriska fördelningen är att sannolikheten, för att få en observation som är större än den största skadan, som har inträffat hittills, är lika med noll. Detta är väldigt olämpligt eftersom man vet att större skador än den hittills största mycket väl kan inträffa och inom återförsäkringen är man speciellt intresserad av just sannolikheten för sådana skador. Därför kommer vi att använda den empiriska fördelningen bara som hjälpmedel för att i den här rapporten kunna avgöra vilken parametrisk fördelning som passar bäst.

3.4 Kriterier vid val av skadekostnadsfördelning

Målsättningen är att hitta en fördelning som stämmer så bra överens med den empiriska fördelningen som möjligt. Ett bra sätt för att utesluta den fördelning som inte passar till data, är att räkna ut den empiriska och parametriska fördelningen för observationerna och plotta dem i samma diagram. Ju närmare de ligger varandra desto bättre passar den parametriska fördelningen med givna data.

För att få ytterligare ett mått på hur väl en fördelning passar den empiriska fördelningen kommer vi att studera Chi-2-statistikan (Lindgren [4]) samt Kolmogorov-Smirnovs statistika (Hjorth [5], sid.132).

Testet görs här bara för att göra en jämförelse mellan olika modeller och för att få en bild på hur väl en fördelning passar och ej för att förkasta en eventuell modell. Nackdelen med chi-2 testet är att indelningen av data i k intervall blir godtycklig och på det sättet är resultatet inte entydigt. Det är därför vi även betraktar ett avståndsmått som man definierar direkt på den empiriska och den parametriska fördelningsfunktionen.

Kolmogorov-Smirnovs test går ut på att jämföra den empiriska fördelningsfunktionen $G(x)$ med motsvarande parametrisk fördelningsfunktion $F(x)$. Dess test-statistika definieras enligt

$$D^{+} = \sup_{x}\left(G_n(x) - F(x)\right)$$

och

$$D^{-} = \sup_{x}\left(F(x) - G_n(x)\right)$$

$$D = \max(D^{+}, D^{-})$$

4 Simulering

4.1 Slumptalsgenerering

Simulering ligger ofta till grund för prognoser. Med simulering menas att man beskriver ett system eller förlopp med en matematisk modell som vanligen realiseras i ett dataprogram. Eftersom en simuleringsmodell endast är en mer eller mindre bra approximation av verkligheten kan simulering aldrig helt ersätta verkliga experiment. I detta fall kommer vi att simulera dels antalet skador och dels skadekostnader för dessa. Skadestorleken för dessa skador simuleras enskilt från den skattade skadefördelningen.

Vid all stokastisk simulering utgår man från slumptal. Slumptal som är oberoende och likformigt fördelade över (0,1) får vi med hjälp av SAS inbyggda slumptalsgenerator. För att erhålla andra fördelningar, använder vi den så kallade ”Invers fördelningsfunktion” metoden. Denna metod bygger på iakttagelsen att vilken kontinuerlig fördelning X än har så blir fördelningsfunktionen värde i den dragna punkten alltid U(0,1), och kan därför genereras direkt genom att dra ett slumptal U.

Slumptalet X erhålles som det värde där *F (X) =U*
U är likformigt fördelad på (0,1)
P (U£ u) = u, 0£ u£ 1

Det gäller vidare att
$$P(X \le x) = P(U \le F(x)) = F(x)$$
Om F är inverterbar kan vi alltså sätta
$$X = F^{-1}(U)$$
1) Ett Paretofördelat slumptal *X* kan erhållas som
$$X = F^{-1}(U) = a * \left[\frac{1}{(1-U)^{1/g}} - 1 \right]$$
(se Hogg och Klugman [7])

2) Lognormal

SAS har en algoritm för att generera normalfördelade slumptal, nämligen funktionen *RANNOR* som ger *N (0,1)*-fördelade slumptal. Om R är ett sådant slumptal så blir

$Y = m + s * R$ ett $N\left(m, s^{2}\right)$- fördelat slumptal.

Ett slumptal *X* som är lognormal fördelat med parametrarna μ och s^{2} erhålles nu som

$$X = e^{m+s*R} = e^{Y}$$

3) Ett slumptal som är Poissonfördelat med parameter l får vi hjälp av funktion *RANPOI*.

Ett års skador har simulerats 10 000 gånger

4.2 Fördelning för den totala kostnaden

Modellen som vi kommer att använda här är den kollektiva modellen, i vilken man bortser från vilka försäkringstagare som drabbas av skador och endast betraktar följden av skador som drabbar skadegrenen.

S är summan av ett stokastiskt antal stokastiska variabler och har en sammansatt fördelning. Som vi tidigare har sagt så betecknar *N* det totala antalet skador. *N* är Poissonfördelat med parameter l. X_i betecknar ersättningsbeloppet över C_1. I så fall kommer den totala ersättningen för skadegrenen att bli

$$S = \sum_{i=1}^{N} X_i$$

Vidare antar vi att alla X_i är oberoende, likafördelade stokastiska variabler med fördelningsfunktion F och att dessa är oberoende av den stokastiska variabel *N*.

S är det totala beloppet som återförsäkringsbolaget får betala ut för skadegrenen och den har en sammansatt Poissonfördelning. Vi använder beteckningen *S Î SaPo (l ,F).*

5 Resultat

5.1 Civilförsäkring

Totalt fanns det 154 observationer som ingick i analysen, vilka har inträffat från 1995 och fram till 31 december, 2001. Den totala kostnaden under denna period var på *122 017 387* kr.

Tabell 1 Resultat av parameterskattningar

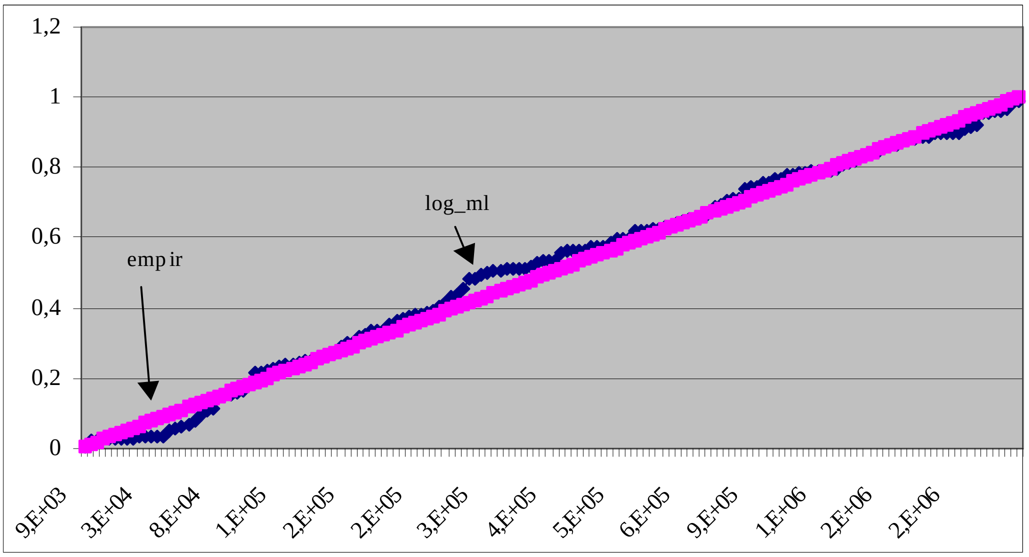
ML-metodens skattningar via ML-ekvationerna

Lognormal		
	my	*12.769*
	sigma	*1.387*
Pareto		
	alfa	*3 726 668*
	gamma	*6.062*

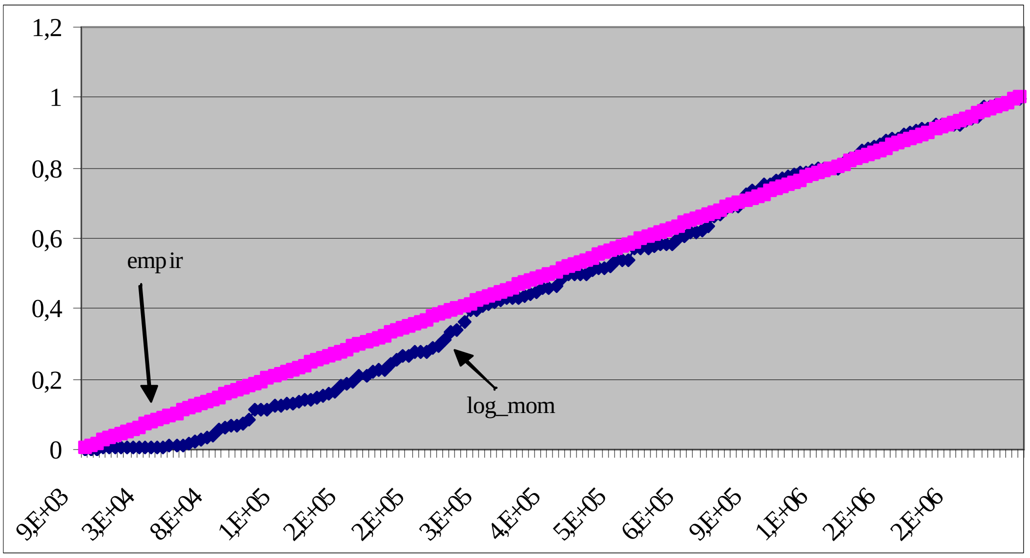
Momentmetodens skattningar

Lognormal		
	my	*12.999*
	sigma	*1.081*
Pareto		
	alfa	*2 096 168*
	gamma	*3.646*

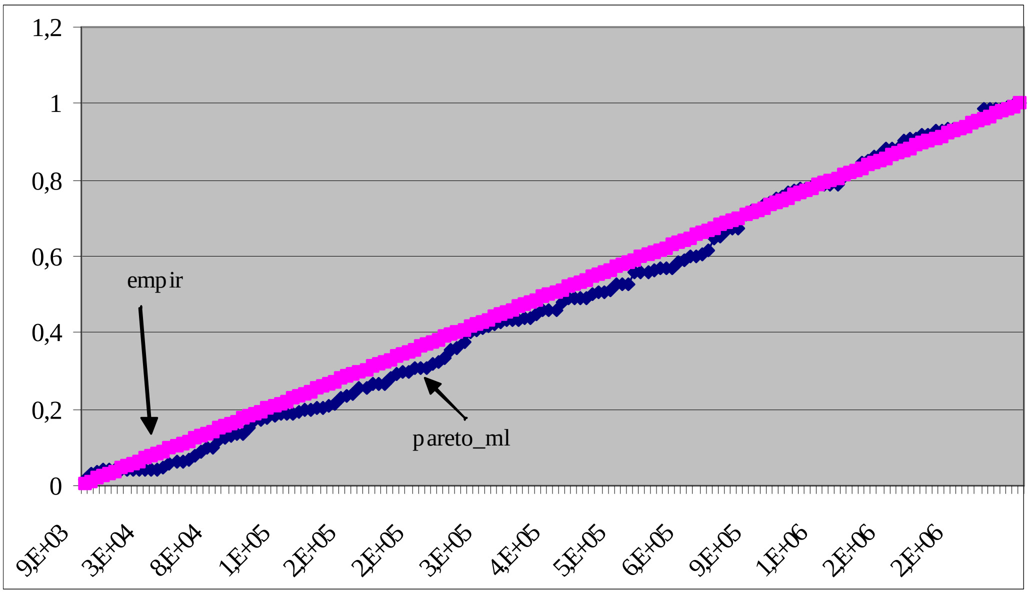
5.1.1 Jämförelse mellan fördelningar



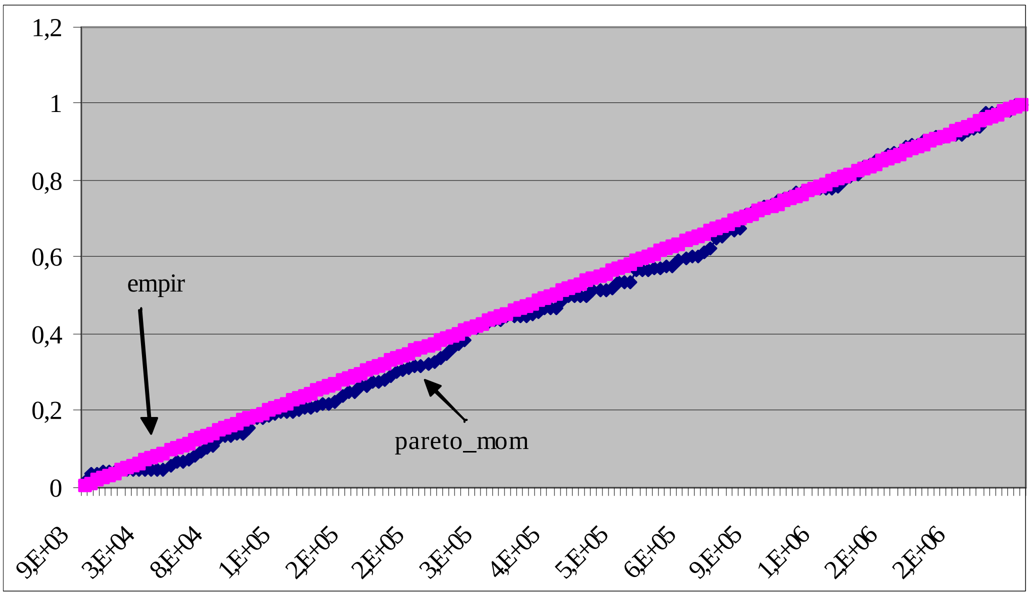
Figur 1 visar empirisk och lognormal fördelning parametrar skattade med ML- metoden.



Figur 2 visar empirisk och lognormalfördelning med momentmetoden



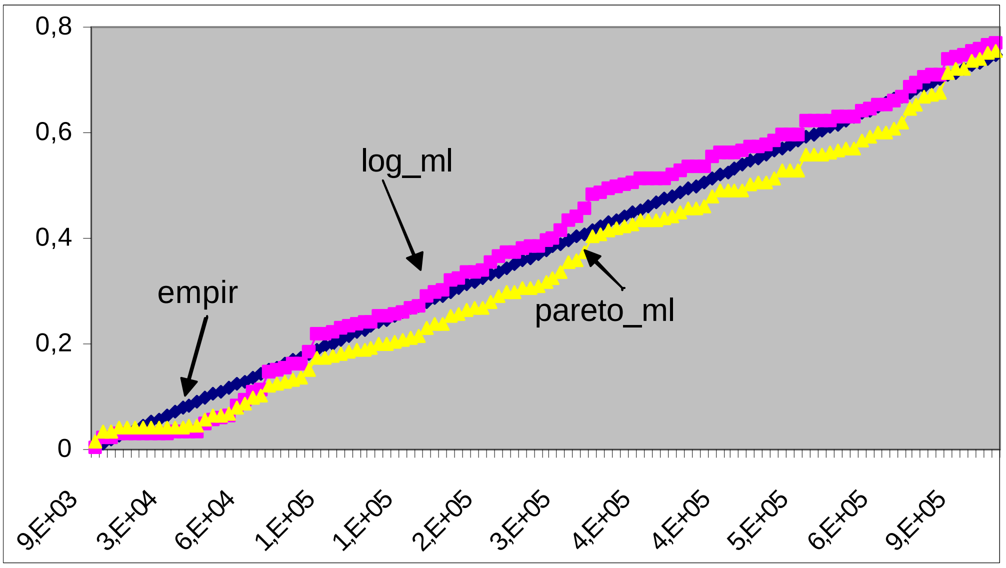
Figur 3 visar empirisk fördelning och Paretofördelning med ML-metoden.



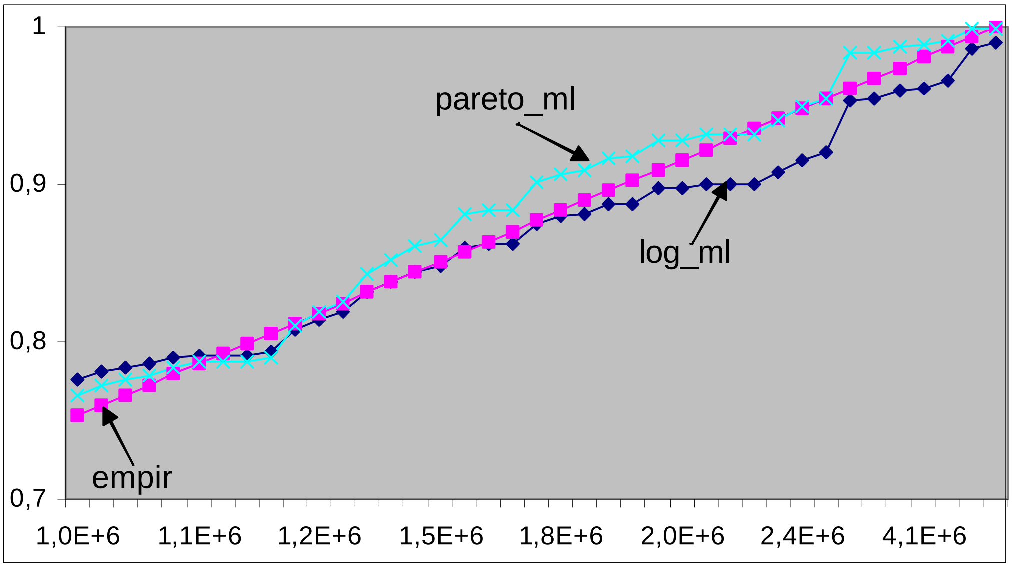
Figur 4 visar empirisk fördelning och Paretofördelning med momentmetoden.

Val av fördelning

Figurer 1 och 3 visar att det inte är helt självklart vilken fördelning som har den bästa anpassningen. Speciellt när det gäller fördelningen för de skador vars skadekostnad överstiger två miljoner var det lite svårt att bedöma vilken fördelning som passar bäst eftersom Pareto och lognormalfördelning följer varandra och den empiriska rör sig runt dessa två. Däremot ser man direkt att lognormal med momentmetoden gav den sämsta anpassningen.
För att kunna bedöma med hjälp av grafer vilken fördelning som passar bäst har vi indelat skadekostnaden i följande intervall: 0 till 1.00 Mkr, 1.00 Mkr till 10.00 Mkr.



Figur 5 visar den empiriska samt lognormal och Pareto fördelningen med parametrar skattade med ML-metoden för skadekostnad < 1.00 Mkr.



Figur 6 visar den empiriska samt lognormal och Pareto fördelningen med parametrar skattade med ML-metoden för skadekostnad > 1.00 Mkr.

5.1.2 Resultat av chi-2 och Kolmogorov-Smirnovs teststatistikan

	Q_statistika	*K_S_statistika*
Lognormal ML-metoden	*28.629*	*0.039*
Lognormal momentmetoden	*78.609*	*0.101*
Pareto ML-metoden	*26.586*	*0.057*
Pareto momentmetoden	*23.383*	*0.082*

Slutsats

Valet av fördelningen utifrån ovanstående grafer är lognormalfördelning med parametrar skattade med ML-metoden för skador vars skadekostnad ligger i intervallet 0 till 1.00 Mkr. Det är den fördelning som stämmer bäst med den empiriska fördelningen. När det gäller fördelningen för de skador vars skadekostnad överstiger 1.00 Mkr var det lite svårt att bedöma vilken fördelning passar bäst eftersom både Pareto och lognormalfördelning följer varandra och den empiriska rör sig runt dessa två. Valet blir lognormalfördelning med ML metoden då den fördelningen ligger närmare i de flesta fall. Lognormalfördelning med parametrarna skattade med momentmetoden ger sämsta anpassning.

5.1.3 Kontroll av simuleringens noggrannhet

För att kontrollera simuleringens noggrannhet jämför vi dess väntevärde och standardavvikelse för *s* med de exakta analytiska värdena enligt A5.

	Log_ML		Log_mom	
	Approx.	Simulerade	Approx.	Simulerade
Skevhet	0.209		0.246	
Väntevärde	20 209 310	20 272 377	17 431 055	17 537 257
Standardavvikelse	11 275 943	11 654 924	6 663 877	6 604 996

	Pareto_ML		Pareto_mom	
	Approx.	Simulerade	Approx.	Simulerade
Skevhet	0.243		0.246	
Väntevärde	16 196 503	16 117 325	17 431 092	17 554 914
Standardavvikelse	5 451 500	5 414 468	6 663 896	6 675 363

Tabell 3 jämförelse av den totala skadekostnaden

Överensstämmelsen är riktigt bra.

5.1.4 Jämförelse mellan standardavvikelse i excess of loss med aggregerat och excess loss återförsäkring

Där:

C_1 är *ett självbehåll* för excess of loss med aggregerad återförsäkring

C_2 är *ett självbehåll* för excess of loss återförsäkring

m_2 är *väntevärde* för den totala skadekostnaden för excess of loss återförsäkring

s_2 är *standardavvikelse* för den totala skadekostnaden för excess of loss återförsäkring

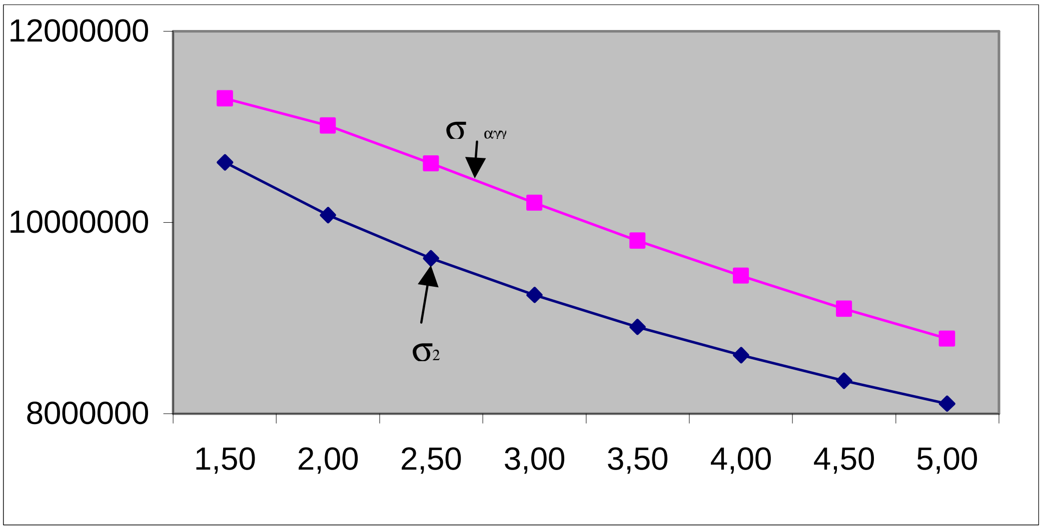
m_agg är *väntevärde* för den totala skadekostnaden för excess of loss med aggregerat självbehåll

s_agg är *standardavvikelse* för den totala skadekostnaden för excess of loss med aggregerat självbehåll

Lognormalfördelning med parametrar skattade med ML-metoden

C_1	C_2	μ_2	σ_2	μ_agg	σ_agg	aggregat	$\sigma\_agg/\sigma 2$
1,00	1,50	13147154	10629585	13147154	11295971	7040740	1,063
	2,00	9862080	10079661	9862080	11011912	10610586	1,092
	2,50	7860871	9626717	7860871	10615645	13102394	1,103
	3,00	6497461	9241715	6497461	10204322	15066547	1,104
	3,50	5501730	8908664	5501730	9810933	16721707	1,101
	4,00	4742249	8611158	4742249	9442168	18176604	1,097
	4,50	4140702	8343946	4140702	9100126	19494044	1,091
	5,00	3659534	8104661	3659534	8786498	20695948	1,084

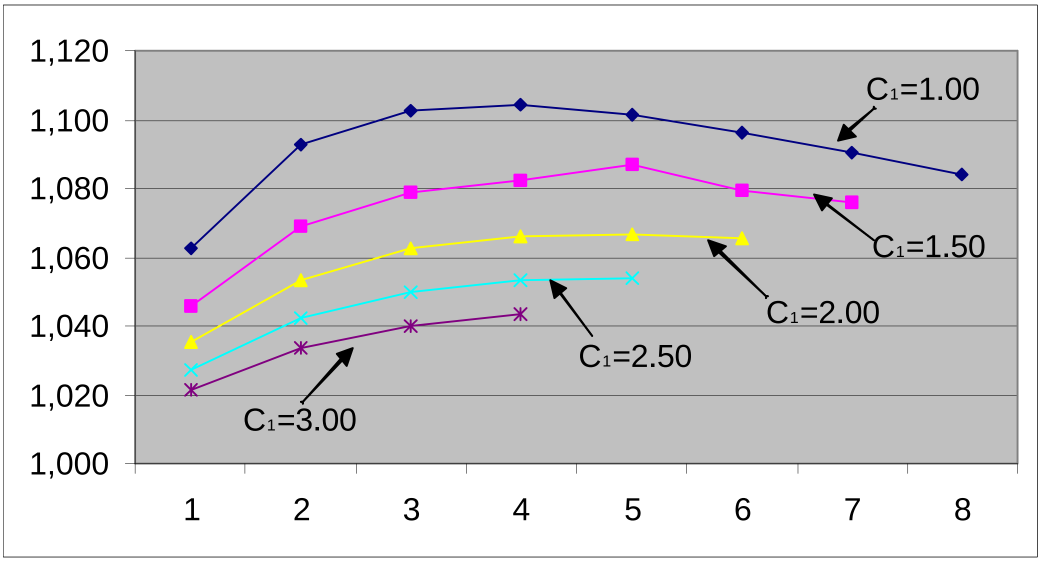
Tabell 4 Självbehåll, väntevärde, standardavvikelse samt kvoten mellan s_agg och s_2 för de två återförsäkringskontrarakter.



Figur 7 visar standaravvikelsen för de två återförsäkringskontrakt då $C_1 = 1.00$ Mkr.

C_1	C_2	μ_2	σ_2	μ_agg	σ_agg	aggregat	σ_agg/σ2
1,50	2,00	9862080	10079661	9862080	10539135	3374474	1,046
	2,50	7860871	9626717	7860871	10291001	5673294	1,069
	3,00	6497461	9241715	6497461	9970642	7492902	1,079
	3,50	5501730	8908664	5501730	9639785	9032498	1,082
	4,00	4742249	8611158	4742249	9361088	10393352	1,087
	4,50	4140702	8343946	4140702	9005611	11637153	1,079
	5,00	3659534	8104661	3659534	8718730	12771642	1,076
2,00	2,50	7860871	9626717	7860871	9965903	2135018	1,035
	3,00	6497461	9241715	6497461	9737008	3816660	1,054
	3,50	5501730	8908664	5501730	9464250	5251908	1,062
	4,00	4742249	8611158	4742249	9181918	6527653	1,066
	4,50	4140702	8343946	4140702	8900140	7704594	1,067
	5,00	3659534	8104661	3659534	8635840	8781336	1,066
2,50	3,00	6497461	9241715	6497461	9495600	1541981	1,027
	3,50	5501730	8908664	5501730	9284624	2870891	1,042
	4,00	4742249	8611158	4742249	9041504	4067759	1,050
	4,50	4140702	8343946	4140702	8789405	5177096	1,053
	5,00	3659534	8104661	3659534	8543819	6205397	1,054
3,00	3,50	5501730	8908664	5501730	9098362	1217946	1,021
	4,00	4742249	8611158	4742249	8901857	2325123	1,034
	4,50	4140702	8343946	4140702	8679490	3368668	1,040
	5,00	3659534	8104661	3659534	8455679	4342952	1,043
3,50	4,00	4742249	8611158	4742249	8760381	1017537	1,017
	4,50	4140702	8343946	4140702	8572724	1987197	1,027
	5,00	3659534	8104661	3659534	8370647	2908821	1,033
4,00	4,50	4140702	8343946	4140702	8459853	894245	1,014
	5,00	3659534	8104661	3659534	8282422	1757796	1,022
4,50	5,00	3659534	8104661	3659534	8192201	804388	1,011

Tabell 5 Självbehåll, väntevärde, standardavvikelse samt kvoten mellan aggregerad och vanlig standardavvikelse för de två återförsäkringskontrakt med olika excesspunkter



Figur 8 visar kvoten mellan standaravvikelsen för de 5 först excesspunkterna

Slutsats

Standardavvikelsen för den aggregerade återförsäkringskontrakt är större än standardavvikelsen för den vanliga återförsäkringskontrakt. Om vi ökar C_1 så avtar skillnaden mellan dem. Om man däremot ökar C_2 så är denna skillnad nästan konstant.

5.2 Företagsförsäkring

Data

Från 1995 och fram till 31 december inträffade 198 observationer som ingår i analysen. Deras kostnad uppgick till *483 884 339* kr.

Tabell 6 Resultat av parameterskattningar

ML-metodens skattningar via ML-ekvationerna

Lognormal

my	*13.673*
sigma	*1.714*

Pareto

alfa	*18 395 000*
gamma	*9.144*

Momentmetodens skattningar

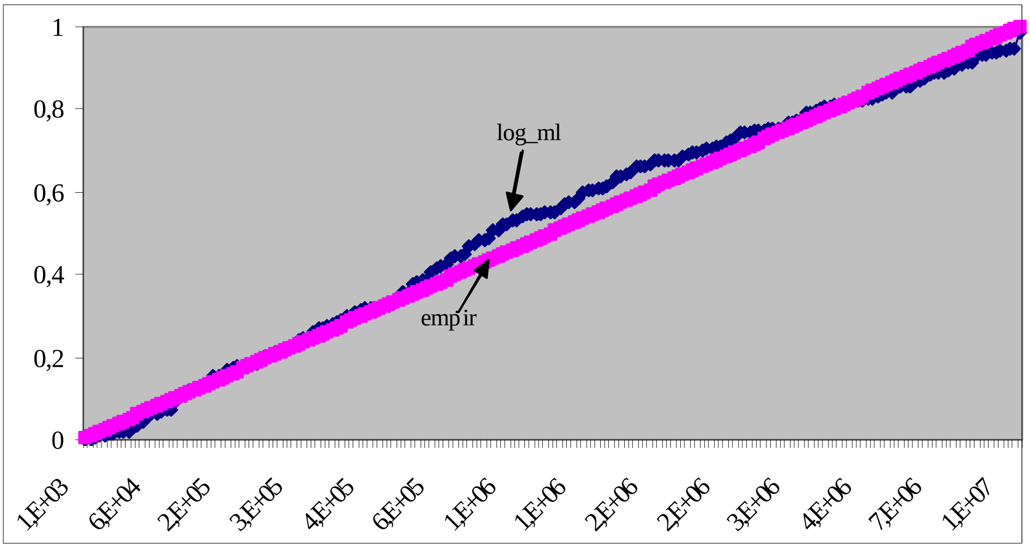
Lognormal

my	*14.115*
sigma	*1.559*

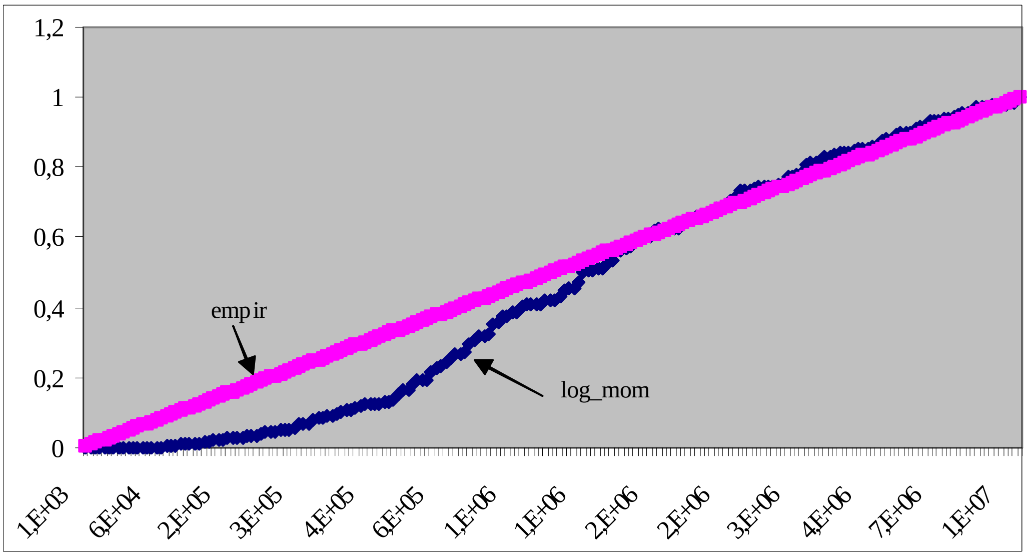
Pareto

alfa	*2 606 670*
gamma	*2.067*

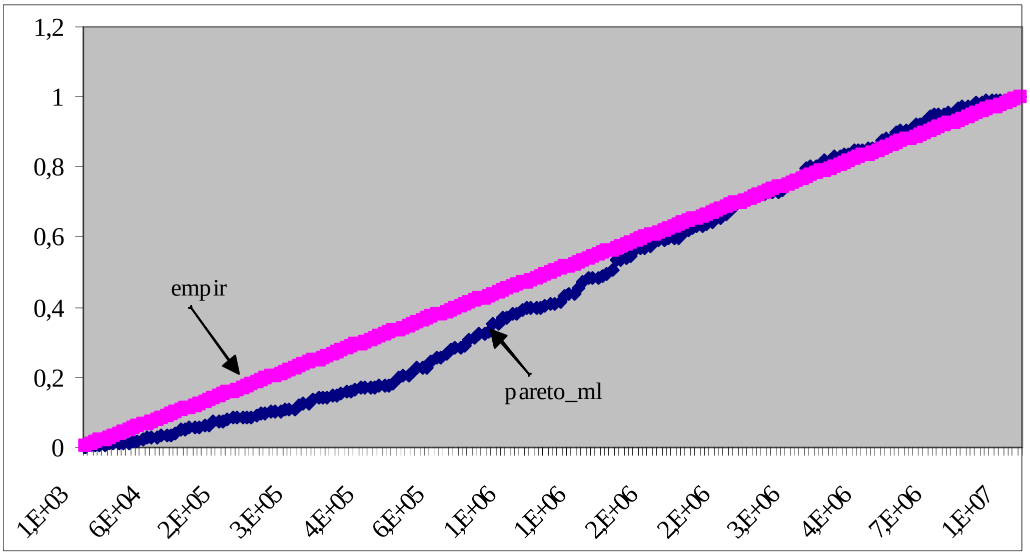
5.2.1 Jämförelse mellan fördelningar



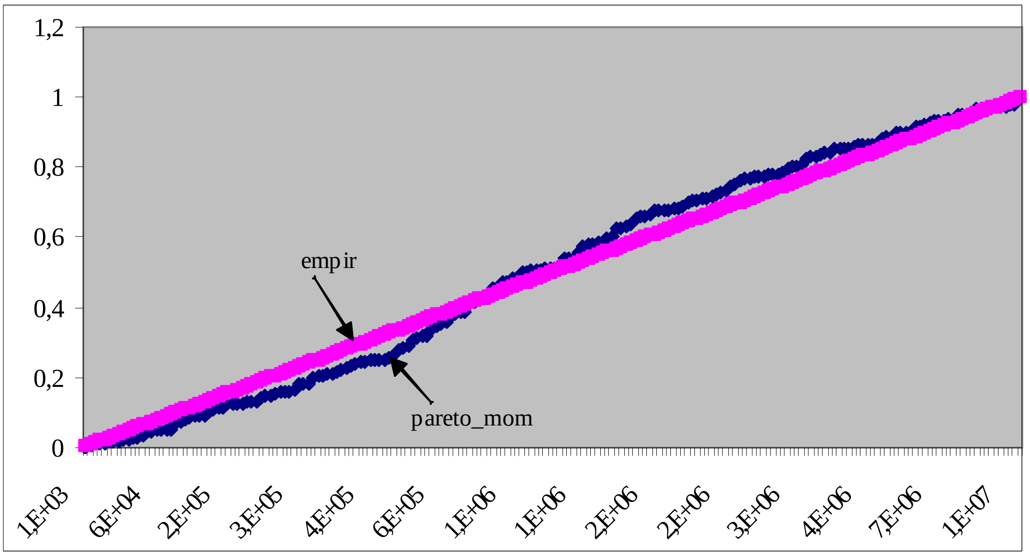
Figur 9 visar empirisk och lognormal fördelning med parametrar skattade med ML- metoden.



Figur 10 visar empirisk och lognormalfördelning med momentmetoden.



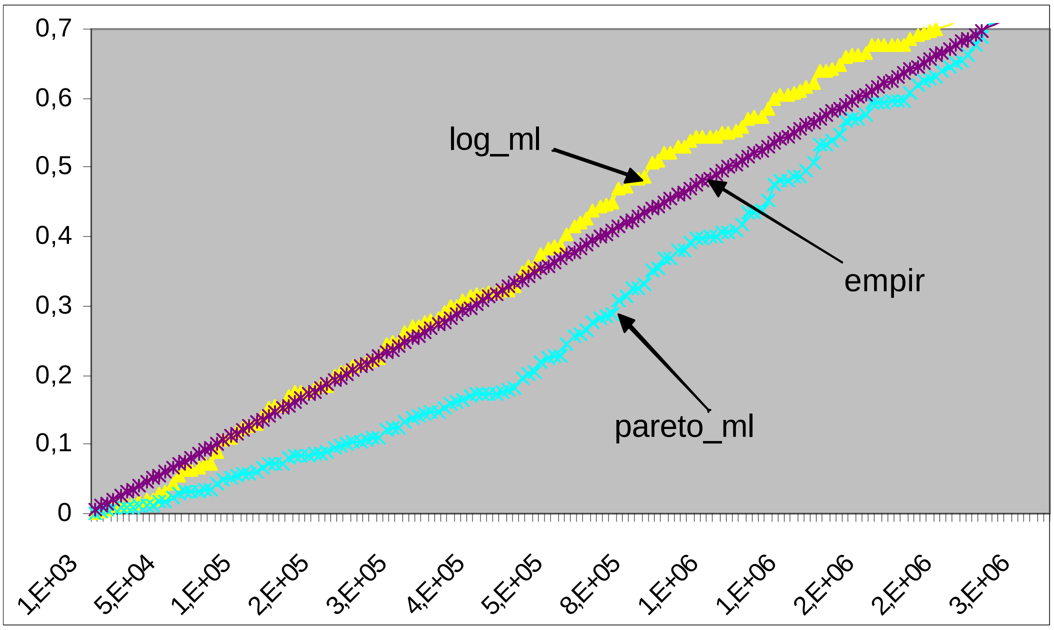
Figur 11 visar empirisk fördelning och paretofördelning med ML-metoden.



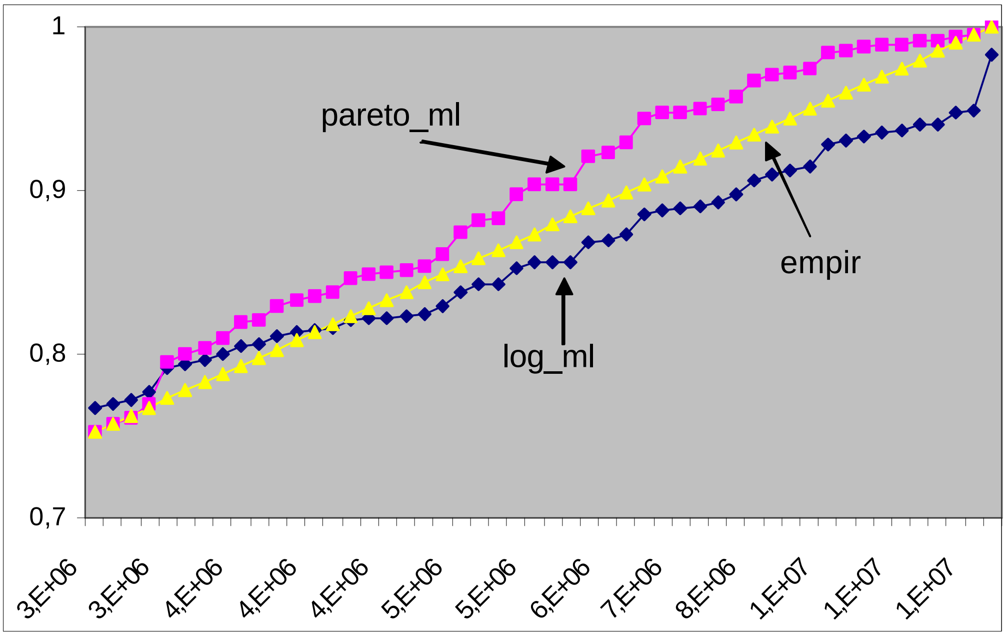
Figur 12 visar empirisk fördelning och Paretofördelning med momentmetoden.

Slutsats

Pareto och lognormalfördelning med parametrar skattade med ML-metoden passar ungefär lika bra till data. För att se om det finns någon skillnad i olika intervall har ersättningskostnaden indelats i följande intervall: 0 till 3.00 Mkr och 3.00 Mkr till 35.00 Mkr.



Figur 13 visar skadekostnad i intervallet 0 till 3.00 Mkr.



Figur 14 visar skadekostnad i intervallet 3.00 Mkr till 35 Mkr.

5.2.2 Resultat av chi-2 och Kolmogorov-Smirnovs teststatistikan

	Q_statistika	*K_S_statistika*
Lognormal ML-metoden	*31.514*	*0.067*
Lognormal momentmetoden	*923.110*	*0.197*
Pareto ml-metoden	*61.029*	*0.151*
Pareto momentmetoden	*34.347*	*0.074*

Slutsats

Valet av fördelningsfunktionen utifrån ovanstående resultaten är lognormalfördelning med parametrar skattade med ML-metoden för skador vars skadekostnad ligger i intervallet 0 till 3.00 Mkr. När det gäller de skador vars skadekostnader överstiger tre miljoner så är det Pareto fördelningen som passar bäst. Den fördelningen som passar sämst är lognormal med momentmetoden. Det som var lite förvånande här är att Pareto med momentmetoden har också väldigt bra anpassning för skadekostnader som är mindre än 3.00 Mkr. Vi väljer lognormalfördelning med ML-skattningar.

5.2.3 Kontroll av simuleringens noggrannhet

	Log_ML		Log_mom	
	Approx.	Simulerade	Approx.	Simulerade
Skevhet	0.123		0.217	
Väntevärde	106 602 209	107 171 965	69 126 334	68 838 329
Standardavvikelse	87 118 592	99 289 061	23 542 324	25 252 156

	Pareto_ML		Pareto_mom	
	Approx.	Simulerade	Approx.	Simulerade
Skevhet	0.211		Går ej att räkna ut	
Väntevärde	63 891 023	63 945 593	69 128 416	63 829 040
Standardavvikelse	18 139 285	17 992 567	73 561 691	18 101 103

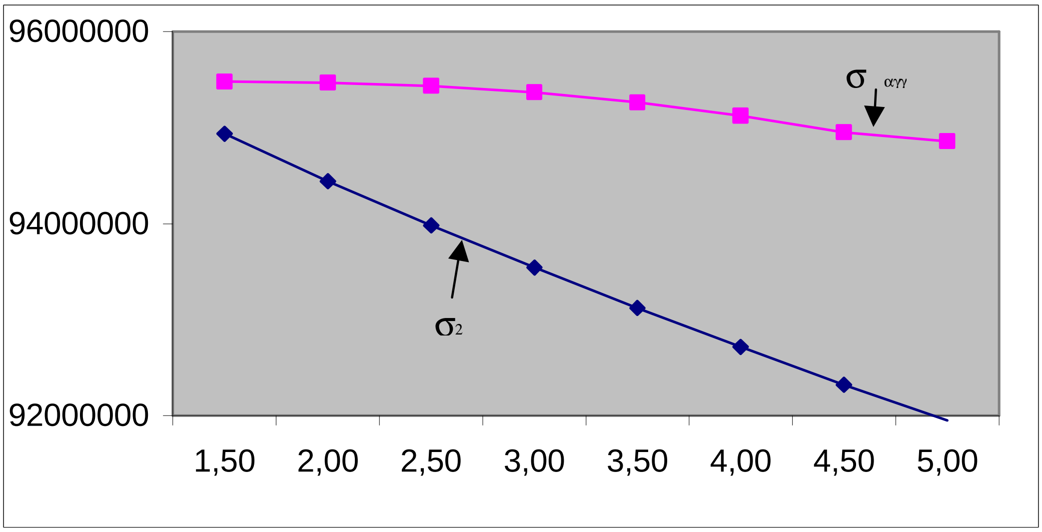
Tabell 7 jämförelse av den totala skadekostnaden

5.2.4 Jämförelse mellan standardavvikelse i excess of loss med aggregerat och excess of loss återförsäkring

Lognormalfördelning med parametrar skattade med ML-metoden

C_1	C_2	μ_2	σ_2	μ_agg	σ_agg	aggregat	σ_agg/σ2
1,00	1,50	97510079	94935882	97510079	95478585	11082999	1,006
	2,00	89896806	94443221	89896806	95468984	18706198	1,011
	2,50	83390733	93979568	83390733	95433653	24651309	1,015
	3,00	79147427	93539834	79147427	95365163	29574987	1,020
	3,50	75045078	93122408	75045078	95261522	33805781	1,023
	4,00	71497440	92716188	71497440	95121377	37535807	1,026
	4,50	68380015	92321579	68380015	94951043	40884946	1,028
	5,00	65597595	91948041	65597595	94857607	43941480	1,032

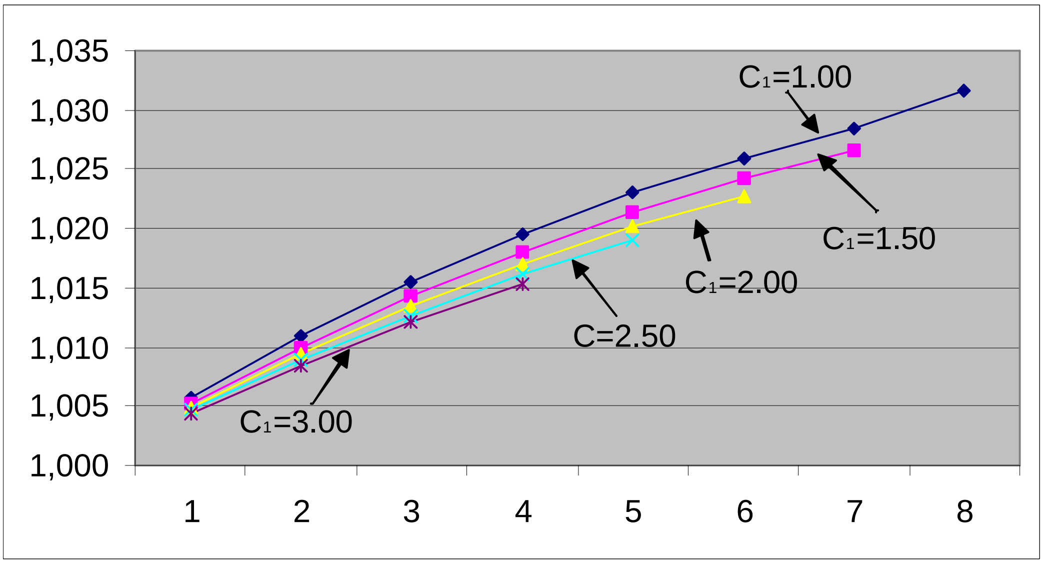
Tabell 8 Självbehåll, väntevärde, standardavvikelse samt kvoten mellan s_agg och s_2 för de två återförsäkringskontrarakter.



Figur 15 visar standaravvikelsen för de två återförsäkringskontrakt då $C_1 = 1.00$ Mkr.

C_1	C_2	μ_2	σ_2	μ_agg	σ_agg	aggregat	σ_agg/σ2
1,50	2,00	89896806	94443221	89896806	94934273	7614946	1,005
	2,50	83390733	93979568	83390733	94919003	13537846	1,010
	3,00	79147427	93539834	79147427	94877251	18429966	1,014
	3,50	75045078	93122408	75045078	94803860	22622837	1,018
	4,00	71497440	92716188	71497440	94695626	26310321	1,021
	4,50	68380015	92321579	68380015	94554616	29619169	1,024
	5,00	65597595	91948041	65597595	94389794	32634255	1,027
2,00	2,50	83390733	93979568	83390733	94439726	5909939	1,005
	3,00	79147427	93539834	79147427	94419265	10777079	1,009
	3,50	75045078	93122408	75045078	94371245	14938427	1,013
	4,00	71497440	92716188	71497440	94292665	18587476	1,017
	4,50	68380015	92321579	68380015	94181137	21855412	1,020
	5,00	65597595	91948041	65597595	94042023	24833471	1,023
2,50	3,00	79147427	93539834	79147427	93973554	4850325	1,005
	3,50	75045078	93122408	75045078	93946809	8985436	1,009
	4,00	71497440	92716188	71497440	93893412	12601708	1,013
	4,50	68380015	92321579	68380015	93810045	15831217	1,016
	5,00	65597595	91948041	65597595	93697985	18770688	1,019
3,00	3,50	75045078	93122408	75045078	93530930	4113259	1,004
	4,00	71497440	92716188	71497440	93498451	7702509	1,008
	4,50	68380015	92321579	68380015	93440401	10897684	1,012
	5,00	65597595	91948041	65597595	93353521	13801432	1,015
3,50	4,00	71497440	92716188	71497440	93109387	3564312	1,004
	4,50	68380015	92321579	68380015	93072170	6731413	1,008
	5,00	65597595	91948041	65597595	93009971	9600391	1,012
4,00	4,50	68380015	92321579	68380015	92698738	3140687	1,004
	5,00	65597595	91948041	65597595	92657765	5979930	1,008
4,50	5,00	65597595	91948041	65597595	92300838	2811126	1,004

Tabell 9 Självbehåll, väntevärde, standardavvikelse samt kvoten mellan aggregerad och vanlig standardavvikelse för de två återförsäkringskontrakt med olika excesspunkter



Figur 16 visar kvoten mellan standaravvikelsen för de 5 först excesspunkterna

Slutsats

Standardavvikelsen för det aggregerade återförsäkringskontraktet är lite större än standardavvikelsen för vanlig excess of loss. Här är medelvärdet så stort att skillnaden mellan metoderna blir lite. Den ökar med ökande C_2.

5.3 Kaskoförsäkring

Data

Den totala antalet skador som ingår i analysen var 27 och deras kostnad uppgick till *12 112 126* kr.

Tabell 10 Resultat av parameterskattningar

Ml-metodens skattningar via Ml-ekvationerna

Lognormal

my	*12.635*
sigma	*0.942*

Pareto

alfa	*1 006 650*
gamma	*2.696*

Momentmetodens skattningar

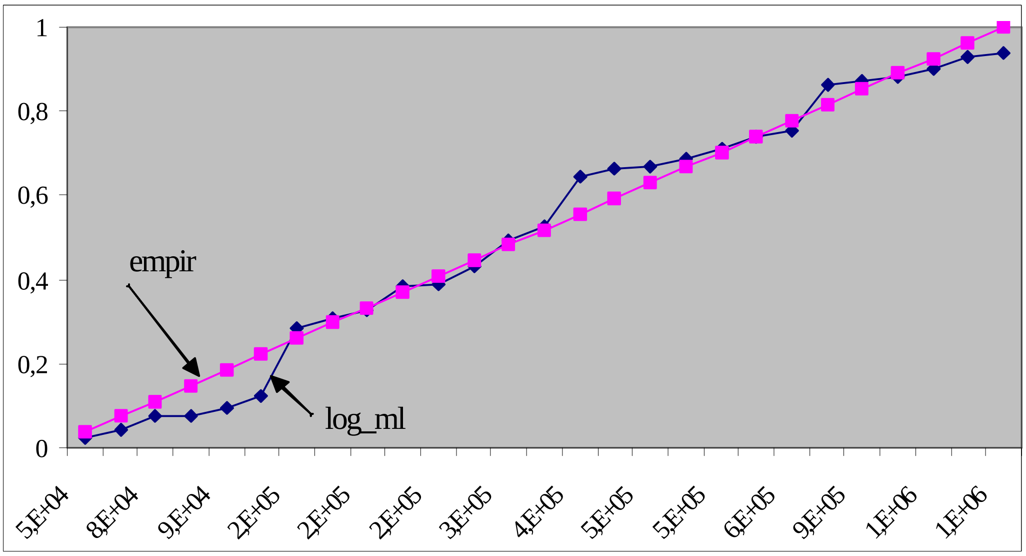
Lognormal

my	*12.768*
sigma	*0.702*

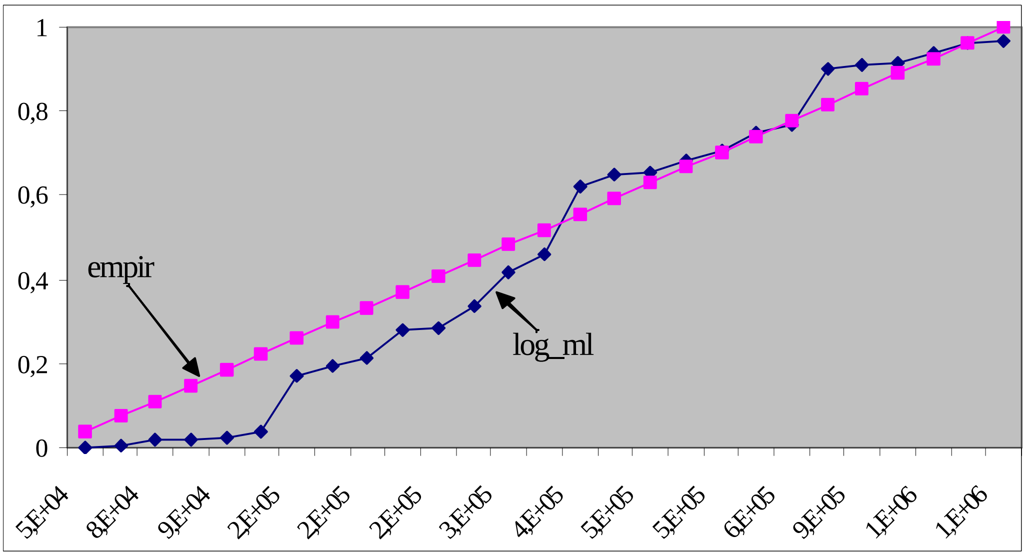
Pareto

alfa	*322 486*
gamma	*2.400*

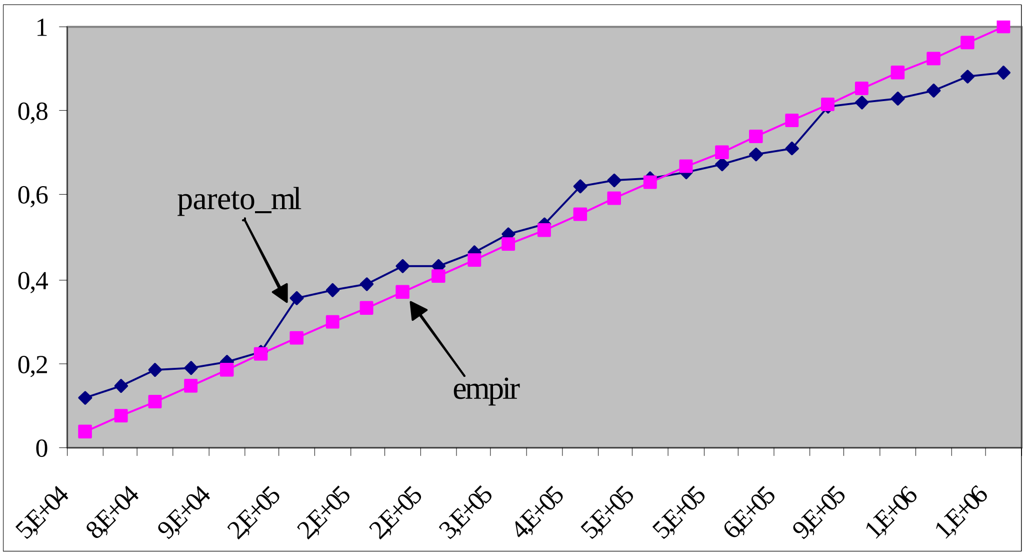
5.3.1 Jämförelse mellan fördelningar



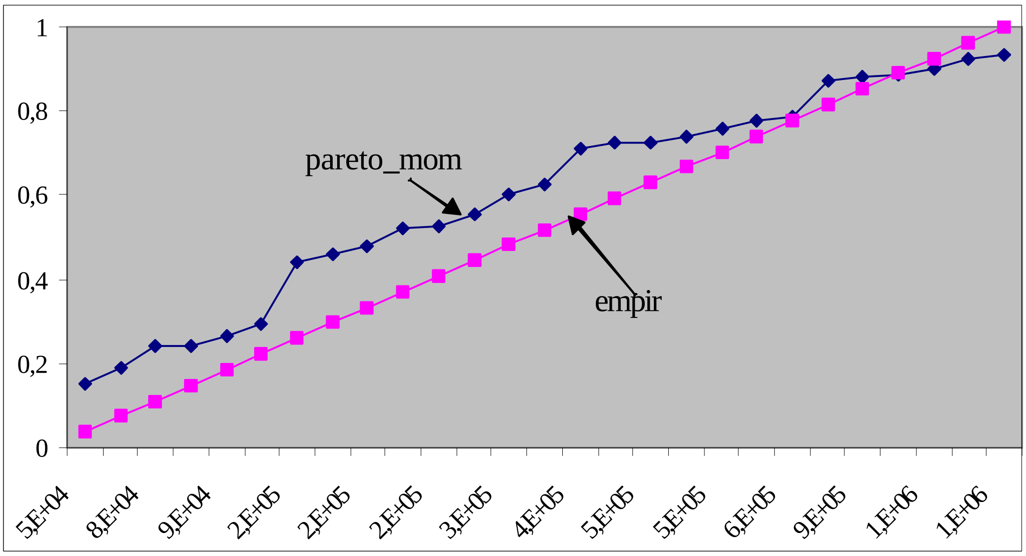
Figur 17 visar empirisk och lognormal fördelning med parametrar skattade med ML- metoden.



Figur 18 Empirisk och lognormalfördelning med momentmetoden



Figur 19 visar empirisk fördelning och Paretofördelning med ML-metoden.



Figur 20 visar empirisk fördelning och Paretofördelning med momentmetoden.

5.3.2 Resultat av chi-2 och Kolmogorov-Smirnovs teststatistikan

	Q_statistika	*K_S_statistika*
lognormal ML-metoden	9.171	0.090
lognormal momentmetoden	33.080	0.159
Pareto ML-metoden	7.189	0.107
Pareto momentmetoden	11.435	0.150

Chi-2 samt Komogorov-Smirnovs värdena i tabell 2 visar att både lognormal och Pareto fördelningen med ML-metoden har bra anpassning i denna skadegren.

Slutsats

Efter att ha studerat plottarna står valet mellan Pareto och lognormalfördelning med ML-metoden. Eftersom det var lite svårt att bedöma tittade vi även på resultat av chi-2 och Kolmogorov-Smirnovs teststatistikan. Lognormal fördelning med momentmetoden har också här det sämsta anpassning medan Paretofördelning med ML metoden passar bäst.

5.3.3 Kontroll av simuleringens noggrannhet

	Log_ML		Log_mom	
	Approx.	Simulerade	Approx.	Simulerade
Skevhet	0.594		0.496	
Väntevärde	1 845 160	1 854 010	1 730 303	1 730 234
Standardavvikelse	1 463 651	1 447 626	1 127 068	1 153 526

	Pareto_ML		Pareto_mom	
	Approx.	Simulerade	Approx.	Simulerade
Skevhet	Går ej att räkna ut		Går ej att räkna ut	
Väntevärde	2 289 518	2 273845	888 482	876 615
Standardavvikelse	2 573 670	2 457 557	1 196 919	1 033 404

Tabell 11 jämförelse av den totala ersättningskostnaden.

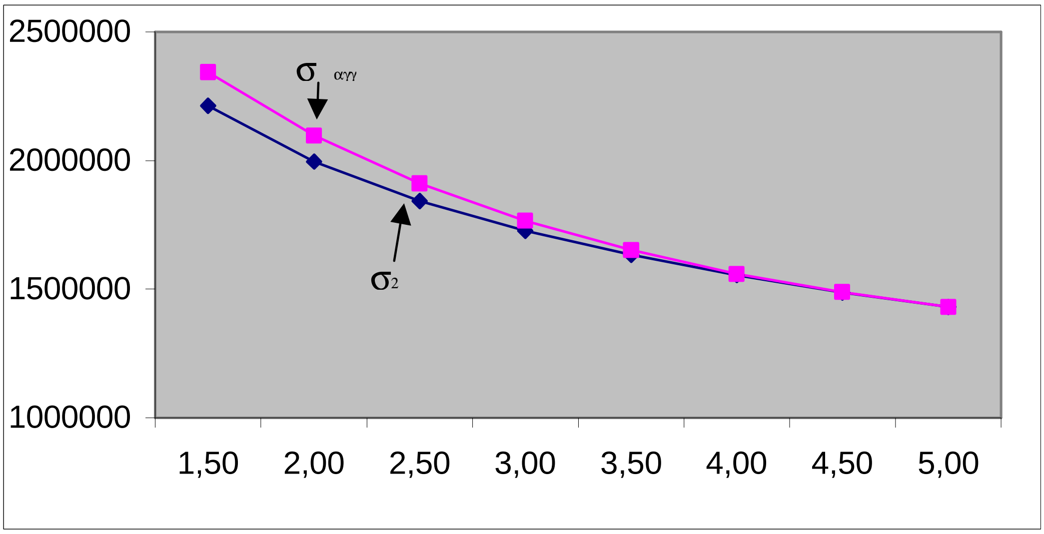
Överensstämmelse är riktigt bra.

5.3.4 Jämförelse mellan standardavvikelse i excess of loss med aggregerat och excess of loss återförsäkring

Paretofördelning med parametrar skattade med ML-metoden

C_1	C_2	μ_2	σ_2	μ_agg	σ_agg	aggregat	σ_agg/σ2
1,00	1,50	1159266	2212944	1159266	2343398	1485683	1,059
	2,00	718470	1996007	718470	2096207	2532068	1,050
	2,50	494574	1841506	494574	1909424	3383738	1,037
	3,00	363804	1725265	363804	1766361	4134903	1,024
	3,50	280443	1632090	280443	1650954	4835664	1,012
	4,00	224057	1553862	224057	1559319	5479206	1,004
	4,50	184243	1487101	184243	1489101	6071634	1,001
	5,00	155052	1429395	155052	1429797	6621290	1,000

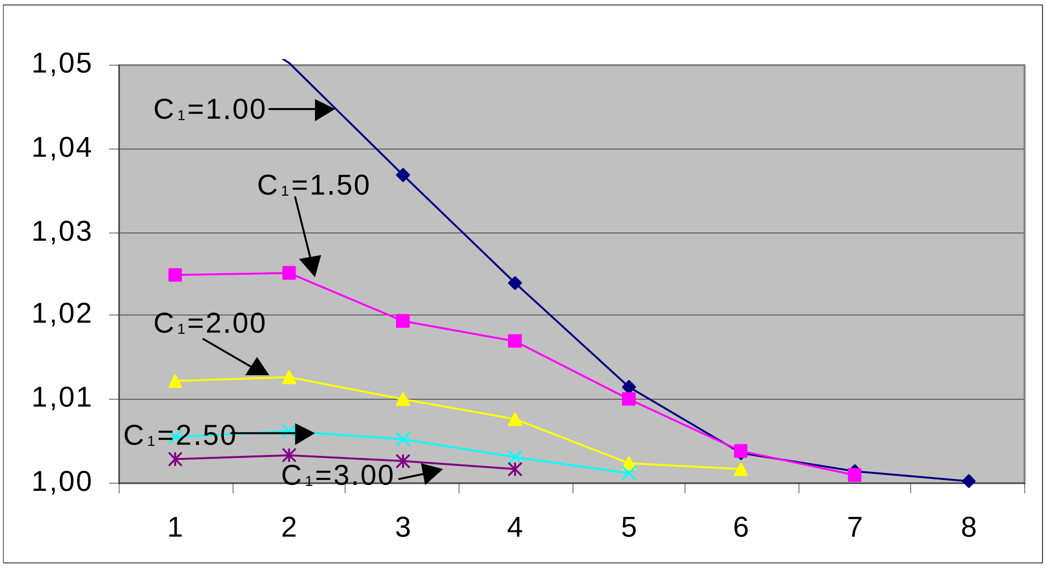
Tabell 12 Självbehåll, väntevärde, standardavvikelse samt kvoten mellan s_agg och s_2 för de två återförsäkringskontrarakter.



Figur 21 visar standaravvikelsen för de två återförsäkringskontrakt då $C_1 = 1.00$ Mkr.

C_1	C_2	μ_2	σ_2	μ_agg	σ_agg	aggregat	σ_agg/σ2
1,50	2,00	718470	1996007	718470	2045505	841710	1,025
	2,50	494574	1841506	494574	1887673	1588190	1,025
	3,00	363804	1725265	363804	1758789	2274150	1,019
	3,50	280443	1632090	280443	1659769	2923475	1,017
	4,00	224057	1553862	224057	1569341	3540765	1,010
	4,50	184243	1487101	184243	1492911	4127032	1,004
	5,00	155052	1429395	155052	1430650	4692513	1,001
2,00	2,50	494574	1841506	494574	1864121	654238	1,012
	3,00	363804	1725265	363804	1747028	1284401	1,013
	3,50	280443	1632090	280443	1648377	1892585	1,010
	4,00	224057	1553862	224057	1565798	2479056	1,008
	4,50	184243	1487101	184243	1490570	3059152	1,002
	5,00	155052	1429395	155052	1431936	3625413	1,002
2,50	3,00	363804	1725265	363804	1734644	582796	1,005
	3,50	280443	1632090	280443	1642110	1156238	1,006
	4,00	224057	1553862	224057	1562140	1724163	1,005
	4,50	184243	1487101	184243	1491899	2292306	1,003
	5,00	155052	1429395	155052	1431029	2847266	1,001
3,00	3,50	280443	1632090	280443	1636771	550109	1,003
	4,00	224057	1553862	224057	1559135	1101874	1,003
	4,50	184243	1487101	184243	1490890	1654292	1,003
	5,00	155052	1429395	155052	1431698	2195178	1,002
3,50	4,00	224057	1553862	224057	1556527	537481	1,002
	4,50	184243	1487101	184243	1490021	1076279	1,002
	5,00	155052	1429395	155052	1431815	1608491	1,002
4,00	4,50	184243	1487101	184243	1488690	528255	1,001
	5,00	155052	1429395	155052	1431270	1053493	1,001
4,50	5,00	155052	1429395	155052	1430395	518694	1,001

Tabell 13 Självbehåll, väntevärde, standardavvikelse samt kvoten mellan aggregerad och vanlig standardavvikelse för de två återförsäkringskontrakt med olika excesspunkter



Figur 22 visar kvoten mellan standaravvikelsen för de 5 först excesspunkterna

Slutsats

Standardavvikelsen för det aggregerade återförsäkringskontraktet är nästan samma som standardavvikelsen för det vanliga återförsäkringskontraktet. Skillnaden mellan metoderna avtar då man ökar C_2 .

6 Slutsats

Målet med detta arbete har varit att jämföra effekten av olika typer av återförsäkringslösningar inom skadeförsäkring. Jämförelsen går ut på att undersöka hur skadekostnadens varians påverkas av valet av kontrakt, givet att väntevärdena hålls lika. För att kunna göra jämförelsen bestämdes först en parametrisk fördelning för skadebeloppen inom återförsäkring. De fördelningar som studerades var Pareto- och lognormalfördelning. Efter att ha studerat de två olika kontrakten kan man sammanfatta resultatet för de olika skadegrenarna som följer.

Civilförsäkring

För denna skadegren visade lognormalfördelningen den bästa anpassningen till skadekostnaderna. Standardavvikelsen för det aggregerade återförsäkringskontraktet är genomgående större än standardavvikelsen för det vanliga excess of loss-kontraktet. Om vi ökar självbehållet i den förra metoden C_1 så avtar skillnaden mellan dem. Om man däremot ökar självbehållet i det vanliga kontraktet C_2 så är denna skillnad nästan konstant.

Företagsförsäkring

Här passade lognormalfördelningen bäst för skadekostnader upp till 3.00 Mkr, däröver är det Pareto-fördelningen som passar bäst. Standardavvikelsen för det aggregerade återförsäkringskontraktet är återigen lite större än standardavvikelsen för vanlig excess of loss. Här är medelvärdet så stort att skillnaden mellan metoderna blir liten. Skillnaden ökar med ökande C_2.

Kaskoförsäkring

Paretofördelning gav bäst anpassning till skadekostnaderna för denna skadegren. Standardavvikelsen för det aggregerade kontraktet är nästan densamma som standardavvikelsen för det vanliga kontraktet. Skillnaden mellan metoderna avtar då man ökar C_2.

Den övergripande slutsatsen blir alltså att skillnaden i standardavvikelse för de två kontrakten är liten utom för civilförsäkring.

7 Referenser

[1] Johansson, B. (1997): *Matematiska modeller inom sakförsäkring*, Kompendium, Matematisk Statistik, Stockholms Universitet

[2] Blom, G., Holmquist, B. (1998): *Statistikteori med tillämplingar,* Studentlitteratur, Lund, tredje upplagan

[3] Råde, L.(1987*)*: *Simulering*, Studentlitteratur, Lund.

[4] Lindgren, B. (1997): *Statistical Theory*, Fourth edition

[5] Hjorth, U. (1998): *Statistisk slutledning i ekonomi och teknik*, Studentlitteratur, Lund, tredje upplagan

[6] Gustafsson, B. (1993): *Återförsäkring i skadeförsäkring*, IFU, 5: e upplagan

[7] Hogg, Robert V. och Stuart A. Klugman (1984): Loss Distributions, Wiley

Appendix

Sannolikhetsfördelningar

Nedan presenteras de sannolikhetsfördelningar som har använts i denna rapport.

A1 ***Lognormalfördelning,***

$X \in LogN(m, s)$

Fördelnings- och täthetsfunktion för *X*

$$F(x) = \Phi\left(\frac{\ln x - m}{s}\right)$$

$$f_x(x) = \frac{1}{\sqrt{2p}\, s x} \exp\{-\frac{(\ln x - m)^2}{2s^2}\}, x > 0$$

Väntevärde: $E(X) = e^{m + \frac{1}{2}s^2}$

Varians : $Var(X) = e^{2m + s^2}(e^{s^2} - 1)$

$$Skevhet(X) = \frac{e^{\frac{9}{2}s^2} - 3e^{\frac{5}{2}s^2} + 2e^{\frac{3}{2}s^2}}{(e^{s^2}(e^{s^2} - 1))^{\frac{3}{2}}}$$

Parameterskattningar med momentmetoden

$$\hat{m} = 2\log(\bar{x}) - \frac{1}{2}\log(s^2 + \bar{x}^2)$$

$$\hat{s}^2 = \log(s^2 + \bar{x}) - 2\log(\bar{x})$$

Parameterskattningar med maximum-likelihood-metoden

$$\hat{m} = \frac{1}{n}\sum_{k=1}^{n} \log x_k$$

$$\hat{s}^2 = \frac{1}{n}\sum_{k=1}^{n} (\log x_k)^2 - \hat{m}^2$$

A2 *Paretofördelningen,*
$X \in Pa(a, g)$

Fördelnings- och täthetsfunktion för *X*

$$F(X) = 1 - \left(\frac{a}{a + x}\right)^{g}, x > 0$$

$$f(x) = \frac{g a^{g}}{(a + x)^{g+1}}, x > 0$$

Väntevärde: $E(X) = \dfrac{a}{g - 1} \qquad , g > 1$

Varians $:Var(x) = \dfrac{g a^{2}}{(g - 1)^{2}(g - 2)}, g > 2$

$$Skevhet(X) = \frac{2\sqrt{(g - 2)}}{g^{1.5}(g - 3)} * (g(g + 1)), g > 3$$

Parameterskattningar med momentmetoden

$$\hat{a} = \bar{x}\frac{s^{2} + \bar{x}^{2}}{s^{2} - \bar{x}^{2}}$$

$$\hat{g} = \frac{2\hat{a}^{2}}{s^{2} - \bar{x}}$$

Fungerar ej om γ<2. Ty variansen existerar inte då.

Parameterskattningar med maximum-likelihood-metoden

$\hat{a}$ löses numerisk ur $\dfrac{1}{a(\hat{a})} + b(\hat{a}) = 1$ där

$$a(\hat{a}) = \frac{1}{n}\sum_{k=1}^{n}\frac{\hat{a}}{\hat{a} + x_k}$$

och

$$b(\hat{a}) = \frac{1}{n}\sum_{k=1}^{n}\log\left(\frac{\hat{a}}{\hat{a} + x_k}\right)$$

$$\hat{g} = \frac{a(\hat{a})}{(1 - a(\hat{a}))}$$

A3 *Empiriska fördelningen*

Den antagande vi utgår från är att vi observerar utfallen $x_1,\dots,x_n$ av *n* observationer på de oberoende stokastiska variablerna $X_1,\dots X_n$, alla med samma fördelningsfunktion *F*.
Den empiriska fördelningsfunktionen skattas som andelen observationer som är mindre än eller lika med x. D.v.s.

$$\hat{F}(x) = \sum_{k:x_k \le x} \frac{1}{n}$$

A4 *Poisson-fördelning,*

$X \in Po(l)$, λ>0

Sannolikhetsfördelning: $P(X = k) = e^{-l}\dfrac{l^k}{k!}, k = 0,1,2,\dots$

Väntevärde: $E(X) = l$

Varians: $Var(X) = l$

$$Skevhet(X) = \frac{1}{\sqrt{l}}$$

Parameterskattningar med momentmetoden

$$\hat{l} = \bar{x}$$

Parameterskattningar med maximum-likelihood-metoden

$$\hat{l} = \bar{x}$$

A5 Fördelning för den totala kostnaden

Ms, Mn, M betecknar de momentgenererande funktionerna för *S*, *N* respektive *X* som antas existera. Då gäller att

$$M_S(t) = E\left[e^{tS}\right] = \sum_{k=0}^{\infty} E\left[e^{tS} \middle| N = k\right] P(N = k)$$
$$= P(N = 0) + \sum_{k=1}^{\infty} E\left[e^{t(X_1+\ldots+X_k)} \middle| N = k\right] P(N = k)$$
$$= P(N = 0) + \sum_{k=1}^{\infty} E\left[e^{t(X_1+\ldots+X_k)}\right] P(N = k)$$
$$= P(N = 0) + \sum_{k=1}^{\infty} E\left[\prod_{i=1}^{k} e^{tX_i}\right] P(N = k) =$$
$$= P(N = 0) + \sum_{k=1}^{\infty} \left[\prod_{i=1}^{k} E\left[e^{tX_i}\right]\right] P(N = k) =$$
$$= \sum_{k=0}^{\infty} (M(t))^k P(N = k) = \sum_{k=0}^{\infty} e^{k \log M(t)} P(N = k)$$
$$= M_N(\log M(t)).$$

där Ψ är den kumulantgenererande funktion som definieras av $\Psi(t) = \log M(t)$.

Eftersom $\Psi_S(t) = l(t-1)$ i Poissonfallet blir den kumulantgenererande funktion för S

$$\Psi_S(T) = \Psi_N(\Psi(t)) = l(e^{\Psi(t)} - 1) = l * (M(t) - 1)$$

Genom att derivera detta uttryck en, två, respektive tre gånger och sätta in t = 0, erhåller vi följande relationer mellan kumulanterna $\{k_j^s\}$, $\{k_j^N\}$ och $\{k_j\}$ för *S*, *N* och X_1 *:*

$$k_1^S = k_1^N k_1,$$
$$k_2^S = k_2^N (k_1)^2 + k_1^N k_2,$$
$$k_3^S = k_3^N (k_1)^3 + 3k_2^N k_1 k_2 + k_1^N k_3.$$

Det följer att den j: te kumulanten för S är

$$k_j^S = \Psi_S^{(j)}(0) = l * M^{(j)}(0) = l * a_j$$

där

$$a_k = E\left[X_1^k\right]$$

$$E\left[S\right] = m_S = k_1^S = l * a_1$$

$$Var\left[S\right] = s^2 = k_2^S = l * a_2$$