



Stockholms  
universitet

Insjuknande och avveckling för svensk  
lång sjukförsäkring 1939-2010  
- Historisk översikt och osäkerhets-  
analys för moderna data

Veronicka Hjorter

Masteruppsats 2010:2  
Matematisk statistik  
Juni 2010

[www.math.su.se](http://www.math.su.se)

Matematisk statistik  
Matematiska institutionen  
Stockholms universitet  
106 91 Stockholm



Matematisk statistik  
Stockholms universitet  
Masteruppsats 2010:2,  
<http://www.math.su.se/matstat>

# Insjuknande och avveckling för svensk lång sjukförsäkring 1939-2010 - Historisk översikt och osäkerhetsanalys för moderna data

Veronica Hjorter\*

Juni 2010

## Sammanfattning

Sjukförsäkring har två viktiga riskmoment att ta hänsyn till, risken att insjukna och sannolikheten att kvarstå som sjuk över tid. Dessa risker har historiskt modellerats på olika sätt. I första delen av detta arbete har modeller från 1939 fram till idag sammanställts och jämförts med varandra. I arbetets andra del har en ny avvecklingsfunktion (som anger sannolikheten att kvarstå som sjuk över tid) skattats fram ur Folksams data. Tekniken som använts i modellframtagandet är densamma som användes då avvecklingsfunktionen senast skattades fram. Osäkerheten i skattningen studeras genom konfidensband framtagna med hjälp av bootstrapteknik.

---

\*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige. E-post: [veronickahjorter@gmail.com](mailto:veronickahjorter@gmail.com). Handledare: Mia Deijfen och Ola Hössjer.

## Sammanfattning

Sjukförsäkring har två viktiga riskmoment att ta hänsyn till, risken att insjukna och sannolikheten att kvarstå som sjuk över tid. Dessa risker har historiskt modellerats på olika sätt. I första delen av detta arbete har modeller från 1939 fram till idag sammanställts och jämförts med varandra. I arbetets andra del har en ny avvecklingsfunktion (som anger sannolikheten att kvarstå som sjuk över tid) skattats fram ur Folksams data. Tekniken som använts vid modellframtagandet är densamma som användes då avvecklingsfunktionen senast skattades fram. Osäkerheten i skattningen studeras genom konfindensband framtagna med hjälp av bootstrapteknik.

## Innehåll

|  |           |
|--|-----------|
| <b>1 Förord</b>  | <b>5</b>  |
| <b>2 Inledning</b>   | <b>6</b>  |
| <b>3 Teori</b>   | <b>7</b>  |
| 3.1 Sjukförsäkring . . . . .                                     | 7         |
| 3.2 Insjuknandeintensiteten . . . . .                            | 8         |
| 3.3 Avvecklingsintensiteten . . . . .                            | 8         |
| 3.4 Avvecklingsfunktionen . . . . .                              | 9         |
| 3.5 t-frekvens . . . . .   | 10        |
| 3.6 Förväntad sjuklängd . . . . .                                | 11        |
| 3.7 Överlevnadsfunktionen . . . . .                              | 12        |
| <b>4 Grundsystemen</b>   | <b>13</b> |
| 4.1 1939 års grunder för lång sjukförsäkring . . . . .           | 13        |
| 4.2 1954 års grunder för lång sjukförsäkring . . . . .           | 15        |
| 4.3 1965 års grunder för lång sjukförsäkring . . . . .           | 17        |
| 4.4 1973 års grunder för lång sjukförsäkring . . . . .           | 18        |
| 4.5 1984 års grunder för lång sjukförsäkring . . . . .           | 20        |
| 4.6 1990 års sjukförsäkringsmodell . . . . .                     | 21        |
| 4.7 Den nya modellen . . . . .                                   | 22        |
| 4.7.1 Frivillig försäkring . . . . .                             | 23        |
| 4.7.2 Obligatorisk försäkring . . . . .                          | 24        |
| 4.7.3 Frivillig mot obligatorisk . . . . .                       | 24        |
| <b>5 Jämförelse av grunderna</b>                                 | <b>27</b> |
| 5.1 Skillnaden mellan män och kvinnor i insjuknandeintensiteten  | 27        |
| 5.2 Skillnader i avvecklingsfunktionen . . . . .                 | 32        |
| <b>6 Konfidensband med hjälp av bootstrap</b>                    | <b>38</b> |
| 6.1 Datamaterialet . . . . .                                     | 39        |
| 6.2 Simulering med hjälp av icke-parametrisk bootstrap . . . . . | 41        |
| 6.3 Skattningsteknik för avvecklingsfunktionen . . . . .         | 41        |
| 6.4 Kaplan-Meier . . . . .                                       | 42        |
| 6.5 Resultat . . . . .   | 44        |
| 6.5.1 Jämförelse av datamaterialet . . . . .                     | 44        |
| 6.5.2 Konfidensband . . . . .                                    | 46        |
| 6.6 Kommentarer . . . . .  | 52        |
| <b>7 Diskussion och slutsats</b>                                 | <b>54</b> |

|   |           |
|---|-----------|
| <i>INNEHÅLL</i>   | 4         |
| <b>8 Appendix</b>   | <b>56</b> |
| 8.1 Härledning av $\kappa_{[x]+t} = -\frac{d}{dt} \ln \lambda_{[x]+t}$ . . . . .                | 56        |
| 8.2 Härledning av $\lambda_{[x]+t}^{(k)} = \frac{\lambda_{[x]+t+k}}{\lambda_{[x]+k}}$ . . . . . | 56        |
| <b>9 BILAGA 1 - Överlevnadsfunktionerna</b>   | <b>59</b> |
| <b>10 BILAGA 2 - Observerade t-frekvenser 2002</b>  | <b>60</b> |

## 1 Förord

Denna uppsats utgör mitt masterarbete i matematisk statistik vid Stockholms universitet.

Jag vill tacka mina handledare för vägledning, kunskap och inspiration genom uppsatsens gång; Gunnar Andersson, doktor i matematisk statistik, föreläsare på Stockholms universitet och framstående inom aktuariebranschen i Sverige, Maria Deijfen, docent och forskarassistent i matematisk statistik och Ola Hössjer, professor i matematisk statistik på Stockholms universitet.

Jag vill även tacka de personer som ingått i sjuklighetsundersökningsgruppen då de har bidragit till diskussion i olika problem. Bland dem vill jag ge ett extra stor tack till Christian Stotzer Salmeron som var källan till hur avvecklingsmodellen tagits fram och till Erik Alm med sina insiktsfulla synpunkter. Samt ett sista tack till företaget Folksam.

## 2 Inledning

Förr var försäkringsbolag som förmedlade lång sjukförsäkring tvungna att följa allmänna riktlinjer, med några få undantag. Dessa riktlinjer presenterades som grunder där det bland annat angavs en insjuknandeintensitet och en avvecklingsfunktion. Insjuknandeintensiteten, som ofta skrivs som  $\nu_x$ , anger intensiteten för en  $x$ -årig individ att bli sjuk, givet att denne är frisk. Avvecklingsfunktionen,  $\lambda_{[x]+t}$ , som är en analytisk utjämning av observationer anger sannolikheten för en  $x$ -årig individ att kvarstå som sjuk  $t$  år efter att han eller hon insjuknat.

1990 släpptes dessa regler och bolagen fick själva skatta sina modeller. Problemet är att de flesta bolag har för lite data för att kunna göra tillförlitliga skattningar. I och med detta satte *Försäkringstekniska forskningsnämnden (FTN)* ihop en grupp med lämpliga aktuarier som fick kallas för SUS-gruppen. De som har ingått i arbetsgruppen är Erik Alm, Gunnar Andersson, Bengt von Bahr, Erland Ekheden, Åsa Larson, Christian Salmeron, Ellinor Samuelsson och Arne Sandström. Deras mål har varit att med hjälp av alla försäkringsbolags data ta fram en ny uppdaterad modell för lång sjukförsäkring. Försäkringsbolagen får sedan själva välja hur de vill hantera denna.

Då den nya modellen ska publiceras ska den även jämföras mot de äldre grundsystemen. En av avsikterna med denna uppsats har varit att genomföra just detta. Syftet är att jämföra de olika grundsystemens avvecklingsfunktion och även skillnaderna mellan könen då det gäller insjuknandet. Underlaget för de olika grundsystemen är begränsad vilket försvårar jämförelsen. Modellerna är inte heller helt jämförbara då bestånden och politiken bakom sjukförsäkringen har förändrats genom tiden.

Arbetets andra syfte har varit att utifrån Folksams data skatta en ny avvecklingsfunktion enligt samma metod som använts i den nyaste modellen och utifrån bootstrapteknik se på osäkerheten i denna skattning genom konfidensband. Denna metod skulle kunna vara till hjälp då reserver för framtida betalningar ska avsättas.

Uppsatsen är upplagd på så sätt att den grundläggande teorin presenteras i kapitel 3 och därefter, i kapitel 4, redogörs för de olika grundsystemen från 1939 och framåt. I kapitel 5 jämförs dessa grundsystem mot varandra, dels genom att se på avvecklingsfunktionen och förväntad tid att betala ut sjukpenning och dels genom att se på hur skillnaderna mellan kvinnor och mäns insjuknandeintensitet har förändrats genom tiden. I kapitel 6 hittas del två av detta arbete, nämligen osäkerhetsanalysen av moderna data som består av Folksamsdata från 2000-01-01 till och med 2007-12-31.



## 3 Teori

### 3.1 Sjukförsäkring

I Sverige har idén med sjukförsäkring funnits sedan mitten av 1800-talet. Syftet med försäkringen är att försäkra sig mot inkomstförluster vid långa sjukdomsfall. Utan en privat sjukförsäkring fås i dagens läge, enligt försäkringskassans hemsida, en sjukpenning på cirka 80 procent av lönen med diverse begränsningar. Många av dem som blir sjukskrivna får däremot inte ut 80 procent av lönen eftersom beloppet som betalas ut per månad är begränsat. Även hur lång tid sjukersättningen kan erhållas är begränsad.

En privat sjukförsäkring täcker upp mellanskillnaden av lönen och försäkringskassans ersättning. Ofta rör det sig om att den försäkrade ska få ut 90 procent av lönen efter en karenstid på tre månader. Vad som gäller kan skilja sig mellan avtalen. En arbetstagaren bör dock inte tjäna på att vara sjuk, detta fick man erfara under senare delen av 1980-talet då sjuktalet var höga i Sverige. Då kunde sjukersättningen uppgå till 100 procent av lönen, vilket troligen var en stor bidragande faktor till de höga sjuktalet. Ett tak infördes därför i början på 1990-talet om att sjukersättningen max fick motsvara 90 procent av lönen.

Bland de olika privata sjukförsäkringarna kan det även skilja sig på hur stor andel den försäkrade ska vara sjukskriven för att ha rätt till ersättning. För Länsförsäkringar och Aspis ligger denna sjukskrivningsgrad på 25 procent, medan t.ex. Folksam ligger på 50 procent.

Tittar vi tillbaka i tiden hade den försäkrade rätt till ersättning om han eller hon uppfyllde kraven för invaliditet. Vad som menades med invaliditet varierar mellan de olika grunderna och ofta finns ingen definition. Allmänt brukar invaliditet definieras som att arbetsförmågan ska vara nedsatt till minst  $m$  procent på grund av sjukdom eller olycksfall, där  $m$  kan vara, och oftast är, 50 procent.

När försäkringen upphör gälla och då den försäkrade ej har rätt till mer ersättning gäller enligt avtal. På de olika försäkringsbolagens hemsidor verkar den vanligaste slutåldern vara 65 år, det vill säga vid pensionsåldern.

Då sjukförsäkring studeras finns två riskmoment att ta hänsyn till, nämligen insjuknande och avveckling. Antaganden om insjuknande används vid prissättningen av en försäkring, men inte för att bestämma de tekniska avsättningarna. I det fallet används avvecklingsfunktionen.

Den modell som SUS-gruppen använder, och som denna uppsats kommer använda sig utav, är den med de tre tillstånden

- (2) = Frisk,
- (1) = Sjuk,
- (0) = Avliden.

Övergångar mellan tillstånden frisk och sjuk kan ske upprepade gånger, medan tillståndet avliden är ett absorberande tillstånd. Hoppet till respektive tillstånd antas inte bero på historien, därav har vi att göra med en Markovprocess. Intensiteten att lämna tillståndet sjuk beror både på åldern och tiden vi tillbringat i just tillstånd (1), i och med detta har vi att göra med en så kallad semi-Markovprocess.

### 3.2 Insjuknandeintensiteten

Hoppet från tillstånd (2) till (1) kallas för insjuknandeintensiteten, förr även kallad invalidiseringsintensiteten. Denna definieras enligt följande.

**Definition 3.1** *Intensiteten för en  $x$ -årig person att bli sjuk, givet att denne är frisk, skrivs som*

$$\nu_x = \lim_{t \rightarrow 0} \frac{1}{t} P(\text{sjuk vid ålder } x + t \mid \text{frisk vid ålder } x).$$

Insjuknandeintensiteten är ett mått i förhållande till hela antalet levande vid samma ålder  $x$ . Antalet  $x$ -åriga personer som under tidsperioden  $\Delta x$  blir arbetsförmögna kan beräknas genom

$$Nl(x) \cdot \nu_x \cdot \Delta x. \quad (1)$$

$Nl(x)$  står då för antalet levande  $x$ -åriga kvinnor eller män, det vill säga både friska och sjuka individer. Två olika insjuknandeintensiteter brukar anges, en för kvinnor och en för män.

Insjuknandeintensiteten skattas fram med hjälp av observerade t-frekvenser (t-frekvensen förklaras i kapitel 3.5) vilket resulterar i att insjuknandeintensiteten även kan komma att bero på karenstiden. En person som har en försäkring med en kort karenstid får då en högre intensitet att insjukna än en med en längre karenstid. För att visa att insjuknandeintensiteten beror på karenstiden används beteckningen  $\nu_x^{(k)}$ .

Ovan har beskrivits att insjuknandeintensiteten beror på faktorerna ålder och kön. Detta är inte hela sanningen. Insjuknande beror även i stor grad på ekonomi och politiska beslut, troligen även på produkt och kundunderlag. Eftersom dessa parametrar inte är med i modellerna kommer det visa sig vara svårt att jämföra de olika insjuknandeintensiteterna. Skillnader mellan män och kvinnor kan ses som mer generella och därför kommer främst dessa skillnader att studeras då det gäller insjuknandeintensiteten.

### 3.3 Avvecklingsintensiteten

Som tidigare nämndes har en sjukförsäkring tre tillstånd som ska tas hänsyn till. Tillstånd (1) kan lämnas på två sätt, vart och ett med en respektive intensitet. Antingen till tillstånd (2) med intensitet  $\mu_{12}(x, t)$  eller till tillstånd

(0) med intensitet  $\mu_{10}(x, t)$ . Orsaken till avveckling är mindre viktig inom sjukförsäkring, ty båda händelserna leder till att utbetalningarna upphör innan försäkringstiden tagit slut. Skulle hänsyn tas till avvecklingsorsak skulle modellerna dessutom bli allt mer komplicerade. Då händelserna är diskjunta kan intensiteterna summeras. Summan kallas avvecklingsintensiteten och betecknas

$$\kappa_{[x]+t} = \mu_{12}(x, t) + \mu_{10}(x, t) \quad (2)$$

där  $[x]$  representerar ålder vid insjuknande och  $t$  durationen (sjuktiden) i år.

Om  $T_x$  är den  $x$ -åriga individens totala sjukdomsfallslängd kan avvecklingsintensiteten även definieras som ett gränsvärde.

$$\kappa_{[x]+t} = \lim_{dt \rightarrow 0} \frac{P(T_x \leq t + dt | T_x > t)}{dt}. \quad (3)$$

I avsnittet nedan kommer avvecklingsfunktionen,  $\lambda_{[x]+t}$ , definieras. Det går då att visa, se Appendix 8.1, att

$$\kappa_{[x]+t} = -\frac{\lambda'_{[x]+t}}{\lambda_{[x]+t}} = -\frac{d}{dt} \ln \lambda_{[x]+t}. \quad (4)$$

Vilket är medför

$$\lambda_{[x]+t} = e^{-\int_0^t \kappa_{[x]+s} ds}. \quad (5)$$

### 3.4 Avvecklingsfunktionen

Många länder har valt att modellera avvecklingsintensiteten, exempelvis England. I Sverige har vi däremot modellerat avvecklingsfunktionen.

**Definition 3.2** Låt den stokastiska variabeln  $T_x$  vara en  $x$ -årig individs totala sjukdomsfallslängd vars fördelningsfunktion kan skrivas som

$$F_x(t) = P(T_x \leq t), \quad t \geq 0.$$

**Definition 3.3** Avvecklingsfunktionen, som är den betingade sannolikheten att en person som är  $x$  år kvarstår som sjuk i  $t$  år efter insjuknandet, kan skrivas som

$$\lambda_{[x]+t} = P(T_x > t) = 1 - F_x(t), \quad t \geq 0.$$

Avvecklingsfunktionen är ett mått på hur stor sannolikheten är att en  $x$ -årig individ inte har tillfrisknat efter  $t$  år och kan till exempel ha formen

$$\lambda_{[x]+t} = a_x e^{-80t} + b_x e^{-13t} + c_x e^{-4,5t} + d_x (0,15e^{-0,3t} + 0,85e^{-0,01t}) \quad (6)$$

där

$$\begin{aligned} a_x &= 1 - b_x - c_x - d_x, \\ b_x &= 0,12, \\ c_x &= 0,006e^{0,04x}, \\ d_x &= 0,00065 + 0,000018 \cdot e^{0,13x}. \end{aligned}$$

Detta är den avvecklingsfunktion som Folksam började använda år 1990. De fem exponentialfunktionerna i avvecklingsfunktionen representerar avvecklingstakten för en "viss sjukdom". Termen  $e^{-80t}$  representerar en hög avvecklingstakt och  $e^{-0,01t}$  en låg.

Vi kan även komma att se på den betingade avvecklingsfunktionen, det vill säga den betingade sannolikheten att en individ kvarstår som sjuk  $t$  år efter karenstiden  $k$ . Denna betecknas  $\lambda_{[x]+t}^{(k)}$ . I detta fall har avvecklingsfunktionen betingats med att individen har varit sjuk under karenstiden.

Då det verkar ha varit vanligt med en karenstid på tre månader är frågan hur tillförlitlig skattningarna för  $\lambda_{[x]+t}$  är då  $t < 0,25$ . I detta arbete ska de äldre avvecklingsfunktionerna jämföras med den nya framtagna avvecklingsfunktionen. Den nya är endast definierad för  $t \geq 0,25$ , därför kommer jämförelsen att göras med  $\lambda_{[x]+t}^{(k)}$  där  $k = 0,25$ .

Om  $\lambda_{[x]+t}$  är given kan den betingade sannolikheten,  $\lambda_{[x]+t}^{(k)}$ , för ett fixt  $t$  beräknas genom

$$\lambda_{[x]+t}^{(k)} = \frac{\lambda_{[x]+t+k}}{\lambda_{[x]+k}}. \quad (7)$$

Se Appendix 8.2 för härledning av sambandet.

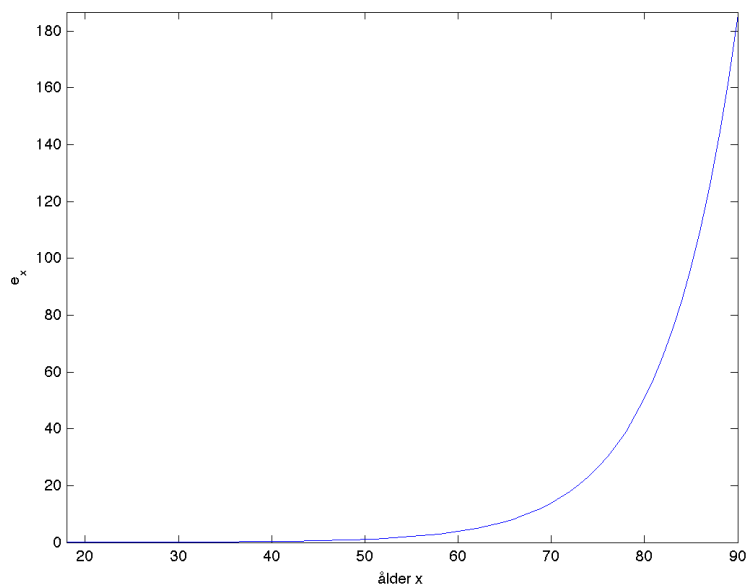
Skrivsättet  $\lambda_{[x]+t}$  är inte att föredra, speciellt vid beräkningar då det gäller att ha koll på hakparenteserna. I och med detta har det nu blivit vanligare att skriva avvecklingsfunktionen som  $\lambda_x(t)$  och avvecklingsintensiteten som  $\kappa_x(t)$ .

### 3.5 t-frekvens

Sannolikheten att en individ, vid  $x$  års ålder, insjuknar under en tidsperiod  $\Delta x$  beräknas med hjälp av insjuknandeintensiteten genom  $\nu_x \cdot \Delta x$ . Sannolikheten att en individ som insjuknat vid  $x$  års ålder kvarstår som sjuk  $t$  år efter insjuknandet definieras liksom tidigare som  $\lambda_{[x]+t}$ . Den kombinerade sannolikheten blir då, se Dillner (1969),

$$\eta_{[x]+t} \cdot \Delta x = \nu_x \cdot \lambda_{[x]+t} \cdot \Delta x, \quad (8)$$

Försäkringsförbundet (2002) beskriver att *t-frekvenser betecknar sannolikheten att en försäkringstagare ska insjukna medan han/hon tillhör åldersintervallet  $x$  samt kvarstå som sjuk  $t$  år senare.*



Figur 1:  $e_x$  för 1990 års grunder under antagandet att avvecklingsfunktionen gäller för alla  $t$  och  $x$

Då karenstiden ofta är tre månader, kommer det vara mer intressant att se på sannolikheten att insjukna och kvarstå som sjuk  $t$  år efter karenstiden. T-frekvensen för karenstiden tre månader blir då

$$\eta_{[x]+0,25} = \nu_x^{(k)} \cdot \lambda_{[x]+0,25}. \quad (9)$$

### 3.6 Förväntad sjuklängd

Förväntad sjuklängd för en individ som insjuknar vid  $x$  års ålder beräknas enligt

$$e_x = E[T_x] = \int_0^\infty (1 - F_x(t)) dt = \int_0^\infty \lambda_{[x]+t} dt. \quad (10)$$

Under antagandet att avvecklingsfunktionen för 1990 års grunder, se (6), gäller för alla  $x$  och  $t$  kan förväntad sjuklängd beräknas med hjälp av ekvation (10).

$$e_x = \frac{a_x}{80} + \frac{b_x}{13} + \frac{c_x}{4,5} + d_x \left( \frac{0,15}{0,3} + \frac{0,85}{0,01} \right)$$

I Figur 1 ser vi att förväntad sjuktid inte verkar vara rimlig. Vid höga åldrar borde dödlighetsintensiteten ta över och förväntad sjuktid borde därmed minska. Införs restriktionen att  $[x]+t \leq 120$  fås inte heller en rimlig funktion. Modellen för avvecklingsfunktionen verkar i detta fall inte gälla för höga

åldrar. Detta är rimligt med tanke på att det data ett försäkringsbolag får in oftast bara är upp till den ålder då sjukersättning upphör betalas ut,  $z$ , som oftast är 65 år. Är så fallet bör avvecklingsfunktionen endast gälla då  $x+t < z$ . Beräkning av förväntad sjuktid kommer därför inte vara möjligt för de flesta utav grunderna. Däremot kan förväntad tid att betala sjukersättning beräknas genom

$$e_x^{\$} = \int_0^{z-x} \lambda_{[x]+t} dt. \quad (11)$$

I väntevärdet (11) har ingen hänsyn tagits till karenstiden. I den nya modellen är avvecklingsfunktionen endast definierad för då karenstiden är minst tre månader, vilket måste tas hänsyn till i jämförelsen av grunderna. Det betyder att då grunderna jämförs mot varandra kommer vi att se på förväntade tiden att betala sjukersättning då individen redan varit sjuk i tre månader.

Förväntad tid att få sjukersättning för en individ som insjuknat vid  $x$  års ålder givet att han eller hon varit sjuk i  $m$  år räknas ut enligt

$$e_{x|m}^{\$} = \int_m^{z-x} \frac{\lambda_{[x]+u}}{\lambda_{[x]+m}} du. \quad (12)$$

Med hjälp av (12) kan då förväntad tid att betala sjukersättning då karenstiden är tre månader beräknas genom att låta  $m = 0,25$ .

Det kan även vara intressant att titta på förväntad tid att betala sjukersättning då individen varit sjuk i ett år. Denna beräknas på samma sätt, nämligen genom att låta  $m = 1$ .

### 3.7 Överlevnadsfunktionen

I beräkningarna kommer ibland överlevnadsfunktionen  $l_x$  att behövas. Enligt samma beräkningar, som i avvecklingsintensiteten, är överlevnadsfunktionen given då dödlighetsintensiteten,  $\mu_x$ , är känd.

$$l_x = e^{-\int_0^x \mu_s ds} \quad (13)$$

Denna anger sannolikheten att en nyfödd uppnår åldern  $x$ . De överlevnadsfunktioner som behövs i uträkningarna för respektive grunder redovisas i Bilaga 1.

## 4 Grundsystemen

Som tidigare nämntes har sjukförsäkring funnits som idé sedan mitten av 1800-talet. I detta avsnitt kommer fokus ligga på grunderna från åren 1939, 1954, 1965, 1973, 1984, 1990 och de nya framtagna grunderna. Förr var försäkringsbolagen tvungna att följa dessa riktlinjer. Undantag kunde ges till bolag som kunde motivera tillräckligt varför de behövde egna grunder. Ett exempel på ett sådant bolag är Svenska Personal Pensionskassan (SPP) som försäkrade privata tjänstemän. På senare år har dessa regler släppts och bolagen har själva fått välja sina modeller.

Årtalen för respektive grundsystem är inte till 100 procent definitiva, det kan skilja sig åt på något år i annan litteratur. Ett skrivsätt som ofta används för de olika grundsystemen är att skriva G73 istället för 1973 års grunder.

### 4.1 1939 års grunder för lång sjukförsäkring

SPP var det bolag som täckte den största delen av sjuk- och invaliditetsrisk området under denna tid. Bolaget grundades 1917 som en pensionskassa och 1924 infördes rätten till *invalidpension*, som det kallades på den tiden. Dessa grunder togs fram för privata tjänstemän och beräkningarna är endast gjorda att gälla fram till de vanligaste slutåldrarna, 60 och 65 år.

Enligt grunderna från 1939 för SPP har den försäkrade rätt till sjukersättning då han eller hon under tre månader varit fullständigt arbetsförmögen på grund av sjukdom eller 50 till 100 procent arbetsförmögen på grund av olycksfall. Den försäkrade får ersättning fram tills den så kallade invaliditetsgraden understiger 50 procent. Med andra ord, för att få sjukersättning på grund av sjukdom var kravet att vara 100 procent sjukskriven i minst 3 månader. Därefter kunde den försäkrade bli partiellt sjukskriven. Vid olycksfall krävdes däremot endast 50 procent sjukskrivning i tre månader för att ha rätt till sjukersättning. Innan G39 kunde sjukersättning endast fås på grund av sjukdom, men utbetalningarna pågick till och med att invaliditetsgraden understeg 25 procent.

Cramér beskriver att innan 1939 användes insjuknandeintensitet  $i_x$ , vilken var densamma för premiefrielseförsäkring som för vanlig *invaliditetsrönteförsäkring* men att  $x$  byttes ut mot  $x + 5$  för den sistnämnda. När han sedan definierar den nya insjuknandeintensiteten,  $\nu_x$ , nämns inget om denna åldersökning, utan han skriver

*Invalidiseringsintensiteten  $\nu_x$  bland samtliga levande antages här given genom formlerna*

$$10^3 \nu_x = \begin{cases} 6 + 0,1 \cdot (x - 40) & \text{för } x \leq 40, \\ 6 + 0,1 \cdot (x - 40) + 0,04 \cdot (x - 40)^2 & \text{för } 40 < x \leq 65, \end{cases} \quad (14)$$

för män, och

$$10^3\nu_x = \begin{cases} 12 + 0,2 \cdot (x - 40) & \text{för } x \leq 40, \\ 12 + 0,2 \cdot (x - 40) + 0,04 \cdot (x - 40)^2 & \text{för } 40 < x \leq 65. \end{cases} \quad (15)$$

för kvinnor.

Intensiteten för kvinnor antogs alltså vara ungefär dubbelt så hög som den för män. Ingen karensfaktor definieras, men då den försäkrade måste ha varit sjuk i minst tre månader för att ha rätt till ersättning kan det antas att  $\nu_x$  är framtagen för då karenstiden är tre månader, eventuellt är den framtagen oavsett karenstid, men det framgår inte. Intensiteten att insjukna och kvarstå som sjuk efter tre månader kan beräknas med hjälp utav (9).

Avvecklingsfunktionen ställer Cramér upp som

Antalet som invalider kvarlevande  $t$  år efter invaliditetens början antages vara omvänt proportionellt mot  $t + 1$ .

Med *invalider* menar Cramér de sjuka och med *antalet som invalider kvarlevande  $t$  år efter invaliditetens början* kan antas att han syftar på avvecklingsfunktionen multiplicerat med antal kvarlevande vid  $t = 0$ , som nedan betecknas med  $N$ . Den matematiska formuleringen borde då vara

$$N \cdot \lambda_{[x]+t} = \frac{c}{t + 1} \quad (16)$$

där  $c$  är en proportionalitetskonstant. Avvecklingsfunktionen är ställd något oklart. Som tur var gavs några funktionsvärden i utredningen för  $\lambda_{[x]+t}^{(k)}$ , där  $t$  är tiden efter karenstiden tre månader. Det visar sig att

$$\lambda_{[x]+t} = \frac{1}{t + 1}, \quad (17)$$

$$\lambda_{[x]+t_0}^{(k)} = \frac{\lambda_{[x]+0,25+t_0}}{\lambda_{[x]+0,25}} = \frac{1,25}{t_0 + 1,25}. \quad (18)$$

I utredningen diskuterades att ha en avvecklingsfunktion som även berodde på åldern för insjuknandet. Funktionen (17) sades vara betryggande så länge  $x < 50$ , däremot för de övre åldrarna kunde en mer betryggande funktion införas. Man ansåg emellertid att det för det praktiska arbetet var en fördel att inte ha åldern för insjuknandet som parameter, vilket också blev fallet. Cramér tillade dock,

Skulle det längre fram visa sig, att en förstärkning av reserven för pågående invaliditetsfall i högre åldrar är behövlig, kan då en sådan ske utan ändring av grunderna för premieberäkningen.



Med andra ord är det upp till varje bolag att öka reserven för sjukfall med höga åldrar, men detta ska inte påverka beräkningen av premien.

Eftersom avvecklingsfunktionen endast är tänkt att användas för åldrar upp till 65 år kan förväntad sjuklängd inte beräknas. Förväntad tid att betala ut ersättning kan däremot beräknas under antagandet att de försäkrade får betalt till åldern 65. Om individen har varit sjuk i  $m$  år blir förväntad tid att betala ersättning

$$e_{x|m}^{\$} = (1 + m) \cdot (\ln(66 - x) - \ln(1 + m)). \quad (19)$$

## 4.2 1954 års grunder för lång sjukförsäkring

Grunderna för 1954 publicerades i en bok utan författare med titeln *1955 års grunder för livförsäkring m.m.*. Grunderna presenterades även efter det av Petrus Mattsson i Skandinavisk aktuarietidsskrift (1956) och diskuterades av Hultman (1958).

Enligt grunderna uppfyller den försäkrade kraven om invaliditet om arbetsförmågan är nedsatt till minst 50 procent på grund av sjukdom eller olycksfall. I grunderna kan läsas att då sjukdom eller olycksfall inträffar innan denne fyllt 14 år får han/hon rätt till ersättning först efter åldern 14.

I boken *1955 års grunder för livförsäkring m.m.* anges att den försäkrade får ersättning efter 70 år om invaliditetsfall inträffat tidigare. Däremot fås ingen ersättning om en olycka inträffar eller invaliditetsgraden ökar efter 70 års ålder.

Till underlag för G54 hade åren 1934 till 1950 studerats, det vill säga en lång period där det insamlade materialet började bli gammalt. Då insjuknandeintensiteten studerades, påvisades det att kvinnor hade ungefär dubbelt så hög insjuknandeintensitet som män. Frekvensen ökade även med åldern. Insjuknandeintensiteten preciserades enligt

*Vid åldern  $x$  år antages invalidiseringsintensiteten  $\nu_x$ , mätt i förhållande till hela antalet levande vid samma ålder, given genom formeln*

$$10^3 \nu_x = \begin{cases} 0,85 \cdot e^{2,5} \cdot \frac{4,3+0,27 \cdot 10^{0,03x}}{l_x} + 133 & \text{för man,} \\ \left(0,85 \cdot e^{2,5} \cdot \frac{4,3+0,27 \cdot 10^{0,03x}}{l_x} + 133\right) \cdot 1,5 & \text{för kvinna.} \end{cases} \quad (20)$$

I kapitlet som behandlar individuell långsjukförsäkring, ur *1955 års grunder för livförsäkring m.m.*, nämns inget om karenstid. Å andra sidan är intensiteten densamma som insjuknandeintensiteten för en premiefrielseförsäkring där det står angivet att *premiebefrielse medgives vid invaliditet, sedan denna varat minst 3 månader*. Om det finns en karenstid eller ej är väldigt oklart i dessa grunder.

I Mattsson (1977) sidan 71 står det att *de s k försäkringskarensbestämmelserna ersattes med en kort väntetid på 30 dagar*. På följande sida kan

sedan läsas att väntetiden slopades 1957. Hultman (1958, sida 383) skriver i sin tur att försäkringen inte gäller om sjukdomen varar kortare tid än tre månader.

I och med att det verkar oklart om det finns någon karenstid eller ej borde insjuknandeintensiteten rimligtvis vara intensiteten att insjukna oberoende av karenstiden. Det vill säga, samma insjuknandeintensitet används oavsett om försäkringen har en karenstid eller ej. Intensiteten att bli sjuk och sedan kvarstå som sjuk efter 3 månader beräknas enligt (9).

Termen  $l_x$  är som tidigare sannolikheten att en nyfödd uppnår åldern  $x$ . För att kunna beräkna denna används dödlighetsintensiteten för en livsfallsförsäkring enligt G54

$$10^3 \mu_x = 1,5 + 0,041 \cdot 10^{0,042x}. \quad (21)$$

För dödlighetsintensiteten fanns inga bestämda antaganden för kvinnor, bolagen fick själva anpassa efter det bestånd de hade.

Avvecklingsfunktionen definieras i G54 som sannolikheten att kvarstå som invalid  $t$  år efter invaliditetens inträde, där denna ges av

$$\lambda_{[x]+t} = \begin{cases} \frac{\nu_x - 0,133}{\nu_x} \cdot f(t) + \frac{0,133}{\nu_x} \cdot g(t) & \text{för } 0 \leq t < 0,25, \\ \frac{\nu_x - 0,133}{\nu_x} \cdot h(t) + \frac{0,133}{\nu_x} \cdot g(t) & \text{för } t \geq 0,25, \end{cases} \quad (22)$$

där

$$\begin{aligned} f(t) &= e^{-10t} \cdot 1,0225^t, \\ g(t) &= \frac{1,0255^t}{1 + 288 \cdot t^2}, \\ h(t) &= \begin{cases} \frac{e^{-2,5 \cdot 1,0225^{0,25}}}{0,75+t} & \text{för } t < 70 - x, \\ \frac{e^{-2,5 \cdot 1,0225^{0,25}}}{70,75-x} \cdot \frac{l_{x+t}}{l_{70}} & \text{för } t \geq 70 - x. \end{cases} \end{aligned}$$

Det anges en avvecklingsfunktion för då durationstiden är mindre än tre månader och en annan då den är större än tre månader. Avvecklingsfunktionen  $\lambda_{[x]+t}^{(k)}$ , där  $t$  är tiden efter tre månader, kan då skrivas som

$$\lambda_{[x]+t}^{(k)} = \frac{1}{\lambda_{[x]+0,25}} \left( \frac{\nu_x - 0,133}{\nu_x} \cdot h(t + 0,25) + \frac{0,133}{\nu_x} \cdot g(t + 0,25) \right). \quad (23)$$

Eftersom avvecklingsfunktionen beror av insjuknandeintensiteten kommer sannolikheten att kvarstå som sjuk vara olika för kvinnor och män. Sannolikheten är något högre för kvinnor jämfört mot männens.

Den förväntade sjuklängden beräknas enligt

$$e_x = \frac{\nu_x - 0,133}{\nu_x} \left( \int_0^{0,25} f(t) dt + \int_{0,25}^{\infty} h(t) dt \right) + \frac{0,133}{\nu_x} \int_0^{\infty} g(t) dt \quad (24)$$

där

$$\begin{aligned}\int_0^{0.25} f(t)dt &= \int_0^{0.25} e^{-10t} 1,0225^t dt, \\ \int_{0.25}^{\infty} h(t)dt &= \int_{0.25}^{70-x} \frac{e^{-2,5} \cdot 1,0225^{0,25}}{0,75+t} dt + \int_{70-x}^{\infty} \frac{e^{-2,5} \cdot 1,0225^{0,25}}{70,75-x} \cdot \frac{l_{x+t}}{l_{70}} dt, \\ \int_0^{\infty} g(t)dt &= \int_0^{\infty} \frac{1,0255^t}{1+288 \cdot t^2} dt.\end{aligned}$$

Förväntade tiden att betala sjukersättning till en individ som insjuknar vid  $x$  års ålder då karenstiden är tre månader är, om  $m = 0,25$

$$e_{x|m}^{\$} = \frac{\frac{\nu_x - 0,133}{\nu_x} \int_m^{z-x} h(u)du + \frac{0,133}{\nu_x} \int_m^{z-x} g(u)du}{\lambda_{[x]+m}}, \quad m \geq 0,25. \quad (25)$$

### 4.3 1965 års grunder för lång sjukförsäkring

Grunderna för 1965 publicerades, precis som G54, i en bok utan författare, men nu med titeln *Grunder och formler för liv- och sjukförsäkring 1964-/65*. Carl-Gösta Dillner publicerade även en artikel i *Skandinavisk Aktuarietidskrift* 1969.

Ersättning kan maximalt fås tills den försäkrade är 70 år enligt G65 men de vanligaste slutåldrarna var 65 och 67 år, se Dillner (1969). Precis som i G54 fick en försäkrad som var minderårig ersättning först då denne fyllt 14 år om han eller hon drabbades av sjukdom eller olycksfall. Insjuknandeintensiteten var definierad som

*Vid försäkring antages invalidiseringsintensiteten  $\nu_x$  vid åldern  $x$  år, mätt i förhållande till hela antalet levande vid samma ålder, given genom formeln*

$$\nu_x^{(k)} = \begin{cases} r(k) \frac{0,2535}{l_x} & \text{om man,} \\ r(k) \frac{0,2535}{l_x} \cdot 1,3 & \text{om kvinna,} \end{cases} \quad (26)$$

där  $r(k)$  är en karensfaktor och skrivs som

$$r(k) = \begin{cases} \frac{1,2-4,2k}{0,65} & \text{för } 0 \leq k \leq 1/12, \\ \frac{0,95-1,2k}{0,65} & \text{för } 1/12 \leq k \leq 3/12, \\ 1 & \text{för } k \geq 3/12. \end{cases} \quad (27)$$

Förhöjningsfaktorn har i G65 sänkts från 50 procent till 30 procent. Ekhult gav en kommentar till detta; *The data gave little support to this, but we wished to avoid prohibitive premium level for women.*

Återigen behövs dödlighetsintensiteten för att kunna beräkna överlevnadsfunktionen.

$$10^3 \mu_x = \begin{cases} 0,6 + 0,034 \cdot 10^{0,042x} & \text{om man} \\ 0,6 + 0,034 \cdot 10^{0,042(x-4)} & \text{om kvinna} \end{cases} \quad (28)$$

Kvinnor har alltså samma dödlighetsintensitet som män, men med en åldersminskning på fyra år.

Inga särskilda antaganden görs för kvinnor i avvecklingsfunktionen,  $\lambda_{[x]+t}$ , och precis som i G54 definieras denna som sannolikheten att kvarstå som invalid  $t$  år efter invaliditetens inträde.

$$\lambda_{[x]+t} = a_x \cdot e^{-51t} + b_x \cdot e^{-13t} + c_x \cdot e^{-3t} + d_x \cdot e^{-0,52t} + e_x \cdot e^{-0,045t} \quad (29)$$

där

$$\begin{aligned} a_x &= 1 - b_x - c_x - d_x - e_x, \\ b_x &= 0,18 \cdot e^{0,015x}, \\ c_x &= 0,019 \cdot e^{0,028x}, \\ d_x &= 0,00073 \cdot e^{0,055x}, \\ e_x &= 0,0017 + 0,000015 \cdot e^{0,11x}. \end{aligned}$$

Beräknas förväntad sjuktid med hjälp utav formel (10) kommer denna, precis som många av de andra grunderna, att bli växande. Avvecklingsfunktionen gäller inte för höga åldrar, om den gjorde det borde dödlighetsintensiteten ta över. Rimligtvis har avvecklingsfunktionen anpassats för de åldrar data finns för vilket borde vara max upp till 70 år.

Förväntad tid att betala sjukersättning då karenstiden är tre månader kan återigen räknas ut med det betingade väntevärdet genom att låta  $m = 0,25$

$$e_{x|m}^{\$} = \frac{\frac{a_x}{51}(e^{-51 \cdot m} - e^{-51 \cdot (z-x)}) + \dots + \frac{e_x}{0,045}(e^{-0,045 \cdot m} - e^{-0,045 \cdot (z-x)})}{\lambda_{[x]+m}}$$

där  $x$  är åldern för insjuknandet.

#### 4.4 1973 års grunder för lång sjukförsäkring

För 1973 års grunder kommer här redovisas det som presenterades av Dillner i *Scandinavian Actuarial Journal* 1974. Grunderna fanns även publicerade av Källström (1990) i boken *Sverige Reinsurance Company 75 years*. Till hands fanns även de grunder som Skandia Liv använde. Två perioder har studerats för att ta fram dessa grunder, 1962-1965 och 1967-1968, där data har delats in i olika åldersintervall.

Ingen definition av invaliditet angavs av Dillner, det vill säga ingen precisering av hur stor grad arbetsförmågan skulle vara nedsatt till för att ha rätt till sjukersättning. I Skandia Livs grunder finns en sådan definition. Nedsättningen av arbetsförmågan ska vara nedsatt till minst 50 procent på grund av sjukdom eller olycksfall för att ha rätt till ersättning.

Källström (1990) beskrev att sjuktalen kraftigt ökade i slutet på 60-talet. Detta kan också ses i insjuknandeintensiteten om värdena ställs upp mot värdena G64 genererar,

$$\nu_x^{(k)} = \begin{cases} r(k) \frac{0,4}{l_x} & \text{om man,} \\ r(k) \frac{0,4}{l_x} \cdot 1,2 & \text{om kvinna,} \end{cases} \quad (30)$$

där  $r(k)$  är en karensfaktor och skrivs som:

$$r(k) = \begin{cases} 2,3 - 10,8k & \text{för } 0 \leq k \leq 1/12, \\ 1,6 - 2,4k & \text{för } 1/12 \leq k \leq 3/12, \\ 1 & \text{för } k \geq 3/12. \end{cases} \quad (31)$$

Insjuknandeintensiteten är förhöjd med 20 procent för kvinnor. Även i Skandia Livs grunder används en förhöjningsfaktor på 20 procent.

Överlevnadsfunktionen  $l_x$  beräknas återigen med hjälp av dödlighetsintensiteten,

$$10^3 \mu_x = 0,6 + 0,034 \cdot 10^{0,042x}. \quad (32)$$

I 1973 års grunder görs ingen skillnad mellan män och kvinnor i avvecklingsfunktionen. Sannolikheten att kvarstå som sjuk  $t$  år efter insjuknandet anges som

$$\lambda_{[x]+t} = a_x \cdot e^{-80t} + b_x \cdot e^{-13t} + c_x \cdot e^{-1,5t} + d_x (0,15 \cdot e^{-0,3t} + 0,85 \cdot e^{-0,04t}) \quad (33)$$

där

$$\begin{aligned} a_x &= 1 - b_x - c_x - d_x, \\ b_x &= 0,12, \\ c_x &= 0,006 \cdot e^{0,04x}, \\ d_x &= 0,001 + 0,000011 \cdot e^{0,13x}. \end{aligned}$$

Här hade Dillner skrivit den andra exponentiellafaktorn som  $-0,13t$  medan Källström angett  $-13t$ . Då Källströms funktion stämmer bättre överens med de övriga grundernas avvecklingsfunktion än vad Dillners gör, kan det antas att det har krupit sig in ett slarvfel i Dillners artikel.

På samma sätt som tidigare blir det problem att beräkna förväntad sjuktid. Avvecklingsfunktionen verkar inte gälla för högre åldrar än utbetalningstiden. Förväntad utbetalningstid går däremot bra att räkna ut. För slutålder  $z = 65$  år och karenstiden tre månader blir denna

$$e_{x|m}^{\$} = \frac{\frac{a_x}{80} (e^{-80 \cdot m} - e^{-80 \cdot (z-x)}) + \dots + \frac{0,85 \cdot d_x}{0,04} (e^{-0,04 \cdot m} - e^{-0,04 \cdot (z-x)})}{\lambda_{[x]+m}}$$

om  $m = 0,25$ .

#### 4.5 1984 års grunder för lång sjukförsäkring

För att sammanställa 1984 års grunder har boken *Sverige Reinsurance Company 75 years* använts med hjälp av aktuarieföreningensprotokoll från 1983-11-22, då boken innehöll en hel del skrivfel. Till hands fanns även Skandia Livs grunder.

De nya insjuknandeintensiteterna hade reducerats i förhållande till 1973 års grunder. För män med ca 20 procent och för kvinnor med ca 25 procent. Skillnaden mellan män och kvinnor hade även minskat sedan G73. Insjuknandeintensiteten för en  $x$ -årig individ angavs som

$$\nu_x^{(k)} = \begin{cases} r(k) \cdot 0,32 \cdot (1 + e^{-5,7+0,065 \cdot x}) & \text{om man,} \\ r(k) \cdot 0,32 \cdot (1 + e^{-5,7+0,065 \cdot x}) \cdot 1,125 & \text{om kvinna,} \end{cases} \quad (34)$$

där  $r(k)$  är en karensfaktor och skrivs som

$$r(k) = \begin{cases} 2,3 - 10,8k & \text{för } 0 \leq k \leq 1/12, \\ 1,6 - 2,4k & \text{för } 1/12 \leq k \leq 3/12, \\ 1 & \text{för } k \geq 3/12. \end{cases} \quad (35)$$

I Aktuarieföreningens protokoll stod den exponentiella faktorn i insjuknandeintensiteten som  $-5,7$  medan Källström hade angivit den som  $-57$ . Efter att ha ritat kurvan, och även sett på angivna värden i tabell, kunde konstateras att Aktuarieföreningens protokoll var mer tillförlitligt. Även i karensfaktorn fanns en skillnad. För  $1/12 \leq k \leq 3/12$  hade Källström skrivit  $1,6 - 21,4k$  medan Aktuarieföreningen hade skrivit  $1,6 - 2,4k$ . Även i detta fall är Aktuarieföreningens protokoll mer tillförlitligt eftersom  $1,6 - 21,4k < 0$  för  $1/12 \leq k \leq 3/12$ .

Källström beskriver att ITP-försäkrade<sup>1</sup> och AGS-försäkrade hade jämförts med G73. I jämförelsen kunde konstateras att antagandena för lång sjukförsäkring måste förstärkas. Av ITP-försäkringen kunde även ses att det fanns en skillnad mellan män och kvinnor. Då det fanns en brist i antal observationer valdes att stärka avvecklingsfunktionen från 1974.

För avvecklingsfunktionen hade observerats i data att funktionen från 1973 stämde väl fram tills en viss brytpunkt, då förtidspension antogs beviljas. Efter brytpunkten visade data en långsammare avveckling än i G73. Sannolikheten att kvarstå som sjuk  $t$  år efter insjuknandes sattes till

$$\lambda_{[x]+t} = \begin{cases} B(x, t) & \text{för } t \leq j(x), \\ B(x, j(x)) \cdot \frac{C(t)}{C(j(x))} & \text{för } t > j(x), \end{cases} \quad (36)$$

där  $B(x, t)$  är avvecklingsfunktionen enligt 73 års grunder. Skillnaden mellan

<sup>1</sup>ITP (industrins och handelns tilläggspension) är en kollektivavtalad tilläggspension för privatanställda tjänstemän. ITP fanns redan 1960, men 1974 blev det en del av kollektivavtalet mellan privattjänstemänkartellen och Svenskt Näringsliv.

kvinnor och män speglas i funktionerna  $j(x)$  och  $C(t)$ . Enligt aktuariereföreningens protokoll gäller att för män

$$j(x) = \begin{cases} 2,5 & \text{för } x \leq 30, \\ 2,5 - 0,07(x - 30) & \text{för } 30 < x < 55, \\ 0,75 & \text{för } 55 \leq x, \end{cases} \quad (37)$$

och för kvinnor

$$j(x) = \begin{cases} 2,25 & \text{för } x \leq 30, \\ 2,25 - 0,06(x - 30) & \text{för } 30 < x < 55, \\ 0,75 & \text{för } 55 \leq x, \end{cases} \quad (38)$$

och

$$C(t) = \begin{cases} 0,15e^{-0,3t} + 0,85e^{-0,03t} & \text{för man,} \\ 0,15e^{-0,3t} + 0,85e^{-0,015t} & \text{för kvinna.} \end{cases}$$

Här redovisar Källström något annorlunda siffror för kvinnor

$$j(x) = \begin{cases} 2,5 & \text{för } x < 30, \\ 2,5 - 0,06(x - 30) & \text{för } 30 \leq x < 55, \\ 0,75 & \text{för } 55 \leq x. \end{cases} \quad (39)$$

I Skandias grunder, som skiljer sig något från de allmänna grunderna, definieras  $j(x)$  på samma sätt som aktuariereföreningen gjorde. Därför valdes att använda (38).

Förväntad tid att betala ersättning för respektive ålder att insjukna är given genom formeln

$$e_{x|m=0,25}^{\$} = \frac{\int_{0,25}^{j(x)} B(x, u) du + \frac{B(x, j(x))}{C(j(x))} \cdot \int_{j(x)}^{z-x} C(u) du}{\lambda_{[x]+0,25}}. \quad (40)$$

#### 4.6 1990 års sjukförsäkringsmodell

Några gemensamma grunder släpptes inte efter 1984. Bolagen fick utifrån sina egna bestånd bestämma vilken modell som speglade deras data bäst. Avvecklingsfunktionen kunde exempelvis ha formen

$$\lambda_{[x]+t} = a_x e^{-80t} + b_x e^{-13t} + c_x e^{-4,5t} + d_x (0,15e^{-0,3t} + 0,85e^{-0,01t}) \quad (41)$$

där

$$\begin{aligned} a_x &= 1 - b_x - c_x - d_x, \\ b_x &= 0,12, \\ c_x &= 0,006 \cdot e^{0,04x}, \\ d_x &= 0,00065 + 0,000018 \cdot e^{0,13x}, \end{aligned}$$

Den angivna avvecklingsformen är den som Folksam använde 1990.

Förväntad utbetalningstid för slutålder ( $z$ ) 65 år och karenstid 3 månader ( $m = 0, 25$ ) beräknas med hjälp av formeln

$$e_{x|m}^{\$} = \frac{\frac{a_x}{80} (e^{-80 \cdot m} - e^{-80 \cdot (z-x)}) + \dots + \frac{0,85 \cdot d_x}{0,01} (e^{-0,01 \cdot m} - e^{-0,01 \cdot (z-x)})}{\lambda_{[x]+m}}.$$

Förhoppningen var även att hitta en insjuknandeintensitet. Den information som kommit fram är att de gamla grundernas intensiteter användes men med vissa justeringar med hjälp av moderna data. Informationen om hur insjuknandeintensiteten såg ut efter uppdateringarna har inte utgetts.

För att ändå kunna jämföra skillnader i könen då det gäller insjuknandeintensiteterna har observerade t-frekvenser från Försäkringsförbundets sjuklighetsundersökning 2002 tagits med. Data i undersökningen var från tidsperioden 1997-01-01 till 2002-01-01 och bolagen som medverkat var Folksam, SEB Trygg Liv, Länsförsäkringar och Nordea. De t-frekvenser som observerades är bifogad i Bilaga 2. Observera att dessa t-frekvenser inte speglar hur insjuknandet såg ut 1990 utan hur insjuknandet såg ut mellan åren 1997 och 2001.

#### 4.7 Den nya modellen

I den nya sjukförsäkringsmodellen har valts att inte ta fram någon ny insjuknandeintensitet. En av anledningarna är att intensiteten till stor del beror på politiska beslut och det ekonomiska läget. I och med detta är det svårt att göra en tillförlitlig skattning, speciellt då sjukskrivning är en sådan het potatis i politiken som den är idag. Däremot har SUS-gruppen studerat skillnaderna mellan könen i denna intensitet och kunnat konstatera att en sådan skillnad finns, se Figur 5.

Vid skattningen av den nya avvecklingsfunktionen har endast sjukfall som pågått i minst nittio dagar studerats. Eftersom de flesta försäkringar har en karenstid på 90 dagar och inte får in uppgifter om kortare sjukfall är det en rimlig avgränsning. Skulle data för kortare sjukfall tagits med i modelleringen skulle en underskattning av dessa skett, ty mycket av data fallit bort.

Datamaterialet bestod av personer som hade insjuknat mellan åldrarna 18-65 år. För att kunna göra bra skattningar har data valts på så sätt att ålder vid insjuknande har legat mellan 25-63 år. De tidiga åldrarna togs bort på grund av för lite data och de högre åldrarna på grund av att en hel del av data troligen rapporterats in felaktigt. Istället för att rapportera in att den försäkrade gått i pension hade rapporterats in att den försäkrade avvecklats, det vill säga blivit frisk eller avlidit.

Avvecklingsfunktionen skattades fram för respektive kön och typ av försäkring där typ av försäkring var frivillig/obligatorisk försäkring eller frivil-



lig/obligatorisk premiefrielseförsäkring. I detta arbete kommer endast de två första typerna att redovisas.

Avvecklingsfunktionen sattes till

$$\lambda_x(t) = \sum_{i=1}^4 f_i(x) e^{-d_i(t-0,25)}, \quad \text{för } t \geq 0, 25, x+t \leq 65 \quad (42)$$

där

$$f_i(x) = a_i + b_i \cdot e^{c_i x},$$

$$f_4(x) = 1 - \sum_{i=1}^3 f_i(x).$$

Närmare beskrivning av hur parametrarna i avvecklingsfunktionen togs fram följer i avsnittet 6.3. Avvecklingsfunktionerna bör användas med stor försiktighet för åldrarna under 28 år och över 61 år eftersom de anpassades till avvecklingsdata i genomsnittsåldrarna  $x = 28, 33, \dots, 58, 61$  år.

Observera att  $\lambda_x(t)$  inte är direkt jämförbart med de övriga  $\lambda_{[x]+t}$ . Avvecklingsfunktionen togs fram för data där sjuktiden pågått i minst tre månader, därefter justerades denna för att  $t$  skulle stå för tiden efter insjuknandet. Det betyder att  $\lambda_x(0,25) = 1$  men för de övriga grunderna är  $\lambda_x(0) = 1$ . I och med detta blir sannolikheten att kvarstå som sjuk  $t$  år efter karenstiden

$$\lambda_x^{(k)}(t) = \sum_{i=1}^n f_i(x) e^{-d_i t}. \quad (43)$$

Förväntad tid att betala sjukersättning då karenstiden är tre månader ( $m = 0,25$ ) för en  $x$ -årig individ beräknas med hjälp av

$$e_{x|m}^{\$} = \frac{1}{\lambda_x(m)} \sum_i f_i(x) \frac{e^{-d_i \cdot 0,25}}{d_i} (e^{-d_i m} - e^{-d_i(z-x)}). \quad (44)$$

#### 4.7.1 Frivillig försäkring

Data bestod av 55 000 inrapporterade sjukfall. För kvinnor togs följande parameterskattningar fram

$$a = [37,4792 \quad 0,3508 \quad 0,1986],$$

$$b = [-36,831 \quad -9,91 \cdot 10^{-9} \quad -1,36 \cdot 10^{-7}],$$

$$c = [0,000278 \quad -15,5935 \quad 0,2433],$$

$$d = [3,2448 \quad 0,9864 \quad 0,3288 \quad 0,005304].$$

För män

$$\begin{aligned} a &= [0,486 \quad 0,309 \quad 0,2653], \\ b &= [-0,0541 \quad -0,9787 \quad -0,00011], \\ c &= [0,0336 \quad -15,5935 \quad 0,1358], \\ d &= [2,8152 \quad 1,1076 \quad 0,3528 \quad 0,006156]. \end{aligned}$$

#### 4.7.2 Obligatorisk försäkring

Data bestod av 28 000 inrapporterade sjukfall. För kvinnor togs följande parameterskattningar fram

$$\begin{aligned} a &= [47,9138 \quad 23,9747 \quad 10,6129], \\ b &= [-46,93422 \quad -34,3621 \quad -0,00002], \\ c &= [0,000225 \quad 0,000046 \quad 0,144], \\ d &= [2,1132 \quad 0,228 \quad 0,2316 \quad 0,011676]. \end{aligned}$$

För män består parametrarna i

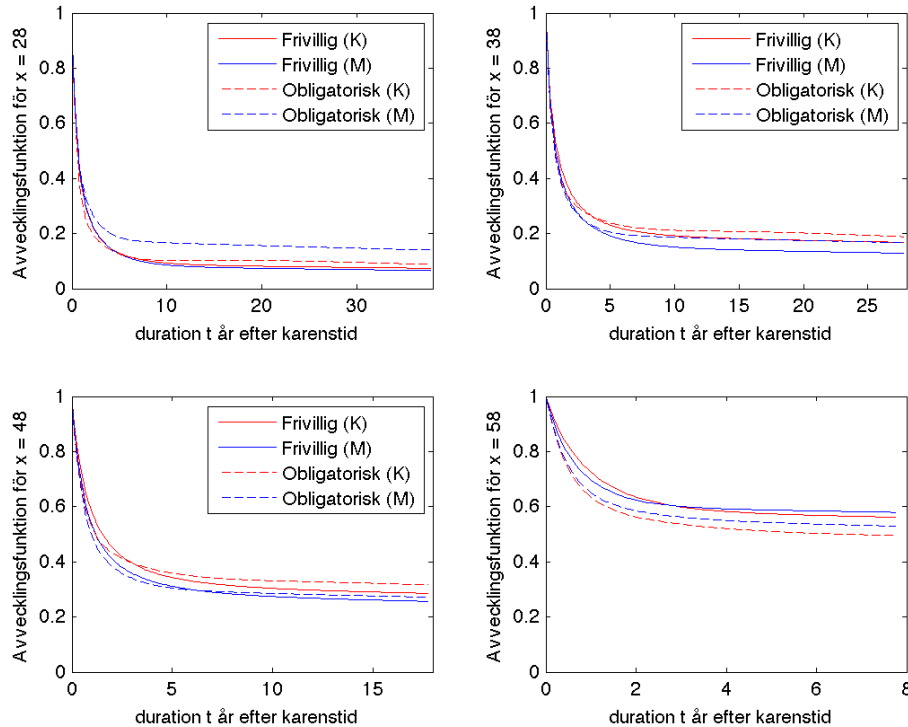
$$\begin{aligned} a &= [54,8588 \quad 187 \quad 0,4999], \\ b &= [-46,9342 \quad -194,8 \quad -0,0033], \\ c &= [0,0023 \quad -0,00062 \quad 0,081], \\ d &= [1,992 \quad 1,9032 \quad 0,6888 \quad 0,006168]. \end{aligned}$$

#### 4.7.3 Frivillig mot obligatorisk

Enligt den nya modellen är sannolikheten att kvarstå som sjuk högre för en man med obligatorisk försäkring än för en man med frivillig försäkring om han insjuknar runt 28 års ålder, se Figur 2. För kvinnor som insjuknar runt 28 års ålder ses inte denna skillnad förrän efter durationstiden 10 år, och även då är skillnaden liten jämfört mot männens. I den obligatoriska försäkringen är skillnaden mellan kvinnor och män i denna ålder stor. Kvinnor tenderar att avvecklas från försäkringen i högre grad än männen. I den frivilliga försäkringen är skillnaden mellan män och kvinnor knappt märkbar och troligen inte relevant.

Bland de som insjuknar runt åldern 38 år stannar större andel kvar som sjuka då de har en obligatorisk försäkring än om de har en frivillig. Vad gäller skillnaden mellan kvinnor och män ser vi att kvinnor i högre grad kvarstår som sjuka. Liknande mönster finns för de som insjuknar runt åldern 48 år, däremot är skillnaden mellan obligatorisk och frivillig försäkring mindre för männen.

Då ålder att insjukna är hög kan en tydlig skillnad ses mellan obligatorisk och frivillig försäkring. De som har en frivillig försäkring tenderar att vara

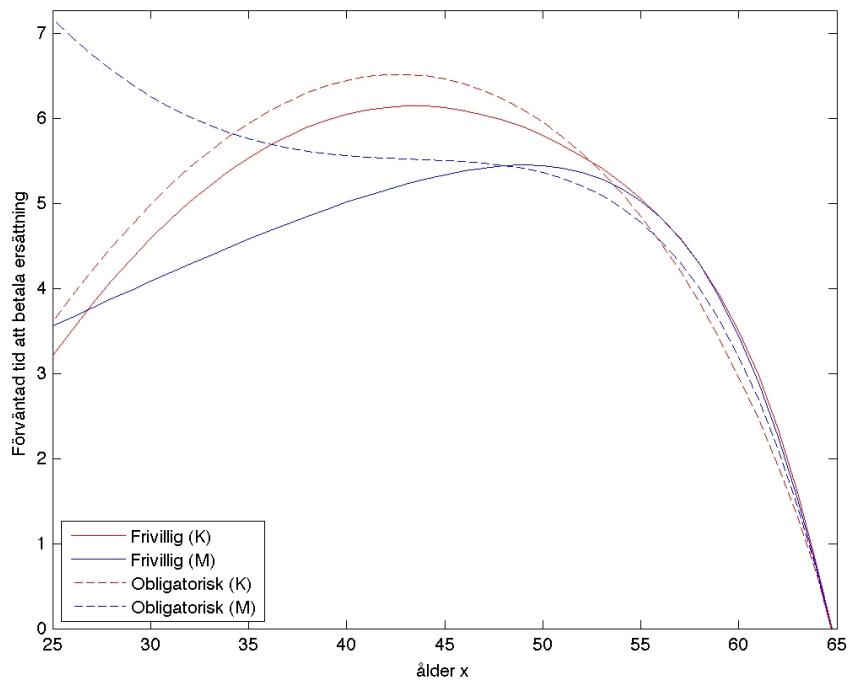


Figur 2: *Avvecklingsfunktionen för frivillig respektive obligatorisk sjukförsäkring.*

sjuka längre. Då ålder vid insjuknande är 58 år kan observeras att andelen män som fortfarande är sjuka efter durationstiden 3 år är större än andelen kvinnor. Observationerna bör dock tas med stor försiktighet. Det finns avvikelser från modellen i data och skillnaderna i figuren är små.

Figur 3 visar en plott över förväntad tid att betala ersättning. Den förväntade tiden är längst för en man med obligatorisk försäkring då ålder att insjukna är under 34 år. Därefter tar kvinnorna, både de med obligatorisk och frivillig försäkring, över. Under insjuknande ålder 50 år verkar de med obligatorisk försäkring alltid att ha en högre förväntad tid att vara sjuk än de med en frivillig försäkring.

Sammantaget kan vi utifrån Figur 3 se att kvinnor har en högre förväntad tid att få ersättning då ålder vid insjuknande är mellan 35-54 år. Vid högre insjuknande åldrar börjar männen få en något längre förväntad tid. De med obligatorisk försäkring tenderar att vara sjuka en längre tid än de med frivillig försäkring upp till insjuknande ålder 50-55 år. Därefter har de med frivillig försäkring en längre förväntad tid att vara sjuka.

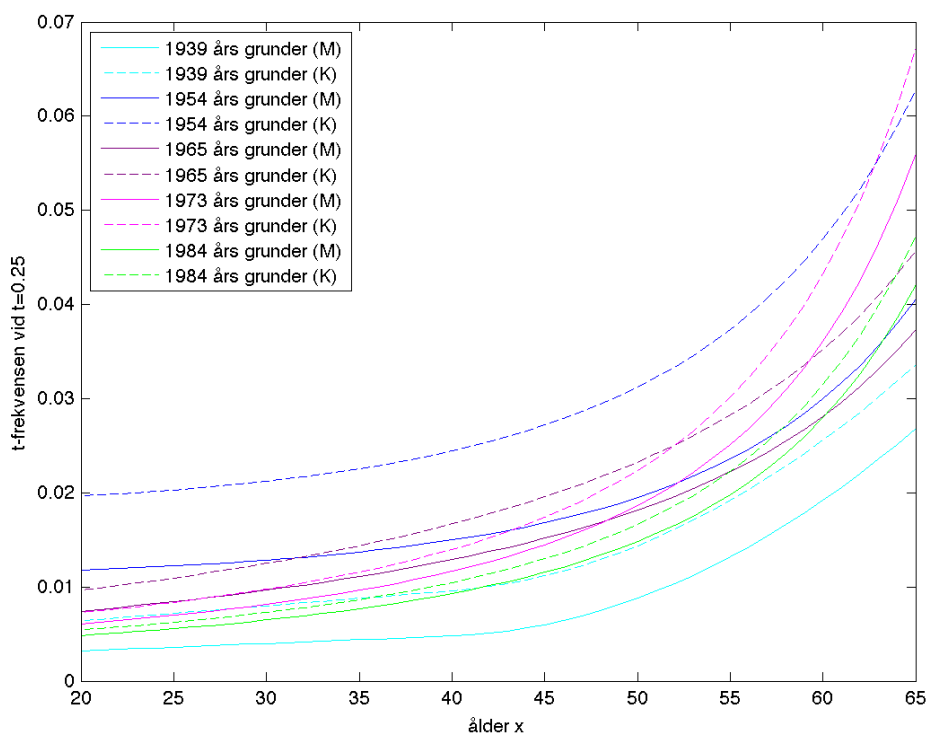


Figur 3: Förväntad tid att betala sjukersättning

## 5 Jämförelse av grunderna

### 5.1 Skillnaden mellan män och kvinnor i insjuknandeintensiteten

Insjuknandeintensiteten i de olika grundsystemen modellerades med hjälp av ålder och kön. I försäkringssyfte är det mer intressant att titta på sannolikheten att insjukna och kvarstå som sjuk efter karenstiden än att titta på enbart insjuknandeintensiteten. Om karenstiden är tre månader fås Figur 4 för t-frekvensen.



Figur 4: *t*-frekvensen för duration tre månader då karenstiden är tre månader. (M) visar männens *t*-frekvens medan (K) visar kvinnornas.

Jämförs kvinnornas *t*-frekvenser kan avläsas från figuren att G39 har lägst antaganden om sannolikheten att insjukna och kvarstå som sjuk efter karenstiden medan G54 har högst antaganden.

*t*-frekvensen har sjunkit nästan en hel procent från G54 till G65. 1973 antogs unga kvinnor ha en lägre *t*-frekvens än 1965 års kvinnor, medan de äldre antogs ha en högre frekvens. I grunderna från 1984 sänktes antagandena ytterligare.

För männen kan ses ett liknande mönster, men hoppet ned från G54 till G65 är mycket mindre jämfört mot kvinnornas.

En orsak till att G39 skiljer sig så markant från G54 är att grunderna bygger på olika bestånd. G39 är framtagna för SPP, det vill säga för privata tjänstemän. Arbetare borde rimligtvis ha en högre insjuknandeintensitet än tjänstemän och detta borde därför vara en bidragande effekt till det stora hoppet till G54. Bestånden kan troligen även ha varierat i övrigt genom tiden, men detta nämns inget om i grundsystemen. Att invaliditetsgraden skiljer sig åt gör även den att insjuknandeintensiteterna är svåra att jämföra.

SUS-gruppen valde att inte ta fram någon ny insjuknandeintensitet i och med att politiken, ekonomin och andra faktorer påverkar för mycket för att kunna göra tillförlitliga skattningar. Istället studerades skillnaderna mellan könen. Då samma problem även borde finnas i de äldre grundsystemen kommer inte någon större vikt läggas vid jämförelsen av dessa, istället kommer nu att fokuseras på skillnaderna mellan könen.

För att studera kvoten mellan kvinnor och mäns insjuknandeintensitet har följande beräkningar gjorts för datamaterialet från 2000 till 2007.

Antal sjuka efter karenstiden  $\Delta x_j$ ,  $SM_{i,j}(w, y)$ , och antal försäkringskontrakt med karenstiden  $\Delta x_j$ ,  $SN_{i,j}(w, y)$ , summeras för åldersgrupperna  $w = [25-29, 30-34, \dots, 60-64]$  över alla bolag, där  $i$  står för bolagsnummer och  $y$  står för kön. Risken att insjukna för respektive åldersgrupp,  $w$ , och kön,  $y$ , skattas genom

$$\hat{P}(w, y) = \frac{\sum_j \sum_i SM_{i,j}(w, y)}{\sum_i \sum_j SN_{i,j}(w, y)}$$

Kvoten av riskerna mellan kvinnor och män definieras som

$$R(w) = \frac{P(w, \text{kvinnor})}{P(w, \text{män})}$$

Då denna kvot togs fram har insjuknandet samlats in från samtliga sjukförsäkringsprodukter och sedan har kvoterna viktas ihop. Till största del är det en karenstid på 3 månader, men det finns även de med kortare och längre karenstid.

Om alla sjukfall som inträffat under tre månader hade tagits med hade  $\hat{P}(w, y)$  varit en skattning av  $\nu_x^{(y)} \Delta x$  där  $\Delta x$  är tidsperioden tre månader. I sådant fall hade  $\Delta x$  kunna divideras bort och en jämförelse av kvoten mellan insjuknandeintensiteterna hade passat. Då många sjukfall säkerligen missats på grund av karenstiden är detta inte riktigt sant.

Kanske vore det därför bättre att jämföra  $R(w)$  med kvoterna av sannolikheten för män och kvinnor att insjukna och kvarstå som sjuka efter tre månader. Dessa är dock inte helt jämförbara heller eftersom det finns de sjukfall som har haft en karenstid som är längre än tre månader, vilket gör att t-frekvensen troligtvis bör vara större då en del sjukfall kan ha missats. Möjligtvis kan det även finnas med sjukfall där sjuktiden varit kortare än tre månader.

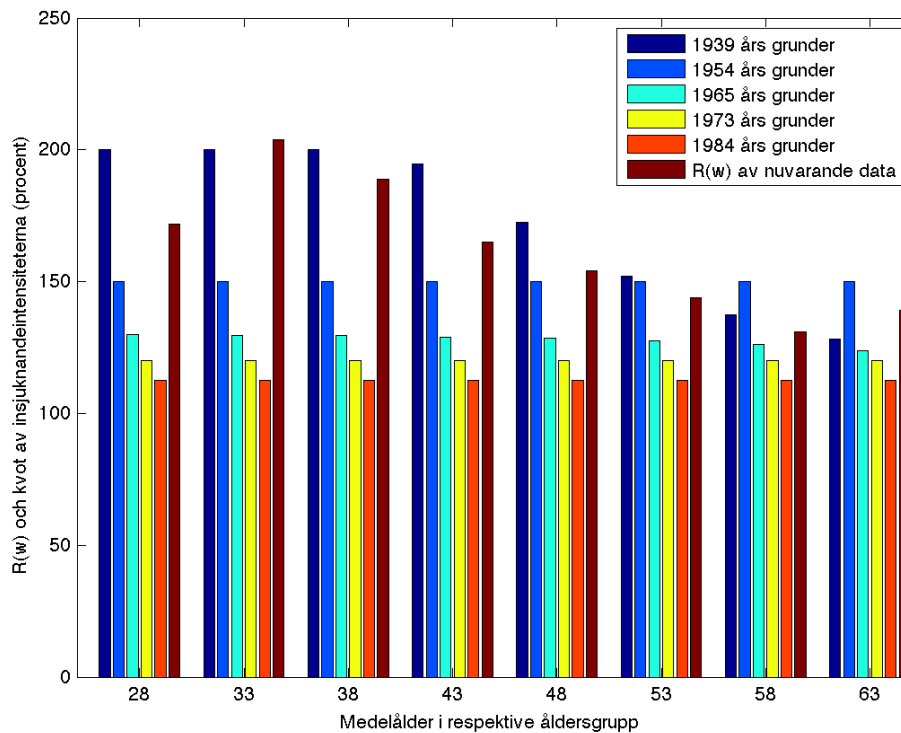
Kvoterna  $R(w)$  kommer därför jämföras dels mot kvoterna av insjuknandeintensiteterna och dels mot kvoterna av t-frekvenserna.

Som ovan nämndes är  $w$  en åldersklass. Den exakta medelåldern för data har inte angivits, däremot finns tillgång till den ungefärliga medelåldern,  $x(w)$ . Jämförelsen med  $R(w)$  sker mot kvoter där insjuknandeålder sätts lika med den ungefärliga medelåldern vid insjuknandet.

Först jämförs  $R(w)$  med kvoten av insjuknandeintensiteterna,

$$\frac{\nu_{x(w)}^{(\text{kvinnna})}}{\nu_{x(w)}^{(\text{man})}}. \quad (45)$$

Denna kvot beräknas för alla olika grundsystem. Karensfaktorn i insjuknandeintensiteten gör sannolikheten att insjukna mindre vid långa karenstider än för dem med kortare karenstider. Denna faktor är en funktion som endast beror på just karenstiden och därför kommer den att kunna divideras bort. Med andra ord, karenstiden spelar ingen roll i jämförelsen av kvoterna av  $\nu_{x(w)}$ .



Figur 5: Kvoten mellan kvinnors och mäns insjuknandeintensitet och  $R(w)$  för dagens data.

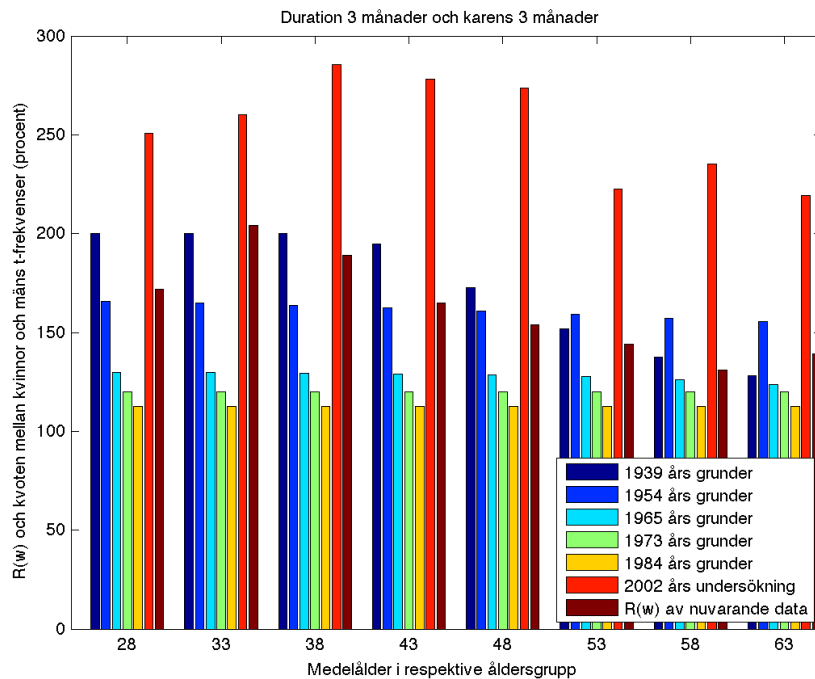
I Figur 5 kan ses hur skillnaderna mellan kvinnors och mäns insjuknan-

deintensitet har minskat stadigt från 1939 ända till 1984 års grunder, med undantag för 1954 års grunder i de högre åldrarna. Dagens data visar en stor ökning av denna skillnad.

För G54, G65, G73 och G84 har skillnaden bestått av en multiplikativ konstant. Det betyder att ingen hänsyn har tagits till hur skillnaderna har legat för olika åldrar. Antingen har åldern inte varit en bidragande faktor i könsskillnaderna eller så har de jämnat ut skillnaderna till den multiplikativa konstanten. Tillbaka i tiden var det i huvudsak män som var försäkrade och underlaget för kvinnor var dåligt. Ett dåligt underlag försvårar självklart skattningen av kvinnors insjuknandeintensitet, dessutom fanns inte samma teknik som vi idag har tillgång till. Hur pass trovärdiga dessa skillnader är kan därför ifrågasättas. Ekhult kommenterade dessutom i G65 att data inte visade minskningen från G54 till G65, om modellerna verkligen speglar de verkliga skillnaderna i alla fallen kan därför också ifrågasättas.

Tidigare nämndes att grunderna kan ha byggt på olika bestånd, det kan även spela roll för de observerade skillnaderna mellan könen.

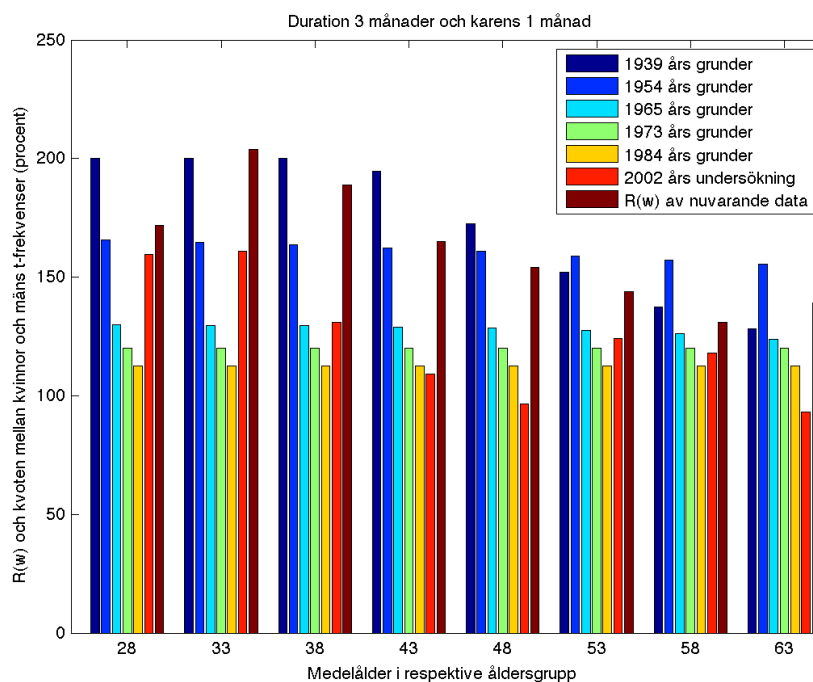
Kvoten av de olika t-frekvenserna har i Figur 6 och 7 ställts upp tillsammans med  $R(w)$ . De observerade t-frekvenserna från 2002 är även med i figurerna.



Figur 6:  $\frac{\eta_{x(w)}^{kvinna}(0,25)}{\eta_{x(w)}^{man}(0,25)}$  tillsammans med observerade t-frekvenserna från 2002 då karenstiden varit tre månader och observerade  $R(w)$  av moderna data

I båda figurerna är durationstiden tre månader, men i Figur 6 är karensti-





Figur 7:  $\frac{\eta_{x(w)}^{kvinnor}(0,25)}{\eta_{x(w)}^{män}(0,25)}$  tillsammans med observerade t-frekvenserna från 2002 då karenstiden varit en månad och observerade  $R(w)$  av moderna data

den 3 månader medan den i Figur 7 är en månad. Kvoten  $R(w)$  är definierad enligt ovan vilket betyder att den innehåller både försäkringar med tre månaders karenstid och en månads karenstid. Hur lång karenstiden är spelar ingen roll för kvoterna i de olika grundsystemen, så länge durationstiden är densamma. Det är därför anmärkningsvärt att karenstiden verkar spela så stor roll för de observerade t-frekvenserna från 2002. Något svar om varför det är på detta sätt har inte hittats.

Kvoten  $R(w)$  borde vara bäst lämpad att jämföras med kvoterna av t-frekvenserna då karenstiden är tre månader, Figur 6. Detta eftersom de flesta försäkringarna i den nya undersökningen har haft en karenstid på tre månader vilket betyder att de troligtvis behövt vara sjuka i tre månader för att observeras.

I båda figurerna kan ses hur kvoten minskar stadigt, liksom kvoterna av insjuknandeintensiteterna, med undantag (som tidigare) för G54 i de höga insjuknande åldrarna. Då karenstiden är tre månader ökar kvoten drastiskt i alla åldersgrupper 2002. Tittar vi däremot i Figur 7, då karenstiden endast är en månad, har skillnaden till och med minskat i några åldersgrupper och i andra gjort en mindre ökning. Kvoten  $R(w)$  av moderna data visar att skillnaden ungefär är lika stor mellan kvinnor och män som den var 1939, under antagandet att modellerna är sanna.

Slutsatsen som kan dras då könen jämförs är att kvinnor är och har varit sjuka i högre grad än männen. Under antagande att modellerna speglar verkligheten kan ses att skillnaderna minskar stadigt med åren för att sedan öka igen någon gång under 1990-talet (alternativt 2000-talet). Eftersom både samhällets uppbyggnad och arbetsfördelning har förändrats genom tiden och bestånden kan ha sett olika ut speglas troligen dessa skillnader även av andra faktorer som inte finns med i modellerna. Dessutom har underlaget för kvinnor varit sämre tillbaka i tiden, vilket gör att det är svårt att uttala sig om skillnaderna mellan könen.

## 5.2 Skillnader i avvecklingsfunktionen

Olika avvecklingsfunktioner för män och kvinnor har angetts i den nya modellen och även i grunderna från 1954 och 1984. I de övriga grunderna ansågs att skillnaden mellan könen inte var tillräckligt stor för att särskilja avvecklingsfunktionerna. Det betyder att data från både män och kvinnor ligger som underlag till de sistnämnda, medan data har delats upp i de förstnämnda. I Figur 8 kan alla dessa avvecklingsfunktioner observeras. De streckade linjerna representerar avvecklingsfunktionen för kvinnor om en sådan finns.

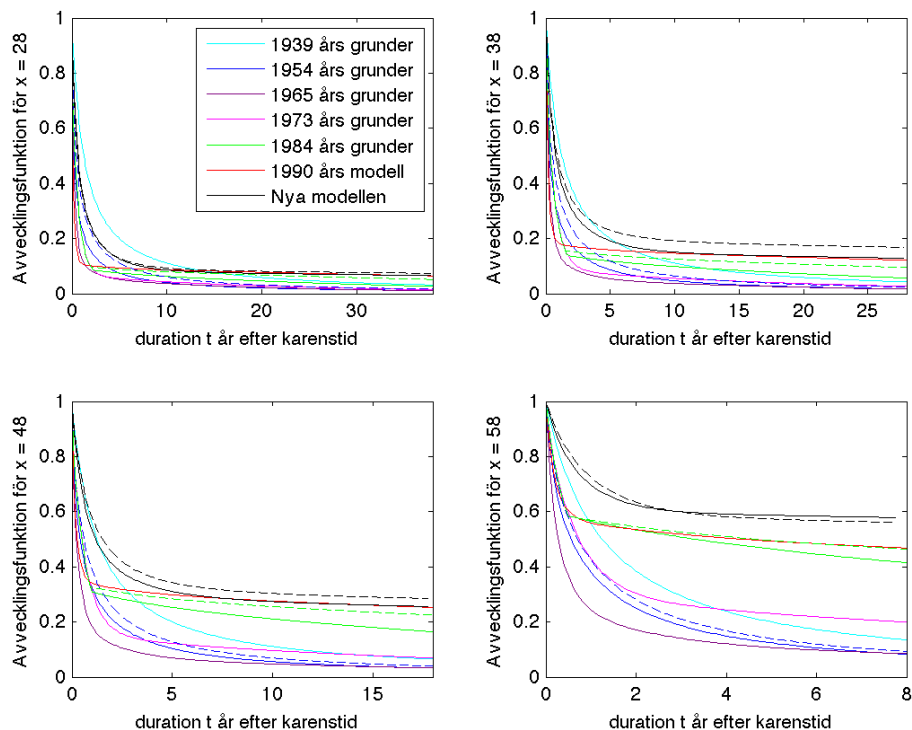
I grunderna framgår det inte om avvecklingsfunktionen är för frivillig och / eller obligatorisk försäkring. Här har valts att jämföra grunderna med den frivilliga försäkringen.

Figur 8 är svår att avläsa, det vore bättre att studera början av avvecklingsfunktionen för sig och svansen för sig. För att ha färre antal kurvor kommer kvinnornas avvecklingsfunktion inte att tas med i följande figurer. Innan vi lämnar figuren kan vi konstatera att en större andel kvinnor tenderar att vara kvar som sjuka i jämförelse med männen. Ett undantag finns och det är i den nya modellen bland de som insjuknar vi höga åldrar. Efter cirka 3 års sjuklängd har männen en något högre andel kvar som sjuka.

För samtliga modeller och åldersgrupper har G39 lägst avveckling de första månaderna efter karenstiden, med undantag från de högsta insjuknande åldrarna där den nya modellen har lägst avveckling, se Figur 9. Högst avveckling upp till insjuknandeålder 40 år har 1990 års modell tätt följt av G65. Därefter har G65 högst avveckling.

Då de första månaderna studeras kan ses hur avvecklingen, från G39, har ökat vid G54. G65 visar en ytterligare ökning, för att i G73 minska. Att G73 och G84 är identiska till en början är enligt definition, vilket även ses i Figur 9. I modellen från 1990 har avvecklingen ökat igen, förutom då ålder att insjukna varit hög. I figuren kan även observeras att från 1990 års avvecklingsfunktion till den nya framtagna avvecklingsfunktionen har det skett en kraftig minskning av avvecklingen de första månaderna.

I Figur 10 kan svansen av avvecklingsfunktionen studeras. Här visas hur stor andel sjuka som är kvar efter två års sjuktid och framåt. Denna andel styrs även av hur avvecklingen sett ut i början av perioden. De olika avveck-

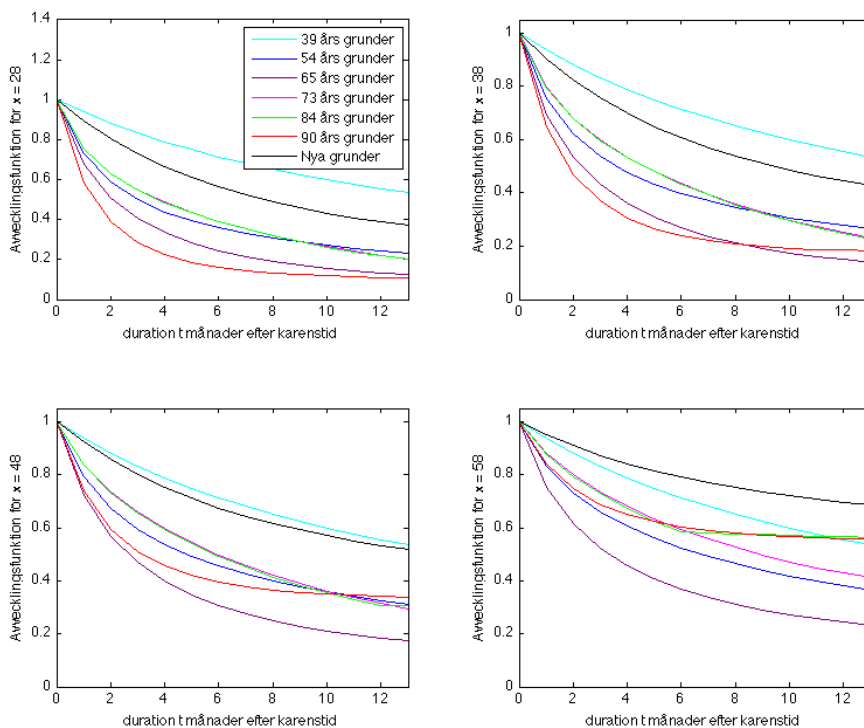


Figur 8: Avvecklingsfunktionen mot durationstid, för de grunder som definierar en avvecklingsfunktion för män och en annan för kvinnor är (-) för män och (- -) för kvinnor.

lingsfunktionerna verkar ha ett liknande mönster för de olika insjuknande åldrarna. Det vill säga de grunder som har högst eller lägst andel kvar vid långa durationstider verkar genomgående ha det för alla åldersgrupperna. G39 skiljer sig däremot åt olika från de övriga modellerna i de olika åldersgrupperna. Eftersom denna avvecklingsfunktion i sig är lite suspekt kommer ingen större vikt läggas vid att jämföra G39 mot de övriga.

Den nya modellen och 1990 års modell har störst andel sjuka kvar vid långa durationstider, deras avvecklingsfunktion är dessutom planare i slutet än de övrigas. G65, därefter G54, har lägst andel sjuka kvar vid de längre durationstiderna. Som väntat har G84 en högre andel kvar än G73, då det var just därför som G84 justerades.

I de grunder som har angett en avvecklingsfunktion för män och en annan för kvinnor kan konstateras, utifrån Figur 11, att ersättning förväntas betalas ut en längre tid för kvinnor än för män i alla åldrar. I den nya modellen kan ses att bland de individer som insjuknar mellan 25-27 års ålder har männen en längre förväntad tid. Detta skulle kunna bero på sättet som



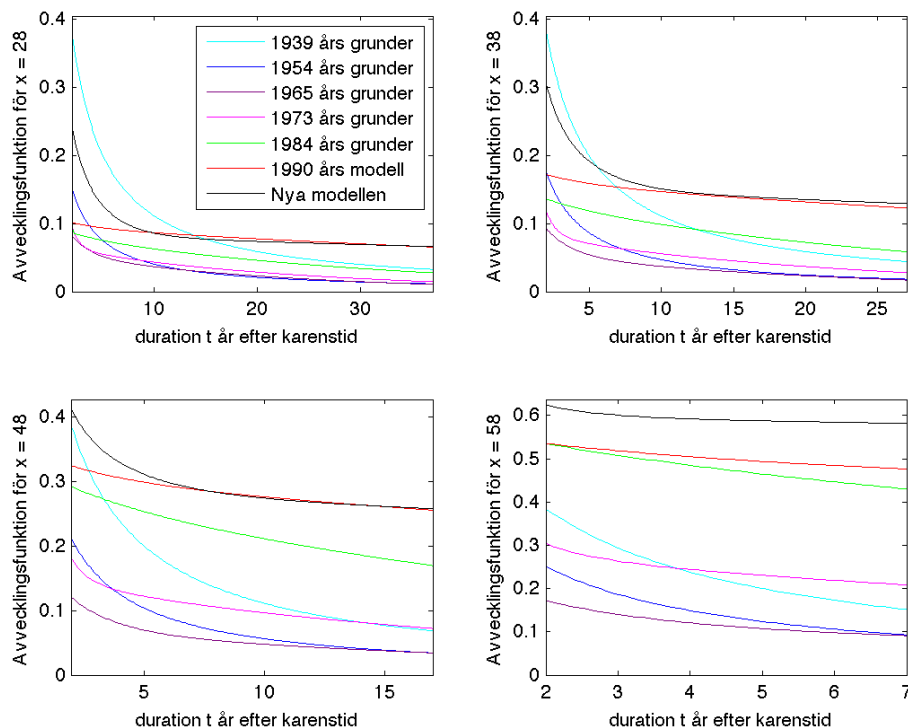
Figur 9: *Avvecklingsfunktionen i början*

avvecklingsfunktionen tagits fram. SUS-gruppen har diskuterat att använda  $\lambda_{28}(t)$  även för åldrar under 28 år, just beroende på hur modellen skattats fram.

Kortast förväntad sjuktid att få ersättning har G65, därefter G54. I topp ligger den nya modellen och sedan G84 och 1990 års modell. I Figur 10 kunde observeras att den nya modellens och 1990 års avvecklingsfunktion nästan hade samma tjocka svans. Här ses en skillnad på nästan ett år i förväntad tid att betala ersättning (1,5 år om kurvan jämförs med kvinnornas). Förklaringen till detta ligger troligen i att det förväntade värdet i 1990 års modell dras ned på grund av en hög avvecklingstakt det första året.

Slutsatserna som kan dras av Figur 11 är att det efter 1973 års grunder verkar ha hänt något. Den genomsnittliga sjuklängden verkar bara bli längre.

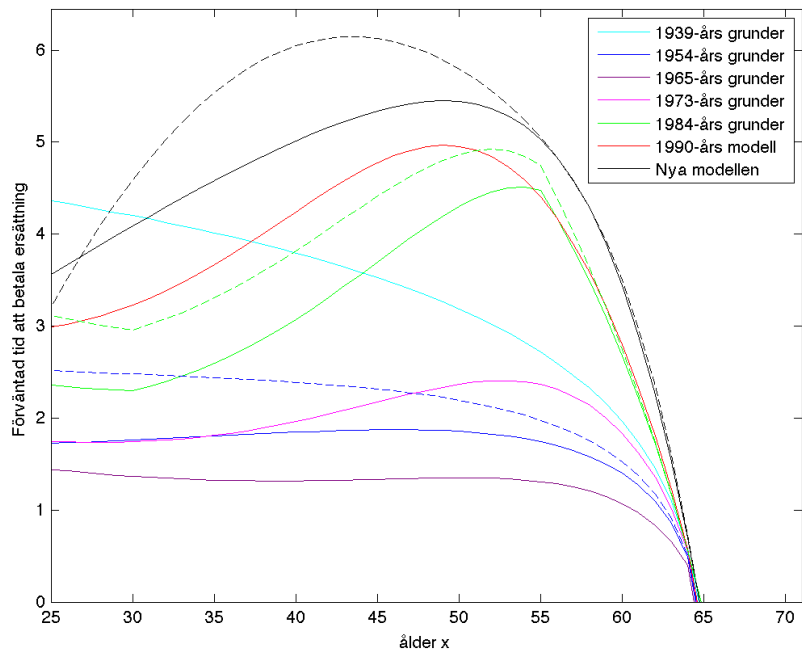
Figur 12 visar hur den förväntade tiden att få sjukersättning ser ut givet att den försäkrade redan varit sjuk ett år. Avvecklingsfunktionen från 1990 ser först lite suspekt ut. Därför gjordes även en numerisk integration och då kunde konstateras att värdena stämde. 1990 års avvecklingsfunktion har en väldigt plan och tjock svans, men det har även de nya framtagna avvecklings-

Figur 10: *Svansen av avvecklingsfunktionen*

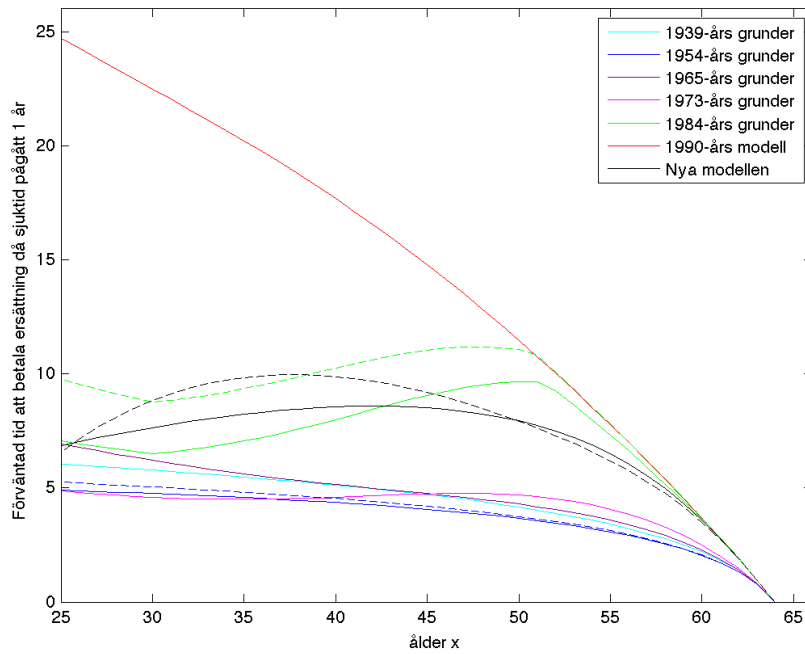
funktionen. Skillnaden är att funktionen från 1990 planar ut mycket tidigare. Kurvan är redan plan efter durationen ett år, medan den nya först planar ut efter 5 år. Slutsatsen är att enligt 1990 års modell förblir en individ oftast sjuk om denne redan varit sjuk i ett år.

Alla modellerna visar en högre förväntad sjuktid i Figur 12 än i Figur 11. Detta är rimligt. Ju längre en person varit sjuk desto mindre blir sannolikheten att han eller hon blir frisk. Efter 1990 års modell har den nya modellen och G84 längst förväntad tid att få sjukpenning. G54 har lägst följt av G73, G39 och G65.

Återigen ser vi samma mönster som i Figur 11, det verkar som om det hänt något efter G73. De som har varit sjuka en längre period verkar komma tillbaka till arbetet i mindre grad än tidigare.



Figur 11: Förväntad tid att betala ersättning då sjukdom pågått i tre månader,  
(-) Män, (- -) Kvinnor



Figur 12: Förväntad tid att betala ersättning då sjukdom pågått i 1 år, (-) Män, (- -) Kvinnor

## 6 Konfidensband med hjälp av bootstrap

Andra delen av detta arbete har bestått av att utifrån Folksams data skatta fram en ny avvecklingsfunktion enligt samma metod som SUS-gruppen. Därefter vill vi se på osäkerheten i modellen, både för avvecklingsfunktionen och förväntad tid att betala ersättning, vilket kommer göras med hjälp av bootstrap-konfidensintervall.

Bootstrap introducerades av Efron första gången 1979. Idén är att genom datorsimuleringar ersätta komplicerade och ofta felaktiga approximationer av varianser och andra osäkerhetsmått. Till en början var det många teoretiker som ifrågasatte denna metod, men senare har det kunnat visas att bootstrap ofta fungerar bättre än de traditionella metoderna.

Den bootstrapteknik som har valts för denna rapport är den icke-parametriska, vars fördel är att den inte utgår från någon fördelning. Tekniken fungerar på så sätt att vi har tillgång till ett stickprov  $\mathbf{z} = (z_1, \dots, z_n)$  från  $n$  stycken individer som är ett utfall av oberoende likafördelade variabler  $Z_1, \dots, Z_n$  med okänd fördelningsfunktion  $F(v)$ . Till varje individ finns  $p$  egenskaper specificerade, vilket betyder att  $z_i = (z_{1i}, \dots, z_{pi})$  och  $Z_i = (Z_{1i}, \dots, Z_{pi})$ .

Vi är intresserade av fördelningen, speciellt spridningen, hos en statistika  $T(\mathbf{z})$ . För att få en uppfattning om denna betraktas stickprovet som en skattning  $\hat{F}_n(v)$  av fördelningsfunktionen  $F(v)$ . Fördelningsfunktionen skattas då med hjälp av den empiriska fördelningsfunktionen

$$\hat{F}_n(v) = \hat{P}_n(Z_i \leq v) = \frac{\sum_{i=1}^n 1\{z_i \leq v\}}{n} = \frac{\sum_{i=1}^n 1\{z_{1i} \leq v_1, \dots, z_{pi} \leq v_p\}}{n} \quad (46)$$

där komponenterna i vektorn  $v = (v_1, \dots, v_p)$  varierar oberoende av varandra mellan  $-\infty$  och  $\infty$ . Från  $\hat{F}_n(v)$  dras nya så kallade bootstrapstickprov oberoende med återläggning av storlek  $n$ , säg  $m$  stycken. Stickproven,  $\mathbf{z}_1^*, \dots, \mathbf{z}_m^*$ , skapas alltså genom att vi drar  $n$  stycken individer från en likformig fördelning på  $1, \dots, n$ . Varje  $\mathbf{z}_j^*$  innehåller då  $n$  stycken  $p$ -dimensionella komponenter. Bootstrapstickprovet  $\mathbf{z}_j^*$  ses som en observation av  $\mathbf{Z}_j^*$ , där  $\mathbf{Z}_1^*, \dots, \mathbf{Z}_m^*$  är oberoende och likafördelade. För varje bootstrapstickprov  $\mathbf{z}_1^*, \dots, \mathbf{z}_m^*$  beräknas sedan  $T(\mathbf{z}_j^*)$  och genom att se på fördelningen för  $T(\mathbf{z}_1^*), \dots, T(\mathbf{z}_m^*)$  fås en uppfattning av fördelningen för statistikan.

För väntevärdet i fördelningen för en given komponent kan ett 95% bootstrap-konfidensintervall definieras till exempel via 2,5% respektive 97,5% percentilerna för denna komponent i  $T(\mathbf{z}_1^*), \dots, T(\mathbf{z}_m^*)$ .

Grunden är med andra ord att varje bootstrapstickprov har samma relation till det givna stickprovet som det givna stickprovet har till populationen. Det ursprungliga stickprovet har bristande information och genom att betrakta det som en population och göra ett stort antal nya stickprov kan något sägas om egenskaperna för stickprovet. Genom detta fås en känsla av spridningen på data.



Denna bootstrapteknik kommer nu att tillämpas på sjukdata, där syftet är att skapa konfidensband för sjuklängderna.

## 6.1 Datamaterialet

Perioden som studerades var 2001-01-01 till och med 2007-12-31. Datamaterialet består av personer med frivillig försäkring som har insjuknat mellan åldrarna 18-65 år. För att kunna göra bra skattningar har data valts på så sätt att ålder vid insjuknande har legat mellan 25-63 år. De tidiga åldrarna togs bort på grund av för lite data och de högre åldrarna på grund av att en hel del av data troligen hade rapporterats in fel. Istället för att rapporterat in att den försäkrade gått i pension hade rapporterats in att denne avvecklats, det vill säga blivit frisk eller avlidit. Sjukfall där sjuktiden varit kortare än nittio dagar tas även bort från ursprungsdata. Detta på grund av att de flesta försäkringar har en karenstid på 90 dagar och därför inte får uppgifter om kortare sjukfall. Skulle data för de kortare sjukfallen tas med kommer det att bli en underskattning av dessa då mycket av data fallit bort.

Rådata har bestått bland annat av insjuknandedatum, födelseår, kön, avvecklingsdatum och avvecklingsorsak. Från detta skapas fyra nya vektorer,  $\mathbf{x}$ ,  $\tilde{\mathbf{d}}$ ,  $\mathbf{y}$ ,  $\tilde{\mathbf{y}}$  (se nedan), med den information som behövs i kommande beräkningar.

Då det endast har rapporterats in födelseår och inte födelsedatum har födelsedatumet satts till första juli. Med hjälp av födelsedatum och insjuknandedatum beräknas

$$x_i = \text{ålder vid insjuknande.} \quad (47)$$

Om en person fortfarande är sjuk efter att observationstiden är slut kan sjuklängden inte beräknas, detta kallas att data är högercensurerat. En person kan även bli avveklad för att försäkringstiden tog slut genom pension, det betyder inte att denna blev frisk, utan också att data är högercensurerat. Låt

$$\tilde{d}_i = \begin{cases} 0 & \text{om observationen är höger censurerad,} \\ 1 & \text{annars.} \end{cases} \quad (48)$$

Den observerade sjuklängden,  $y_i$ , beräknas med hjälp av insjuknandedatum och avvecklingsdatum.

$$y_i = \min(y_i^0, c_i) \quad (49)$$

där  $y_i^0$  är den riktiga sjuklängden och  $c_i$  är det censurerade värdet. Om värdet är högercensurerat vet vi inte  $y_i^0$ .

En del av sjukfallen inträffade innan observationstidens start. Normalt sett är insjuknandedatumet känt trots det, så även i detta fall. Tiden som gått innan observationstidens start är given och har inte utsatts för risk,

$$\tilde{y}_i = \text{sjuktid ej under risk.} \quad (50)$$

Om ett sjukfall inträffar och avvecklas innan observationstiden kommer det inte tas med i uträkningarna, men om sjukfallet avvecklas efter observationstidens start tas det med i beräkningarna. Därav har den sjuktid som passerat innan observationstiden inte utsatts för någon risk att avvecklas. Om hänsyn inte skulle tas till detta skulle de långa sjukfallen överskattas.

Trots att data där insjuknande ålder var över 63 år hade tagits bort, hade SUS-gruppen fortfarande problem med att en del av data verkade ha rapporterats in felaktigt. På samma sätt som ovan verkade en hel del fall rapporterats in som avvecklade då de egentligen var högercensurerade. I och med att denna högercensurering utelämnades blev det en överskattning av avvecklingen för de högre åldrarna. På grund av detta justerades data något genom att låta lutningen på avvecklingskurvan (Kaplan-Meier, (6.4)) få speglas av dödlighetsrisken om

1. durationen är större än 15 år. Ty om en individ varit sjuk i över 15 år är sannolikheten ytterst liten att han/hon kommer bli frisk.
2. insjuknandeålder plus sjuktid är över 63 år ( $x + t > 63$ ). Ty om en individ är äldre än 63 år är sannolikheten mycket liten att avvecklas och borde speglas av dödsrisken.

Studier om dödlighetsrisken för sjuka personer har inte funnits tillgång till. Då denna ändå är ytterst liten kommer den inte ha särskilt stor effekt på avvecklingen och därmed inte ha så stor betydelse för storleken på reserverna. Dödlighetsintensiteten för sjuka personer sattes till något högre än den för friska personer, nämligen till

| För Män          | mätt i promille | För Kvinnor      | mätt i promille |
|------------------|-----------------|------------------|-----------------|
| $x < 55$         | 3,6             | $x < 45$         | 0,96            |
| $55 \leq x < 60$ | 6               | $45 \leq x < 50$ | 1,8             |
| $60 \leq x < 65$ | 9,6             | $50 \leq x < 55$ | 3,24            |
|                  |                 | $55 \leq x < 60$ | 5,16            |
|                  |                 | $60 \leq x < 65$ | 8,04            |

Jämförs det korrigerade data med det okorrigerade ses ingen större skillnad förutom för individer i den äldsta åldersgruppen. Med motiveringen att det är bättre att välja modeller med så lite manipulation av data som möjligt har valts att inte justera med dödlighetsintensiteterna förutom för då åldern är över 63 år.

## 6.2 Simulering med hjälp av icke-parametrisk bootstrap

Ursprungsdata delas upp efter kön. För ett givet kön finns då data  $y_1, \dots, y_n$  där  $y_i$  anger sjuklängden för individ  $i$ . Till varje observation hör också information om ålder vid insjuknande, hurvida observationen är högercensurerad och/eller hur mycket av den observerade sjuktiden som ej utsatts för risk. Det vill säga vi har data på formen  $z_i = (y_i, \tilde{d}_i, \tilde{y}_i, x_i)$ .

För varje kön dras 4 500 stycken bootstrapstickprov fram av storlek  $n$ . Observera att  $n$  är olika för kvinnor och män. Detta görs med hjälp av den empiriska fördelningsfunktionen och är därför väldigt rakt på då den ger lika stor sannolikhet till varje orginaldata värde. Med andra ord görs  $n$  stycken dragningar med återläggning 4 500 gånger.

Till varje bootstrapstickprov skattas nya parametervärden i avvecklingsfunktionen. Hur detta gjorts kan läsas om nedan.

## 6.3 Skattningsteknik för avvecklingsfunktionen

Data, för respektive kön, delas upp i åldersgrupper om fem år efter ålder vid insjuknandet. Åldersgrupp ett innehåller de individer som insjuknat vid ålder 25-29 år, åldersgrupp två 30-34 år och så vidare. Däremot kommer den sista åldersgruppen att innehålla de individer som insjuknat vid ålder 60-63 år. För respektive åldersgrupp har snittåldern,  $x_{(k)}$ ,  $k = 1, \dots, 8$ , beräknats. Det betyder att vi för varje åldersgrupp och kön har tillgång till 4 500 stycken bootstrapstickprov, där storleken hos det  $i$ :te stickprovet ges av det antal observationer i det  $i$ :te återsamlade stickprovet som tillhör den aktuella åldersgruppen för det givna könet.

Parametrarna i avvecklingsfunktionen,  $\theta = \{(a_i, b_i, c_i)_{i=1}^3, (d_i)_{i=1}^4\}$ , tas fram för varje bootstrapstickprov genom att minimera kvadratsumman

$$\sum_{x=x_{(1)}}^{x_{(8)}} \sum_{k=1}^{\tilde{n}_x} (\hat{S}_x(t_k) - \lambda_x(t_k|\theta))^2 \quad (51)$$

där  $\hat{S}_x(t_k)$  är en framtagen Kaplan-Meierkurva för respektive åldersgrupp och  $t_1, \dots, t_{\tilde{n}_x}$  är valda durationslängder med  $t_1 = 0, 25$ . Kaplan-Meier kurvan är en skattning av överlevnadsfunktionen  $S_x = 1 - F_x$  för avvecklingstiden i en åldersklass, där  $F_x$  anger motsvarande fördelningsfunktion. Hur de tas fram kan läsas om i avsnitt 6.4.

Avvecklingsfunktionen är på den form som SUS-gruppen kom fram till, nämligen

$$\lambda_x(t_k) = \sum_{i=1}^4 f_i(x) e^{-d_i(t_k - 0,25)} \quad (52)$$

där

$$f_i(x) = a_i + b_i \cdot e^{c_i \cdot x},$$

$$f_4(x) = 1 - \sum_{i=1}^3 f_i(x).$$

Minimeringsproblemet (51) försökte först lösas enligt ovan beskrivet med hjälp av `fminsearch` i `matlab`, där SUS-gruppens skattningar användes som startvärden och där  $t_k$  sattes till de observerade sjuklängderna  $y_i$ . Det visar sig att det inte riktigt var så enkelt. Problemet är att vi har alldeles för många observationer för de korta sjukfallen. Detta resulterar i att vår avvecklingsfunktion endast stämmer bra till en början. Först angreps problemet genom att lägga in olika vikter, det vill säga ge värden i svansen större betydelse än de i början. `Matlabs fminsearch` är begränsad och denna metod visade sig inte vara den bästa. En bättre metod är att glesa ut de punkter som minimeringen är baserad på genom att se till månader istället för dagar. Durationspunkterna  $t_k$  är med andra ord de observerade sjuklängderna med minst en månads mellanrum. Det framgick även att SUS-gruppen gjort på samma sätt.

#### 6.4 Kaplan-Meier

Avvecklingsfunktionen för respektive åldersklass och kön med Kaplan-Meier ser ut enligt följande

$$\widehat{S}_x(t_k) = \prod_{j=1}^k \frac{n_j - d_j}{n_j} \quad (53)$$

där

$$n_j = \text{antalet under risk vid durationslängd } t_j \text{ och}$$

$$d_j = \text{antalet som avvecklas i intervallet } ]t_{j-1}; t_j].$$

Kaplan-Meier-skattningen brukar normalt gå över de observerade tidpunkterna, se Lee (1992). I och med minimeringsproblemet som diskuterades i föregående stycke är dessa observerade sjuklängder utglesade. Då  $t_j$  är de observerade sjuklängderna, med minst en månads mellanrum, betyder det att flera observationer  $y_i$  kan ske i intervallet  $]t_{j-1}; t_j]$ . För de längre durationslängderna kommer ofta  $t_{j-1} - t_j > 0, 25$ .

För varje åldersgrupp anpassas Kaplan-Meier kurvor. Där tas hänsyn till högercensurering och/eller om en del av sjuktiden inte har varit under risk.

Högercensureringen tas hänsyn i till (53) genom  $d_j$ . Endast de observationer som avvecklas i tidsintervallet  $]t_{j-1}; t_j]$  summeras, vilket betyder att de

högercensurerade sjukfallen inte kommer öka storleken på  $d_j$ . I matematiska termer beräknas antalet som avvecklas i intervallet  $[t_{j-1}; t_j]$  enligt

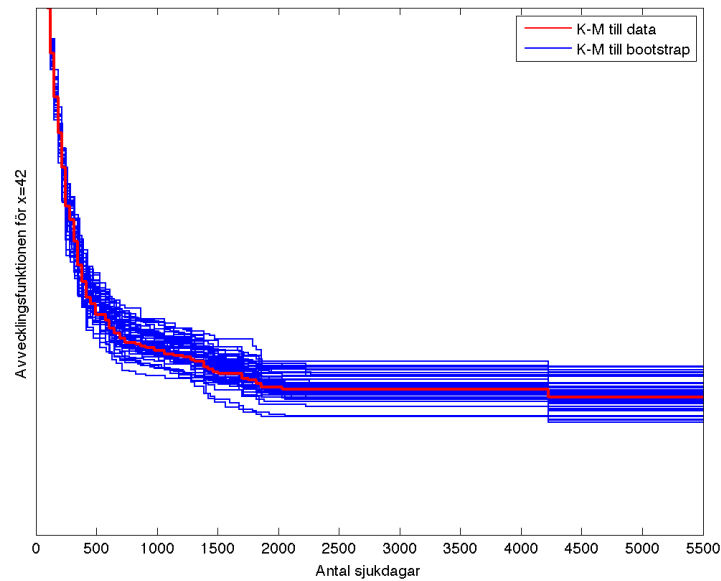
$$d_j = \sum_{y_i \in [t_{j-1}; t_j]} \tilde{d}_i$$

där  $\tilde{d}_i$  är definierad som i (48). Om fallet skulle vara så att det endast inträffar högercensurerade sjukfall under ett intervall  $[t_{j-1}; t_j]$  skulle de ändå bidra till minimeringen genom att det tillkommer en observationspunkt  $t_k$ .

Sjuktid ej under risk tas hänsyn till i (53) genom  $n_j$ , denna beräknas enligt

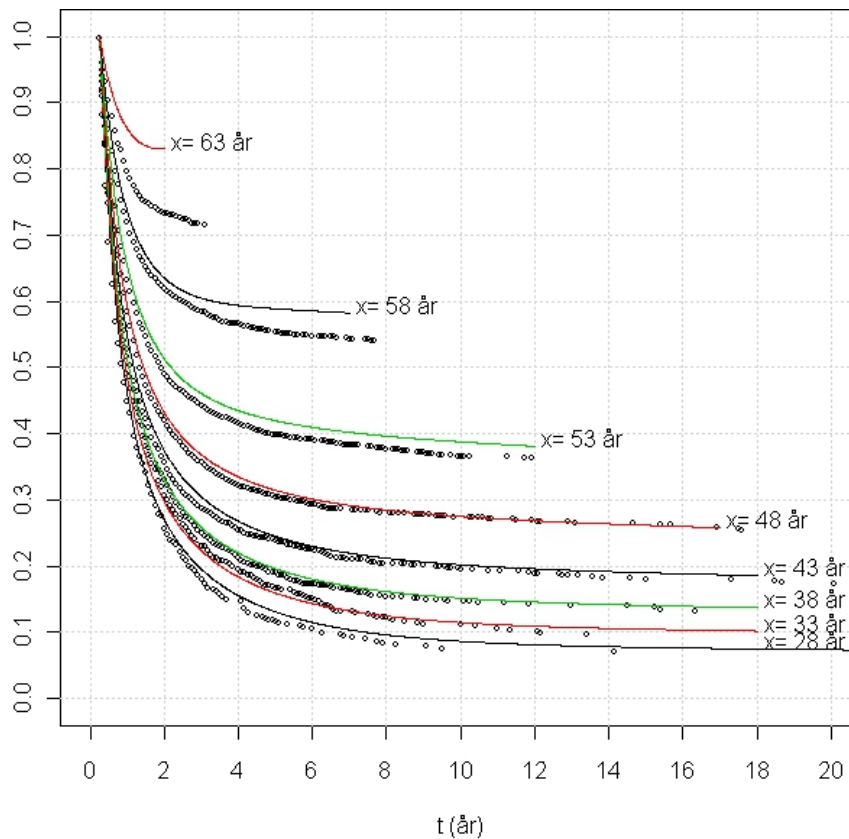
$$\begin{aligned} n_j &= \text{antal sjukfall som inte avvecklats vid durationslängden } t_j \\ &\quad - \text{antal sjukfall som inte är under risk vid durationslängden } t_j \\ &= \sum_{y_i > t_j} I(\tilde{y}_i \leq t_j) \end{aligned}$$

där  $y_i$  och  $\tilde{y}_i$  är definierade som ovan (49, 50).



Figur 13: *Kaplan-Meier* för data och 50 bootstrapstickprov för män i åldersgruppen 40-44 år.

I Figur 13 ser vi hur de olika Kaplan-Meierkurvorna ser ut till de framsimulerade bootstrapstickproven. En jämn spridning runt data kan observeras.



Figur 14: *SUS-gruppens skattningar och Kaplan-Meier för samtliga data, allt för delbeståndet frivillig man.*

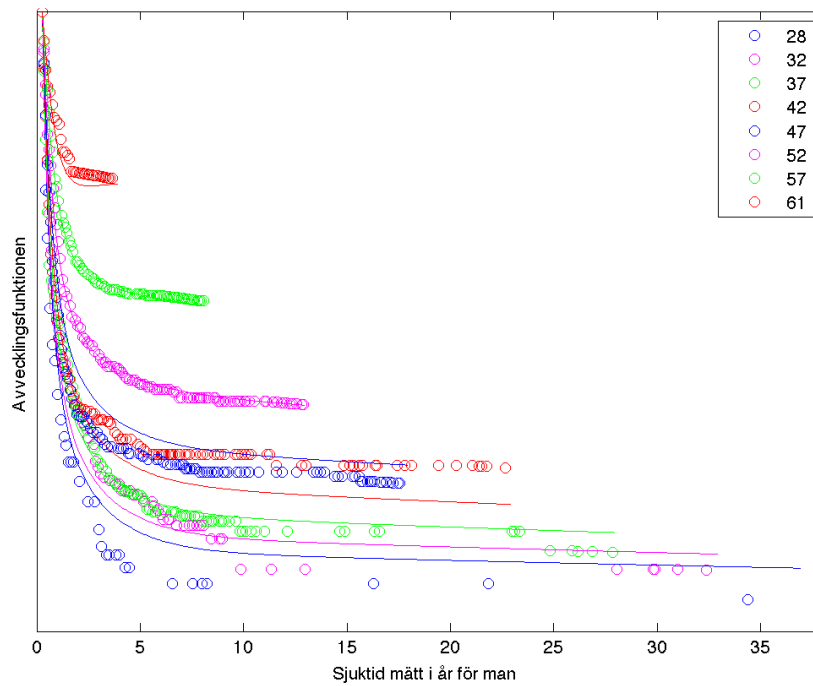
## 6.5 Resultat

### 6.5.1 Jämförelse av datamaterialet

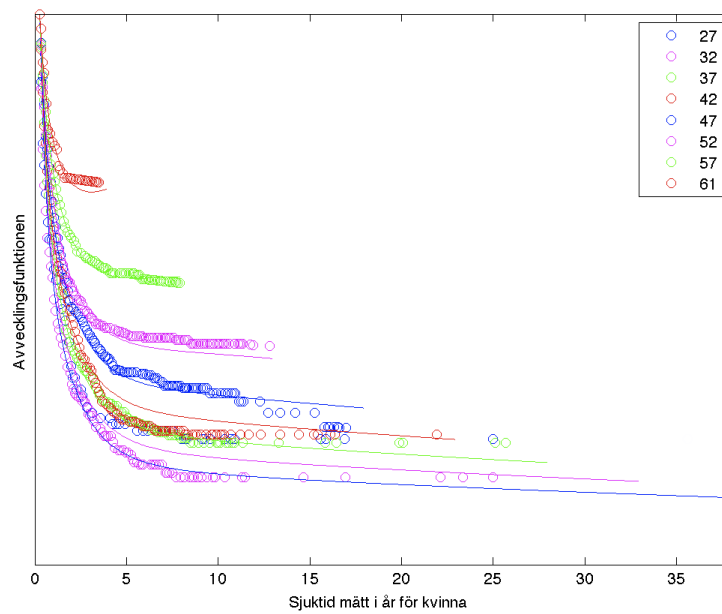
SUS-gruppens framtagna avvecklingsfunktion för män visas i Figur 14 mot de justerade Kaplan-Meierkurvorna för alla bolags data. Anpassningen är bra, men inte perfekt. Liknande figur fås för kvinnor med samma goda anpassning.

Folksams Kaplan-Meier kurvor jämförs i Figurerna 15 och 16 mot de anpassade modellkurvorna  $\lambda_x(t)$  för respektive kön.

Den anpassade modellen för män följer Kaplan-Meierkurvorna väl i de tre äldsta åldersgrupperna, se Figur 15, och även i den åldersgruppen då insjuknande ålder i medel är 37 år. I övrigt är data inte lika välorganiserat. Kaplan-Meierkurvorna visar att de män som insjuknar runt åldern 42 år är kvar som sjuka i högre grad än män som insjuknar runt åldern 47 år. Detta samband kan inte observeras i Kaplan-Meierkurvorna för alla bolags data, se Figur 14. Att det föreligger på detta sätt beror troligen på för lite data. Den



Figur 15: Kaplan-Meier kurvor för män med anpassade modellkurvor, Folksam data.



Figur 16: Kaplan-Meier kurvor för kvinnor med anpassade modellkurvor, Folksam data.

anpassade modellen kan därför inte komma att följa Kaplan-Meier kurvorna lika väl som vi såg i Figur 14.

I männens lägsta åldersgrupp passar modellkurvorna inte särskilt bra. SUS-gruppens framtagna avvecklingsfunktion,  $\lambda_{28}(t)$ , passar i detta fall bättre till Folksams data. Anledningen är troligen att det är för få observationspunkter i denna åldersgrupp, vilket ger för liten vikt till den lägsta åldersgruppen i minimeringen (51). Skulle Folksam använda denna avvecklingsfunktion vid avsättning av reserver skulle reserven för den lägsta åldersklassen bli för stor. Då det inte är så många individer som insjuknar vid låga åldrar jämfört mot de högre åldrarna, blir effekten inte lika stor som om en felskattning hade gjorts där fler individer insjuknat. Nackdelen är dock att de som insjuknar vid låga åldrar kan vara sjuka en längre tid än de som insjuknar vid högre åldrar. Eftersom reserven för den fjärde åldersklassen skulle bli för liten, då den anpassade  $\lambda_{42}(t)$  används, jämnas skillnaderna ut något.

För kvinnor som insjuknar runt åldern 27 år kan observeras i Kaplan-Meierkurvorna, se Figur 16, att en större andel tenderar att förbli sjuka än de som insjuknar runt åldern 32 år. Dessutom ligger andelen som fortfarande är sjuka efter 5 år och framåt nästan lika oavsett om en kvinna insjuknar vid åldern 27, 37 eller 42 år. Förmodligen beror detta återigen på att datamängden är för liten. Modellanpassningen blir helt klart sämre här beroende på utseendet av data.

Vi har i Figurerna 15 och 16 sett att data inte passar lika väl till modellen som alla bolags data passar till den allmänna modellen. Om Folksams data jämförs dels mot sina anpassade modellkurvor, och dels mot modellkurvorna anpassade till hela datamaterialet, kan det observeras att den förstnämnda anpassningen är bättre, konstigt vore det annars.

### 6.5.2 Konfidensband

Konfidensband för avvecklingsfunktionen tas fram genom att för varje givet  $t$  se på 2,5% respektive 97,5% percentilerna i den uppsättning skattade avvecklingsfunktioner som finns till hands. På samma sätt görs för förväntad tid att betala ersättning.

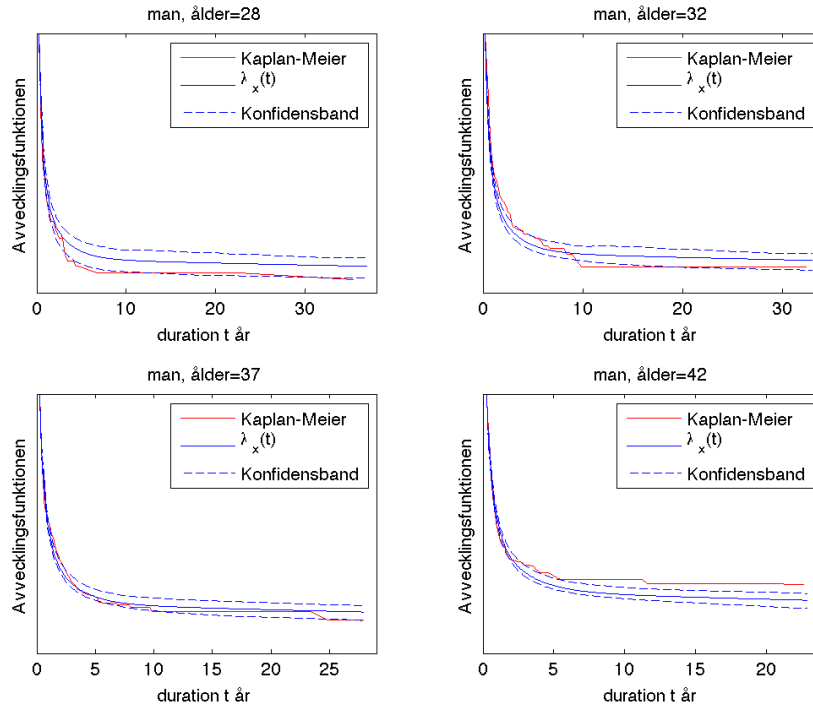
Den allmänna modellen, det vill säga avvecklingsfunktionen som är framtagen till alla bolags data, kommer inte att visas i samma figur som Folksams modell. Detta för att skydda Folksams data. Modellerna liknar varandra, men det förekommer att den allmänna modellen ligger utanför Folksams 95%-iga bootstrap-konfidensintervall.

#### Män

Antal observationer i Folksams ursprungsdata var för män 2 576 stycken. Delas dessa observationerna  $y_i$  upp i respektive åldersklass fås 82, 211, 338, 372, 372, 445, 515 och 241 antal observationer. Minimeringen (51) av data



görs över, enligt ovan beskrivet, valda observationspunkter  $t_k$ . Antalet  $t_k$  i respektive åldersklass, som benämns som  $\tilde{n}_x$  i (51), är 34, 62, 90, 98, 106, 117, 96 och 58. Minimeringen till Folksams ursprungsdata sker då till 661 stycken valda observationspunkter.



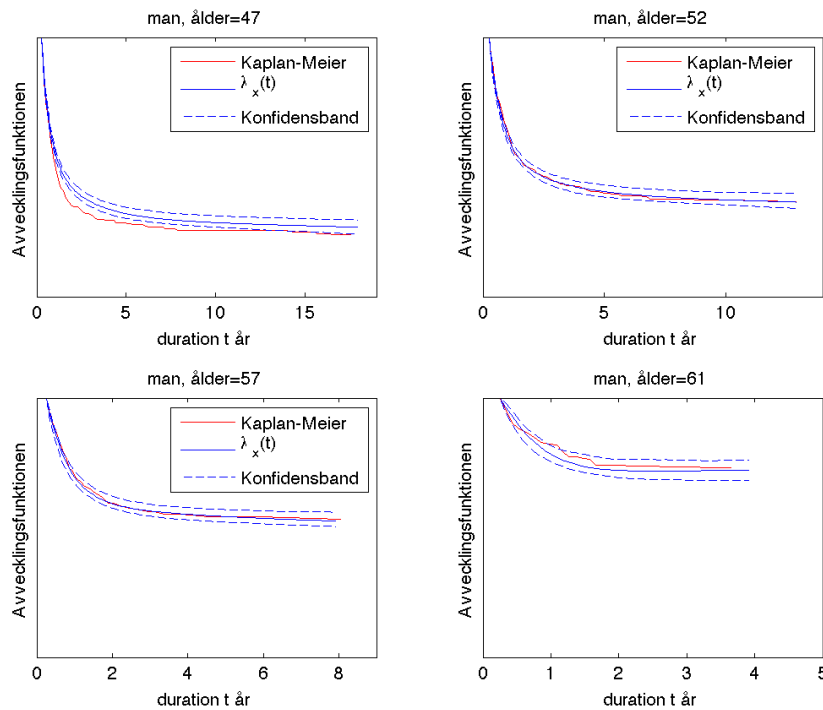
Figur 17: *Avvecklingsfunktionen för män med konfidensband på 95%-nivån*

Då antal observationer är färre i de lägre åldersklasserna, se Figur 15, ser vi som väntat ett bredare konfidensband för avvecklingsfunktionen i dessa åldrar, se Figur 17. Avvecklingsfunktionen, då insjuknande ålder är 28 år, underskattar den avveckling som data visar. Det är därför inte konstigt att konfidensbandet är brett för de högre avvecklingstakterna. Konfidensbandet visar även en stor spridning för de lägre avvecklingstakterna. Detta skulle kunna bero på att spridningen på data visar att många sjukskrivna avvecklas i början av sjukperioden. En annan orsak skulle även kunna vara att punkterna som minimeringen sker till i den yngsta åldersklassen är för få och har för liten inverkan på modellens utseende.

När ålder vid insjuknande är 32 år kan observeras att modellen över-skattar avvecklingen något de första 10 åren, därefter underskattar modellen tillfrisknandet. Kaplan-Meierkurvan ligger dock mestadels innanför konfidensbandet.

Avvecklingsfunktionen för insjuknande ålder 42 år följer Kaplan-Meierkurvan väl, konfidensbandet runt kurvan börjar smalna av jämfört mot de yngre åldersklasserna.

I Figur 15 kunde ses hur Kaplan-Meierkurvorna för åldersklass fyra och fem nästan låg på varandra. Detta är inte möjligt i modellen på grund av dess utformning och framtagande. Därför är det inte oväntat att Kaplan-Meierkurvorna inte ligger innanför konfidensbandet i dessa åldersklasser, se Figur 17 och 18. Om modellenpassningen och konfidensbanden hade tagits fram för alla bolags data hade anpassningen med stor sannolikhet sett bättre ut.



Figur 18: *Avvecklingsfunktionen för män med konfidensband på 95%-nivån*

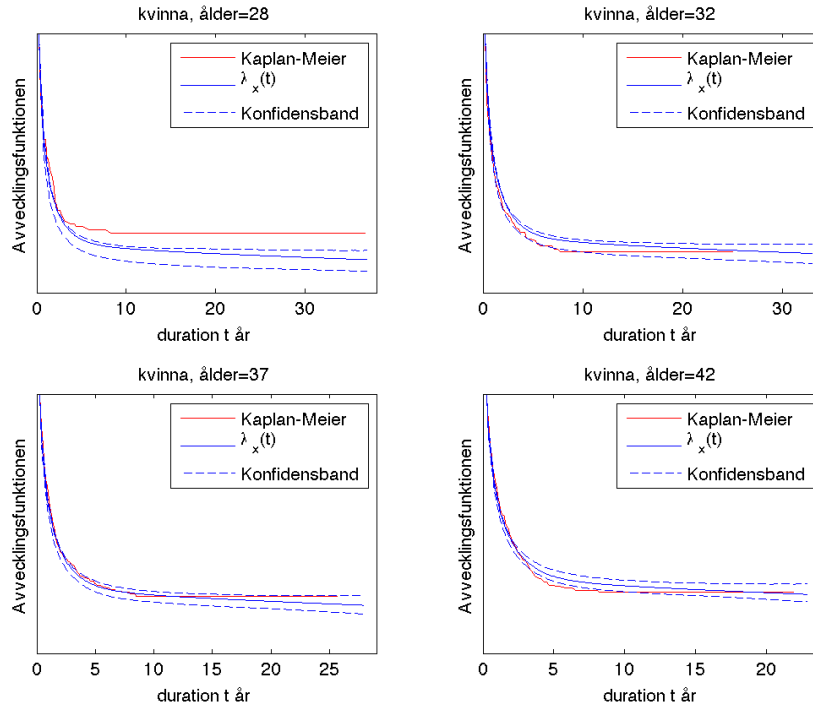
Avvecklingsfunktionen för då ålder vid insjuknande är 52 år eller 57 år följer Kaplan-Meierkurvorna väl. Konfidensbandet är smalt men något bredare uppåt, det vill säga spridningen på data tyder på att modellen snarare överskattar avvecklingen än underskattar.

I den äldsta åldersgruppen ser vi att Kaplan-Meierkurvan ligger praktiskt taget innanför konfidensbanden. Att konfidensbandet är bredare än tidigare kan bero på att antal observationer är färre än i de yngre åldersklasserna, förutom i de yngsta åldersklasserna där antal observationer är minst.

### Kvinnor

Antal observationer  $y_i$  för kvinnor i varje åldersklass var för Folksam sprungsdata 118, 374, 475, 512, 570, 579, 435 och 169. Det vill säga 3 232 observationer där ålder vid insjuknande låg mellan 25 till 63 år. Antal  $t_k$ , som minimeringen (51) sker till för originaldata, blev för respektive ålders-

grupp 50, 82, 94, 95, 116, 109, 89 och 41. Med andra ord sammanlagt 675 stycken valda durationslängder.

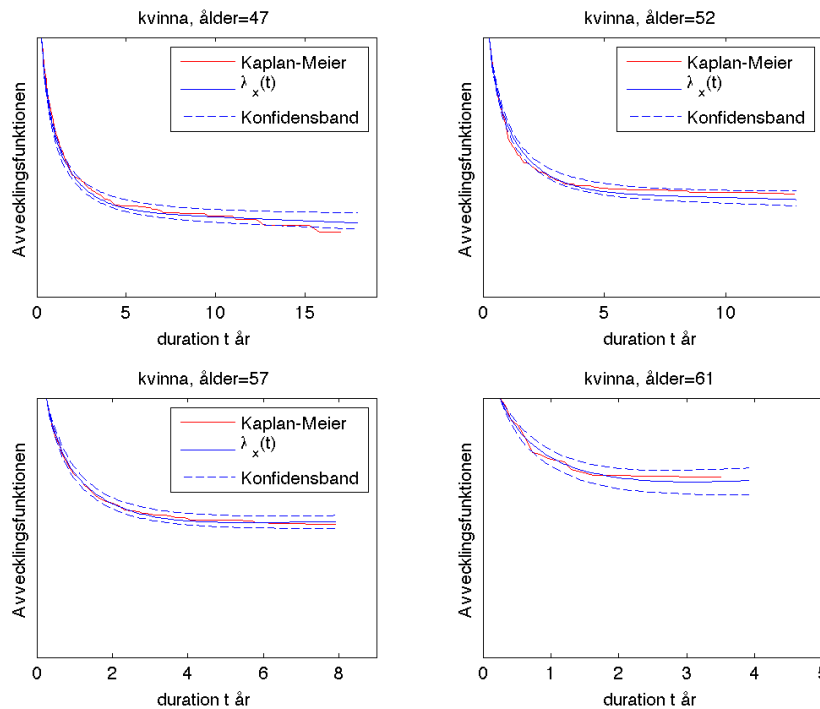


Figur 19: *Avvecklingsfunktionen för kvinnor med konfidensband på 95% nivå*

Redan i diskussionen kring Figur 16 kunde vi ana att det skulle bli svårt för data att ligga innanför konfidensbanden i de lägre åldersklasserna. I den lägsta åldersklassen ses detta tydligt. Även om Folksams data ligger över konfidensbandet så ligger den allmänna modellen, det vill säga modellen till alla bolags data, under konfidensbandet. Att Kaplan-Meierkurvan ser ut på detta sätt kan ha att göra med att datamängden är för liten. Hade vi haft fler observationer kanske det sett annorlunda ut.

Då ålder vid insjuknande är 32 år underskattar modellen avvecklingen något, jämfört mot Kaplan-Meier i åldersgrupp två, men kommer i kapp på slutet. Konfidensbandet visar att spridningen på data främst har en högre avvecklingstakt. Då den anpassade modellen underskattar Kaplan-Meierkurvan verkar detta rimligt.

Vidare kan vi se i de äldre åldersklasserna att Kaplan-Meierkurvorna praktiskt taget ligger inom konfidensbandet. Precis som för männen kan ett brett konfidensband ses då insjuknande ålder är 61 år. Återigen beror det troligtvis på att observationspunkterna här blir färre och har mindre inverkan på modellen.



Figur 20: *Avvecklingsfunktionen för kvinnor med konfidensband på 95%-nivån*

### Jämförelse

Vid jämförelse av de olika könen Kaplan-Meierkurvor kan det observeras att männen i de två övre åldersklasserna tenderar att ha större andel kvar som sjuka, se Figur 21. För den äldsta åldersklassen kan denna skillnad ses nästan omedelbart efter karenstiden och för de som insjuknar runt 57 års ålder uppstår skillnaden efter runt två och ett halvt års sjuktid.

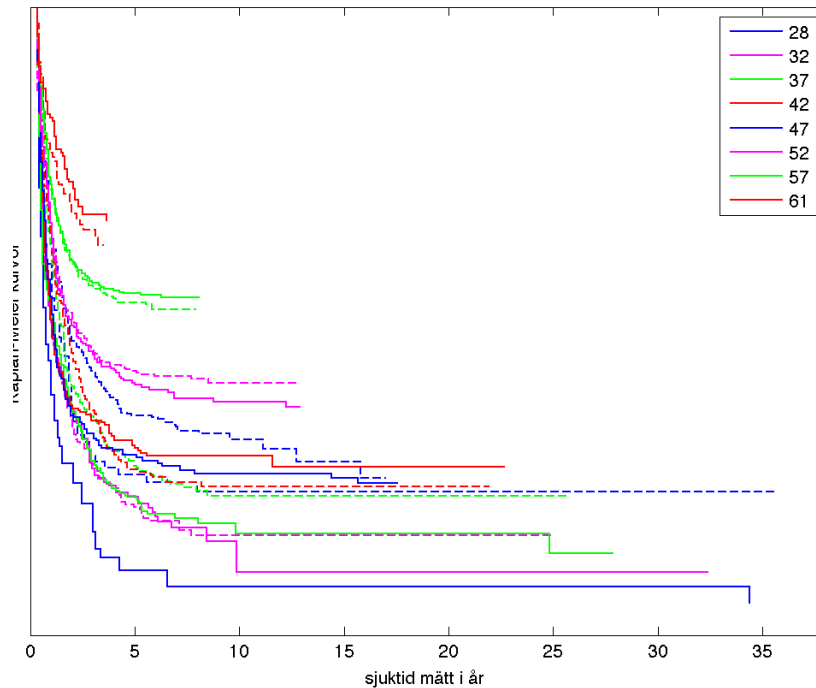
Kring insjuknande ålder 52 och 47 år tenderar däremot kvinnorna att ha större andel sjuka kvar. För de som insjuknar runt åldern 52 år kan skillnaden ses efter drygt 2 års durationstid, medan för dem som insjuknar runt 47 års ålder ses skillnaden betydligt tidigare än så.

Den största skillnaden mellan könen kan observeras i den lägsta åldersklassen. Bland de individer som insjuknar runt ålder 28 år kan vi se i Folksam data att kvinnor kvarstår som sjuka i mycket högre grad än männen. Ungefär två procentenheter skiljer Kaplan-Meierkurvorna åt.

Liknande mönster kan observeras bland de som insjuknar runt åldern 32 år, men här är skillnaderna mer än hälften så stora och den tydliga skillnaden ses inte förrän efter en sjuktid på drygt 10 år.

Även då insjuknande ålder är runt 37 år tenderar kvinnor att kvarstå som sjuka längre tid än män. Skillnaderna börjar redan vid låga durationstider.

Från modellen sett ligger de största skillnaderna mellan män och kvin-



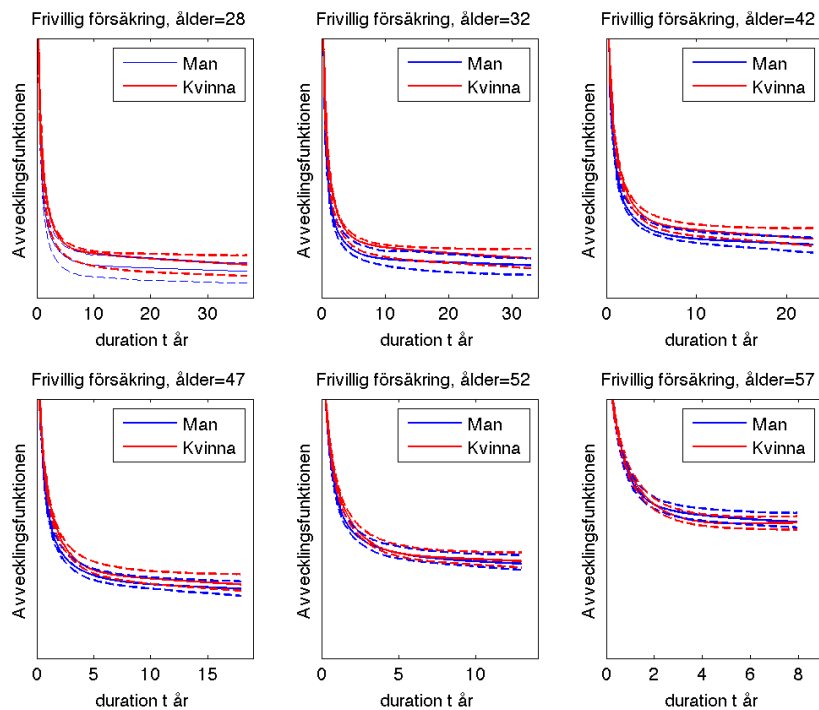
Figur 21: Ej justerade Kaplan-Meierkurvor, (-) Män, (- -) Kvinnor

nor i de lägre åldrarna, se Figur 22. Kvinnor är kvar som sjuka i högre grad än männen. Då ålder vid insjuknande ökar kan vi se i konfidensbanden att skillnaderna mellan könen minskar. Runt insjuknande ålder 52 år ligger konfidensbanden nästa på varandra. Därefter har männen större andel kvar som sjuka.

Ett brett konfidensband för förväntad tid att få sjukersättning kan observeras för både män och kvinnor i de låga insjuknande åldrarna, se Figur 23. Med åldern smalnar konfidensbandet av, detta har att göra med hur väntevärdet räknas ut. Ersättning kan max erhållas upp till 65 år, vilket gör att osäkerheten blir allt mindre med åldern.

I Figuren kan vi även se stora skillnader mellan män och kvinnor. Mäns konfidensband ligger dock alltid över kvinnornas undre konfidensband, vilket tyder på att de inte är statistiskt signifikant skilda. Eftersom vi ser på varje kurva för sig så kan vi inte uttala oss säkert.

Slutsatsen är, att även om modellerna för män och kvinnor skiljer sig åt, visar osäkerheten i modellerna att de ändå inte är så olika. Osäkerhetsanalysen visar även att skillnaderna mellan könen ser ut att minska med insjuknande ålder.

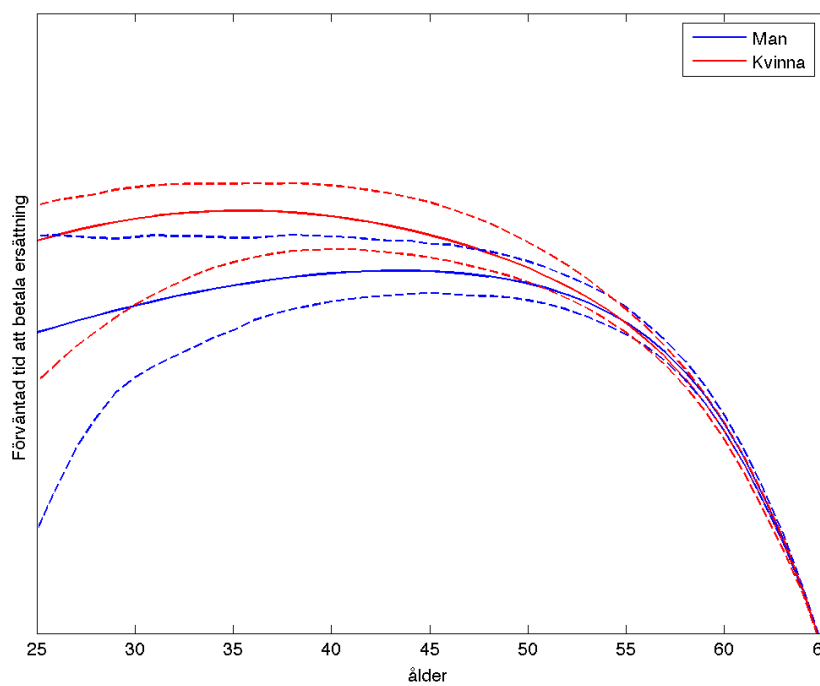


Figur 22: *Avecklingsfunktionen med konfidsband på 95%-nivån*

## 6.6 Kommentarer

Bootstrap skulle kunna användas på fler sätt i denna uppsats. Första tanken var att med hjälp av parametrisk bootstrap bilda konfidsband för de olika grundsystemen och se på skillnaderna. Då vi inte vet hur väl modellerna stämmer och vilka bestånd som använts så är det ett grovt att anta att modellerna är sanna. Då gammalt data inte finns har icke-parametrisk bootstrap endast kunna göras för den nya framtagna modellen. Det hade varit intressant att göra osäkerhetsanalysen för hela datamängden, det vill säga alla bolags data. Tillgång fanns dock endast till Folksams data och då denna metod tar mycket datorkraft och datortid hade det troligtvis behövts en längre period för att ta fram bootstrap-konfidsintervallen för hela datamängden.

Icke-parametrisk bootstrap skulle även kunna ha gjorts med hjälp av betingning på de olika åldersklasserna. Då dras nya bootstrapstickprov fram för respektive kön och åldersklass, och därmed blir antalet observationer i varje åldersklass fixa för alla återsamplings. Denna betingade variant av icke-parametrisk bootstrap ger något smalare konfidsband, eftersom en slumpkomponent eliminerats vid återsamplingen. Den återsamlingsmetod som använts i denna uppsats kan därför ses som ett mer konservativt tillvägagångssätt, som i någon mån kompenserar eventuella systematiska fel hos



Figur 23: Förväntad tid att betala ersättning med konfidensband på 95% nivå

konfidensbanden på grund av otillräcklig anpassning hos den parametriska modellen.

## 7 Diskussion och slutsats

Skillnader mellan mäns och kvinnors insjuknande har kunnat observeras, både i de gamla modellerna och i det nya data. Kvinnor har högre risk att insjukna. Efter 1939 års grunder har skillnaderna stadigt minskat för att sedan öka någon gång på 1990-talet eller 2000-talet. Om detta speglar verkligheten är svårt att säga, modellerna har i flera fall haft dåligt underlag för kvinnor. Dessutom går det inte att utesluta att de observerade skillnaderna även kunnat komma från det yrke kvinnor i största del arbetat inom och/eller hur arbetsfördelningen i samhället har sett ut.

De yrkesverksamma kvinnorna arbetade förr ofta inom fysiska yrken, vilket borde ha ökat risken för sjukskrivningar. Kanske kommer vi att se, i takt med att samhället blir mer jämställt, hur skillnaderna mellan kvinnors och mäns insjuknande minskar. Frågan är dock varför skillnaderna ökade igen på 1990- och 2000-talet. En faktor till ökningen skulle kunna vara att arbetsfördelningen i hemmet hamnat efter utvecklingen i arbetslivet. Även om kvinnor har fått mer jämställda yrken, kan de ha fått ta lika stor del av arbetet i hemmet som tidigare. Det kan då ha resulterat i att kvinnor haft en högre arbetsbörda än män. Att detta skulle haft så stor inverkan som vi har sett i ökningen nu är dock inte troligt.

Samhällsklimatet kan även, på olika sätt, ha varit en bidragande faktor till hur skillnaderna mellan män och kvinnor har speglats. Ekhult kommenterade i 1965 års grunder att den multiplikativa konstanten i insjuknandeintensiteten hade sänkts, trots att data inte visade en minskad skillnad mellan könen. Sänkningen motiverades med att priserna (premierna) inte skulle vara för höga för att kvinnor skulle kunna köpa sjukförsäkringar.

Avvecklingsfunktionen mäter sannolikheten att kvarstå som sjuk. I jämförelsen av de olika modellernas avvecklingsfunktioner och förväntad tid att erhålla sjukersättning har skillnader mellan dem observerats. Personer som varit sjuka en längre tid har förblivit sjuka i allt högre grad sedan 1973 års grunder. Ju längre sjuktid som passerat, desto mindre har sannolikheten att komma tillbaka i arbete blivit. Även den förväntade tiden att erhålla sjukersättning har blivit allt längre sedan 1973 års grunder.

Vissa skillnader i de olika grundsystemens avvecklingsfunktioner kan ha berott på olika beståndssammansättningar. Om ett bestånd som består av en stor andel arbetare jämförs mot ett bestånd med en stor andel tjänstemän har troligen de förstnämnda en större sannolikhet att kvarstå som sjuka. Till exempel, om en individ som arbetat inom ett fysiskt arbete fått en ryggskada borde rimligtvis sannolikheten att börja arbeta igen vara mindre än om han eller hon haft ett kontorsjobb.

En ny avvecklingsfunktion skattades fram i detta arbete, för respektive kön, till Folksamns moderna data. Skillnader kunde då observeras. Förväntad sjuktid var högre för kvinnor än för män, förutom då ålder att insjukna var hög. Vid osäkerhetsanalysen för avvecklingsfunktionerna kunde ses i 6.5.2



att bootstrap-konfidensbanden för kvinnor och män blev alltmer lika med insjuknandeålder. Det vill säga, skillnaderna mellan kvinnors och mäns sannolikhet att kvarstå som sjuka minskade med åldern vid insjuknandet. Bland de som insjuknat vid höga åldrar har männen till och med haft större andel sjuka kvar efter en durationstid runt två år.

En fråga som kan ställas är om vi har tillräckligt många observationer för att kunna lita på konfidensbanden. Ett framtida projekt skulle kunna vara att återsampla alla bolags data, det vill säga ge konfidensband för de skattningar som tagits fram av SUS-gruppen.

## 8 Appendix

### 8.1 Härledning av $\kappa_{[x]+t} = -\frac{d}{dt} \ln \lambda_{[x]+t}$

Kom ihåg följande definitioner:

$$\kappa_{[x]+t} = \lim_{dt \rightarrow 0} \frac{P(T_x \leq t + dt | T > t)}{dt} \quad (54)$$

$$F_x(t) = P(T_x \leq t) \quad (55)$$

$$\lambda_{[x]+t} = P(T_x > t) = 1 - F_x(t) \quad (56)$$

$$f_x(t) = -\frac{d}{dt} \lambda_{[x]+t} = -\lambda'_{[x]+t} \quad (57)$$

$$\begin{aligned} \kappa_{[x]+t} &= \lim_{dt \rightarrow 0} \frac{\frac{P(T_x \leq t+dt, T_x > t)}{dt}}{P(T_x > t)} = \{55, 56\} = \lim_{dt \rightarrow 0} \frac{\frac{F_x(t+dt) - F_x(t)}{dt}}{\lambda_{[x]+t}} = \frac{f_x(t)}{\lambda_{[x]+t}} \\ &= \{57\} = -\frac{\lambda'_{[x]+t}}{\lambda_{[x]+t}} = -\frac{d}{dt} \ln \lambda_{[x]+t} \end{aligned}$$

### 8.2 Härledning av $\lambda_{[x]+t}^{(k)} = \frac{\lambda_{[x]+t+k}}{\lambda_{[x]+k}}$

Observera först att  $\lambda_{[x]+t}^{(k)}$  är avvecklingsfunktionen givet att man varit sjuk under karenstiden  $k$  år,  $t$  är tiden efter karenstiden. Funktionen  $\lambda_{[x]+t}$  är den vanliga avvecklingsfunktionen där  $t$  är tiden efter insjuknandet. Variabeln  $T_x$  står för den totala sjuktiden, inklusive karenstiden.

Om  $t$  är fix gäller

$$\begin{aligned} \lambda_{[x]+t}^{(k)} &= P(T_x > t + k | T_x > k) = \frac{P(T_x > t + k, T_x > k)}{P(T_x > k)} \\ &= \frac{P(T_x > t + k)}{P(T_x > k)} = \frac{\lambda_{[x]+t+k}}{\lambda_{[x]+k}}. \end{aligned}$$

Om karenstiden är tre månader fås

$$\lambda_{[x]+t}^{(k)} = \frac{\lambda_{[x]+t+0,25}}{\lambda_{[x]+0,25}}.$$

## Referenser

- [1] -, *1955 års grunder för livförsäkring m.m.: grunder och formler*, Ivar Hæggströms boktryckeri AB, Stockholm, Sverige.
- [2] -, *Grunder och formler för liv- och sjukförsäkring 1964-/65*, Ivar Hæggströms boktryckeri AB, Stockholm, Sverige.
- [3] ALM, E., ANDERSSON, G., VON BAHR, B., MARTIN-LÖF, A., (2006), *Livförsäkringsmatematik II*, Svenska Försäkringsföreningen, Stockholm, Sverige.
- [4] ANDERSSON, G., (2005), *Livförsäkringsmatematik*, Svenska Försäkringsföreningen, Stockholm, Sverige.
- [5] CRAMÉR, H., FORSMAN, B., JANSSON, T. (1938), *1939 års grunder: utredning och förslag rörande grunder för liv- och pensionsförsäkring*, Stockholm, Sverige.
- [6] DAVISON, A.C., HINKLEY, D.V., (1997), *Bootstrap Methods and their Application*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, England.
- [7] DILLNER, C-G., (1969), New bases for non-cancel lable sicknes insurance in Sweden, *Skandinavisk Aktuarietidskrift*, häfte 3-4, 113-124, Stockholm, Sverige.
- [8] DILLNER, C-G., (1974), New bases for non-cancel lable sicknes insurance in Sweden from 1973, *Scandinavian Actuarial Journal*, No 3, 167-173, Stockholm, Sverige.
- [9] EKHULT, H., (1980), Technique and experience in sickness and disability insurance in Sweden, *Transaction of the 21st International Congress of Actuaries*, No 3, 67-82, Zürich-Lausanne.
- [10] FÖRSÄKRINGSFÖRBUNDET (2002), *Sjuklighetsundersökningen 2002*, Försäkringsförbundet, Stockholm, Sverige.
- [11] HULTMAN, K., (1958), *Livförsäkringsteknik - enligt svenska förhållanden*, Institutet för Försäkringsutbildnings Studiehandböcker, Ivar Hæggströms boktryckeri AB, Stockholm, Sverige.
- [12] KÄLLSTRÖM, L., M.FL., (1990) *Sverige Reinsurance Company 75 years*, Long-term Individual Sickness Insurance According to 1984 Rules (G84), Uppsala, Sverige.
- [13] LEE, E.,T., (1992) *Statistical methods for survival data analysis*, John Wiley & Sons Inc, Canada.

- [14] MATTSSON, P., (1956), New bases for non-cancel lable sicknes insurance in Sweden, *Skandinavisk Aktuarietidskrift*, häfte 3-4, 198-215, Stockholm, Sverige.
- [15] MATTSSON, P., (1977), *Sjukförsäkring i Sverige: Historik, teknik och utvecklingslinjer*, Gotab, Stockholm, Sverige.
- [16] SKANDIA LIV, (1975), 1974 års grunder för individuell sjuk- och olycksfallsförsäkring, VIII B, Resol 1975-06-05, Stockholm, Sverige. (Ej publicerad)
- [17] SKANDIA LIV, (1984), 1984 års grunder för individuell sjuk- och olycksfallsförsäkring, VB, Resol 1984-12-18, Stockholm, Sverige. (Ej publicerad)
- [18] SVENSKA AKTUARIEFÖRENINGEN, (1984), *protokoll 1981-1983*, Protokoll från 1983-11-22, Stockholm, Sverige.
- [19] <http://www.forsakringskassan.se>, (2010) Försäkringskassans hemsida angående sjukpenning, April 2010
- [20] <http://sv.wikipedia.org/>, (2010) angående ITP, April 2010

## 9 BILAGA 1 - Överlevnadsfunktionerna

$$l_x = e^{-\int_0^\infty \mu_s ds} \quad (58)$$

39-års grunder:

$$l_x^{\text{man}} = \exp\left\{-\frac{1}{1000}\left(1,5x + \frac{0,041}{0,042 \cdot \ln 10}(10^{0,042x} - 1)\right)\right\} \quad (59)$$

$$l_x^{\text{kvinnna}} = \exp\left\{-\frac{1}{1000}\left(1,5x + \frac{0,041}{0,042 \cdot \ln 10}10^{-0,042}(10^{0,042x} - 1)\right)\right\} \quad (60)$$

55-års grunder:

$$l_x = \exp\left\{-\frac{1}{1000}\left(1,5x + \frac{0,041}{0,042 \cdot \ln 10}(10^{0,042x} - 1)\right)\right\} \quad (61)$$

65-års grunder:

$$l_x^{\text{man}} = \exp\left\{-\frac{1}{1000}\left(0,6x + \frac{0,034}{0,042 \cdot \ln 10}(10^{0,042x} - 1)\right)\right\} \quad (62)$$

$$l_x^{\text{kvinnna}} = \exp\left\{-\frac{1}{1000}\left(0,6x + \frac{0,034}{0,042 \cdot \ln 10}10^{-0,042 \cdot 4}(10^{0,042x} - 1)\right)\right\} \quad (63)$$

73-års grunder:

$$l_x = \exp\left\{-\frac{1}{1000}\left(0,6x + \frac{0,034}{0,042 \cdot \ln 10}(10^{0,042x} - 1)\right)\right\} \quad (64)$$

## 10 BILAGA 2 - Observerade t-frekvenser 2002

&gt;=1 månads duration:

| Kön | Karens | 25-29   | 30-34   | 35-39   | 40-44   | 45-49   | 50-54   | 55-59   | 60-     |
|-----|--------|---------|---------|---------|---------|---------|---------|---------|---------|
| M   | 1      | 10,994  | 11,831  | 14,0088 | 12,7445 | 13,295  | 17,047  | 16,8609 | 18,6772 |
|     | 3      | 4,1309  | 5,6996  | 6,6988  | 8,8234  | 12,1811 | 20,5239 | 33,2758 | 39,6153 |
|     | >3     |         | 0,8245  | 0,8461  | 1,0215  | 2,9173  | 4,0916  | 9,005   | 13,3196 |
| K   | 1      | 18,7075 | 19,5005 | 17,0238 | 12,7684 | 13,8138 | 20,1613 | 21,374  | 17,801  |
|     | 3      | 10,5227 | 14,7212 | 19,1275 | 23,7186 | 32,1676 | 45,4398 | 76,0145 | 88,4122 |
|     | >3     | 1,2029  | 1,5796  | 3,6041  | 5,0789  | 6,2756  | 9,6386  | 20,5007 | 17,5439 |

&gt;=3 månader duration:

| Kön | Karens | 25-29   | 30-34   | 35-39   | 40-44   | 45-49   | 50-54   | 55-59   | 60-     |
|-----|--------|---------|---------|---------|---------|---------|---------|---------|---------|
| M   | 1      | 10,3069 | 11,0571 | 12,7753 | 11,1514 | 12,1272 | 15,9815 | 15,9556 | 17,9918 |
|     | 3      | 4,0615  | 5,4146  | 6,3599  | 8,2056  | 11,3488 | 19,4044 | 31,0088 | 36,9542 |
|     | >3     |         | 0,8245  | 0,8461  | 1,0215  | 2,8216  | 3,712   | 8,8393  | 12,287  |
| K   | 1      | 16,4399 | 17,7899 | 16,7303 | 12,188  | 11,7117 | 19,8511 | 18,8295 | 16,7539 |
|     | 3      | 10,1814 | 14,0811 | 18,1476 | 22,8251 | 31,0628 | 43,2142 | 72,8929 | 81,0882 |
|     | >3     | 1,2029  | 1,5796  | 3,6041  | 5,0789  | 6,2756  | 8,7508  | 20,1064 | 17,1053 |

&gt;=6 månaders duration:

| Kön | Karens | 25-29  | 30-34   | 35-39   | 40-44   | 45-49   | 50-54   | 55-59   | 60-     |
|-----|--------|--------|---------|---------|---------|---------|---------|---------|---------|
| M   | 1      | 5,268  | 7,2977  | 8,1057  | 6,9033  | 7,5458  | 11,1871 | 11,7687 | 14,2221 |
|     | 3      | 3,1937 | 4,5378  | 4,9444  | 6,661   | 9,8924  | 17,0535 | 26,8797 | 33,4462 |
|     | >3     |        | 0,8245  | 0,7253  | 0,8602  | 2,5347  | 3,5011  | 8,2316  | 11,7708 |
| K   | 1      | 7,3696 | 11,6319 | 11,447  | 9,2861  | 6,006   | 13,9578 | 13,7405 | 12,5654 |
|     | 3      | 8,5888 | 12,041  | 15,4431 | 19,9415 | 27,8333 | 39,3657 | 65,8905 | 71,4099 |
|     | >3     | 0,8019 | 1,5796  | 3,4754  | 4,8044  | 5,8663  | 8,4971  | 18,5295 | 16,2281 |

&gt;=9 månaders duration:

| Kön | Karens | 25-29  | 30-34  | 35-39   | 40-44   | 45-49  | 50-54   | 55-59   | 60-     |
|-----|--------|--------|--------|---------|---------|--------|---------|---------|---------|
| M   | 1      | 4,1228 | 4,4228 | 4,6696  | 4,1597  | 5,0305 | 7,6356  | 8,9397  | 11,8232 |
|     | 3      | 2,2564 | 3,3979 | 3,7482  | 4,8268  | 7,7928 | 14,0869 | 23,7491 | 29,0311 |
|     | >3     |        | 0,7495 | 0,6044  | 0,8602  | 2,3434 | 3,3745  | 7,5686  | 10,8415 |
| K   | 1      | 3,9683 | 6,8423 | 7,6313  | 6,9646  | 4,2042 | 9,9256  | 11,1959 | 11,5183 |
|     | 3      | 5,8586 | 8,4007 | 12,2683 | 15,9613 | 23,244 | 33,338  | 56,357  | 61,47   |
|     | >3     | 0,8019 | 1,4479 | 3,3466  | 3,4317  | 5,457  | 8,2435  | 17,3467 | 15,3509 |

&gt;=1 års duration:

| Kön | Karens | 25-29  | 30-34  | 35-39  | 40-44   | 45-49   | 50-54   | 55-59   | 60-     |
|-----|--------|--------|--------|--------|---------|---------|---------|---------|---------|
| M   | 1      | 3,2066 | 3,096  | 3,7004 | 2,8321  | 4,2221  | 5,8599  | 6,9028  | 9,5956  |
|     | 3      | 1,5621 | 2,6087 | 2,9905 | 3,8615  | 6,1094  | 12,0531 | 20,4296 | 25,7046 |
|     | >3     |        | 0,7495 | 0,6044 | 0,8602  | 2,1521  | 3,0371  | 6,9057  | 9,809   |
| K   | 1      | 1,1338 | 4,4475 | 5,5768 | 5,2234  | 3,6036  | 6,8238  | 7,1247  | 8,377   |
|     | 3      | 3,7541 | 6,2405 | 9,3678 | 13,0777 | 18,5272 | 27,6348 | 49,6921 | 53,6228 |
|     | >3     | 0,6014 | 1,3163 | 3,0892 | 2,4708  | 5,3206  | 7,6094  | 16,7554 | 13,5965 |

Figur 24: Observerade t-frekvenser i promille för 2002 års sjuklighetsundersökning