



Stockholms
universitet

Svenskarnas dödlighet - ett alternativ till Makeham & Lee-Carter

Jimmy Calén

Masteruppsats 2017:11
Försäkringsmatematik
September 2017

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm



Matematisk statistik
Stockholms universitet
Masteruppsats **2017:11**
<http://www.math.su.se/matstat>

Svenskarnas dödlighet - ett alternativ till Makeham Lee-Carter

Jimmy Calén*

September 2017

Sammanfattning

Developing ways of describing death rates and expected life lengths is of crucial relevance for life insurance companies as well as for demographic research in general. In Sweden the well-known Makeham model is most often used to describe mortality intensity as a function of age. The Lee-Carter model is subsequently used to predict future mortality. The two well-established models both have weaknesses. This master thesis investigates the need for an implementation of an alternative way of describing and predicting future mortality rates in the Swedish population. We present evidence for implementation of the alternative method and discuss similarities and differences between the methods.

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

Tack

Jag vill tacka Mathias Lindholm och Filip Lindskog på matematiska institutionen för deras värdefulla, konstruktiva handledning. Ett särskilt tack till mina nära och kära för deras stöttning under studietiden.

Innehåll

1	Inledning.....	5
1.1	Bakgrund	5
2	Beskrivning av data.....	6
3	Matematiska definitioner	8
3.1	Livslängd.....	8
3.2	Dödlighetsintensitet.....	9
3.3	Återstående livslängd	9
3.4	Dödsrisk	9
3.5	Överlevelsefunktion	9
3.6	Förväntad återstående livslängd	11
4	Ansättning av modeller.....	12
4.1	Makeham-modellen.....	12
4.2	Lee-Carter-modellen	12
4.2.1	Stokastisk modell i Lee-Carter-modellen	13
4.3	CH-funktionen	15
5	Matematisk modellering	17
5.1	Tillvägagångssätt	18
5.2	Makeham-modellen.....	19
5.2.1	Resultat av Makeham-anpassning.....	20
5.3	CH-modellen	22
5.3.1	Resultat av CH-anpassning.....	22
5.4	Jämförelse mellan Makeham- & CH-anpassning.....	26
5.5	Lee-Carter-modellen för framskrivning av dödlighetsintensitet	31
5.5.1	Resultat av Lee-Carter-simulering	32
5.6	Framskrivning av parametrar i CH-modellen	36
5.7	Jämförelse mellan Lee-Carter- & CH-modellens framskrivna dödlighetsintensiteter	40
5.8	Omvandling av dödlighetsintensiteter till förväntade livslängder	41
6	Slutord	47
7	Förslag på förbättringar & framtida studier	48
8	Bilagor.....	50
8.1	CH-funktion för dödlighetsintensitet.....	50
8.2	Singular Value Decomposition (SVD)	52
8.3	Matlab-kod för Lee-Carter-modellering	53
8.4	Resultat från Lee-Carter-modellering.....	55
8.5	CH-modellens parametrar.....	59
8.6	Dödlighetsintensitet för svenska män.....	62

8.7	Dödlighetsintensitet för svenska kvinnor	63
9	Litteraturförteckning	64

1 Inledning

Prediktion av dödlighet är en viktig gren av försäkringsbolagens verksamhet gällande framförallt rättvis prissättning av livförsäkringsprodukter. Flertalet olika metoder har genom åren prövats. Detta avsnitt berör primärt en i sammanhanget ny överlevelsefunktion framtagen av Wong & Tsui (2015). Funktionen är olik de vi ofta ser i svensk litteratur. Wong & Tsui har i sin studie *Forecasting life expectancy: Evidence from a new survival function* hittat evidens för att den nya överlevelsefunktionen med stor precision beskriver amerikanska mäns och kvinnors dödlighetsintensitet.

Vi applicerar teorin på svensk dödlighetsdata. Vi jämför sedan precisionen med den vi får genom användande av Makeham-modeller och Lee-Carter-modeller. Prediktion av framtida dödlighetsintensiteter görs sedan medelst tidsserieanalys.

1.1 Bakgrund

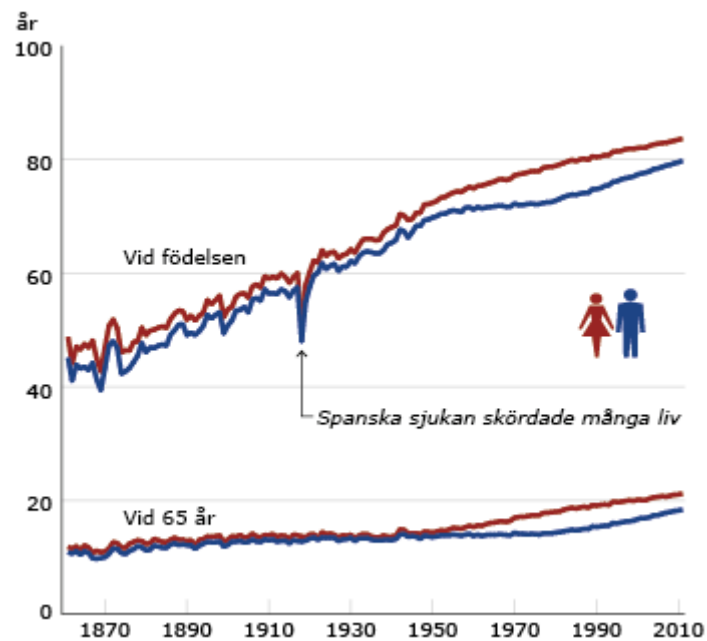
Sveriges befolkning har under mycket lång tid mött en mängd förändringar som resulterat i en aldrig tidigare skådad medellivslängd. Hälsosammare levnadsvanor, utveckling av sjukvård och medicinska behandlingsmetoder är några av de förklarande variablerna. Sedan flera decennier tillbaka (sedan mitten av 1900-talet för kvinnor, sedan slutet av 1900-talet för män) drivs trenden av framförallt minskad dödlighet i pensionsåldern. I motsats till tidigare då vi snarare observerat en minskad barnadödlighet (Folkhälsomyndigheten, 2016).

I de åldrar där det för livförsäkringsbolagens, största behovet av statistiskt säkra skattningar av dödlighet finns, har vi alltså länge sett en avtagande dödlighet som komplicerar såväl prissättning av premier som bolagets egna försäkringstekniska avsättningar¹.

Utveckling i medellivslängd enligt nedan illustration från Statistiska Centralbyrån, SCB (2012).

¹ Benämns ofta som FTA i aktuariella sammanhang

Figur 1



2 Beskrivning av data

Livslängdsantaganden i gällande studie utgår från data sammanställd av SCB benämnd som *Ettårig livslängdstabell för hela riket efter kön och ålder. År 1980 – 2016*. Statistiken är officiell och presenterar följande data, vilka vi kommer definiera närmare i avsnitt 3 och avsnitt 4:

- *Risktid*
- *Antal döda*
- *Antal döda efter födelsedagen*
- *Dödsrisker (promille)*
- *Kvarlevande per 100 000 levande födda*
- *Återstående medellivslängd*

Detta för såväl män som kvinnor år 1980-2016. Vi har alltså fullständig data över befolkningsdödlighet för den svenska populationen under de 37 senaste åren (SCB, 2017).

Vi citerar SCB angående dödsriskerna: ”Till och med 1986 baserades dödsriskerna från 91 år och äldre på Wittsteins formel. Mellan 1987 och 2005 baseras utjämnningen för åldrarna 91 år och äldre på en utjämningsmetod som utvecklades på SCB. Under 2012 skedde en översyn

av utjämningsmodellen vilket medförde en mindre justering av modellen samt att utjämnigen sker från 95 års ålder.” (SCB, 2017)

Välkänt är att det föreligger stora skillnader i dödlighet mellan delgrupper av populationen. En av de mest uppenbara uppdelningarna som kan vara frestande att göra är den könsberoende. Av EU-direktiv gällande lagstiftningen om diskriminering följer dock sedan 21 december 2012, krav på könsberoende premiesättning (Europeiska kommissionen, 2012). Skälet till att vi trots detta analyserar könsseparerad data, är för att det inte föreligger några hinder för försäkringsbolagen att internt i bokföringen göra uppdelningen för att kunna avsätta rätt kapital. Under de senare decennierna har vi dessutom observerat att männen tycks knappa in på kvinnorna – vilket till viss del förklaras av ändrade vanor gällande rökning för framförallt män.

Att vi i denna studie använder data över befolkningsdödlighet och inte över någon specifik undergrupp – är för att ge en demografisk överblick av de hälsoförbättringar det svenska samhället genomgått. Att överföra trender i befolkningsdödlighet till dödlighet inom särskilda delpopulationer görs sedermera genom enklare algebraisk manipulation, varför valet att använda sig av data över befolkningsdödlighet i sig inte är en inskränkning. Tvärtom, så är det ofta nödvändigt att börja så – då data över befolkningsdödlighet generellt är väsentligt mer omfattande än för delpopulationer.

Trots att vi använder den mest omfattande data som finns att tillgå i Sverige – möter vi svårigheter att göra statistiskt säkra skattningar vid särskilt höga åldrar. I enstaka mätpunkter för kvinnor tvingas vi vid ansättning av Lee-Carter-modellen också göra approximationer av dödlighetsintensitet då inte en enda person i en viss ålder avlidit under året som gått. År 1994 avled till exempel inte en enda 8-årig flicka, varför vi i de enstaka fallen tvingas göra approximationer. Allmänt gäller att vid åldrarna 100 år och uppåt är datamaterialet mycket bristfälligt, och SCB (2017) anger i många fall inte ens någon risktid för så pass höga åldrar. Vi inskränker därför vår modellering till åldrarna 0-99 om inte annat görs gällande.

3 Matematiska definitioner

Vi definierar initialt grundläggande notation som genomgående kommer användas i denna studie. Vi utgår i samtliga fall då annat inte görs gällande, från den notation som presenteras i *Livförsäkringsmatematik* (Andersson, 2013).

3.1 Livslängd

Definition 3.1: Låt *livslängden* T för en godtyckligt vald individ vara en icke-negativ kontinuerlig stokastisk variabel med följande fördelningsfunktion:

$$F(x) = P(T \leq x), \quad x \geq 0.$$

Vi definierar med välkänd notation täthetsfunktionen:

$$f(x) = F'(x), \quad x \geq 0.$$

Vi betraktar nu intervallet $(x, x + dx)$. Sannolikheten att en individ som lever vid x , dör i intervallet $(x, x + dx)$ kan om dx är litet anses proportionell mot längden på intervallet. Sannolikheten kan därför uttryckas genom följande approximativa relation:

$$\mu_x \cdot dx \approx \frac{F(x + dx) - F(x)}{1 - F(x)},$$

vilket efter omskrivning kan skrivas som:

$$\mu_x \approx \frac{1}{dx} \frac{F(x + dx) - F(x)}{1 - F(x)},$$

där μ_x representerar en proportionalitetsfaktor. Detta ger:

$$\mu_x = \lim_{dx \rightarrow \infty} \frac{1}{dx} \frac{F(x + dx) - F(x)}{1 - F(x)} = \frac{f(x)}{1 - F(x)}, \quad (1)$$

Detta kallar vi *dödlighetsintensitet* för åldern x . Således följer nästa definition.

3.2 Dödlighetsintensitet

Definition 3.2: Låt T vara en icke-negativ stokastisk variabel som representerar en godtycklig individs livslängd. *Dödlighetsintensiteten* för individen, vid åldern x , definieras som:

$$\mu_x = \frac{f(x)}{1 - F(x)}. \quad (2)$$

För att kunna hantera *återstående livslängd* vid åldern x hos en individ, ansätter vi nästa definition.

3.3 Återstående livslängd

Definition 3.3: Låt T_x vara en icke-negativ stokastisk variabel som representerar *återstående livslängd* vid åldern x för en godtycklig individ. Fördelningsfunktionen för T_x definieras som:

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

Vi har således med användande av 3.1, om *livslängd*, $T = T_0$:

$$P(T_x > t) = P(T > x + t | T > x) = \frac{P(T > x + t)}{P(T > x)}. \quad (3)$$

Vi observerar därmed att om en individ uppnått åldern x kan individens totala livslängd skrivas som $x + T_x$. Vi formulerar därför definitionen av den *ettåriga dödsrisken* enligt följande:

3.4 Dödsrisk

Definition 3.4: Den ettåriga *dödsrisken* q_x , för åldern x definieras som:

$$q_x = P(T_x \leq 1). \quad (4)$$

Vi introducerar nu begreppet *överlevelsefunktion*.

3.5 Överlevelsefunktion

Definition 3.5: Funktionen l_x definieras som *överlevelsefunktionen* till den återstående livslängden T_x genom:

$$l_x(t) = 1 - F(x) = P(T_x > t), \quad t \geq 0, \quad (5)$$

vilket således representerar sannolikheten att en x -årig individ lever i ytterligare t år. Vi kan härmed medelst omskrivning uttrycka följande samband:

$$P(T_x > t) = \frac{l_0(x+t)}{l_0(x)}, \quad t \geq 0,$$

vilket kan uttryckas som:

$$l_x(t) = \frac{l_0(x+t)}{l_0(x)}, \quad t \geq 0. \quad (6)$$

Vi ser alltså att vi genom betingning kan uttrycka överlevelsefunktionen för en x -årig individ i termer av överlevelsefunktionen för en nyfödd, det vill säga $x = 0$.

Vi avser nu definiera sambandet mellan dödlighetsintensiteten μ_x och överlevelsefunktionen l_x . Med hänvisning till definition 3.2 ges vi att eftersom:

$$l'(t) = \frac{d(1-F(t))}{dt} = -f(t),$$

så kan sambandet skrivas som:

$$\mu_x = -\frac{l'(x)}{l(x)} = \frac{-d(\ln l(x))}{dx}. \quad (7)$$

Vilket med användande av $l(0) = 1$, ger:

$$l(x) = \exp \left\{ \int_0^x \mu_s ds \right\}. \quad (8)$$

Vi observerar också att följande samband gäller:

$$l(x) = l(x-1) \cdot (1 - q_{x-1}) = \prod_{i=0}^{x-1} (1 - q_i), \quad x = 1, 2, \dots \quad (9)$$

Vi definierar slutligen *förväntad återstående livslängd* enligt följande:

3.6 Förväntad återstående livslängd

Definition 3.6: Låt T_x vara den återstående livslängden för en godtycklig individ med åldern x . Med förväntad återstående livslängd för T_x menar vi:

$$e_x = E(T_x) = \int_0^{\infty} \frac{l(x+t)}{l(x)} dt, \quad (10)$$

vilket vi med hjälp av Trapetsregeln alternativt Euler-Maclaurins summationsformel kan approximera genom följande:

$$e_x \approx \left[\sum_{i=0}^{\infty} \frac{l(x+i)}{l(x)} \right] - \frac{1}{2}, \quad x=1, 2, \dots \quad (11)$$

alternativt den något mer noggranna

$$e_x \approx \left[\sum_{i=0}^{\infty} \frac{l(x+i)}{l(x)} \right] - \frac{1}{2} - \frac{1}{12} \cdot \mu_x, \quad x=1, 2, \dots \quad (12)$$

För det i sammanhanget kanske viktigaste sambandet för förståelse av den modellering som görs i denna studie presenterar vi också det approximativa sambandet mellan dödlighetsintensitet och dödsrisk. Med bakgrund i ovan definitioner kan vi med hjälp av omskrivningar och Taylorutveckling etablera följande:

$$\mu_{x+\frac{1}{2}} \approx \frac{q_x}{1 - \frac{q_x}{2}}, \quad (13)$$

och analogt:

$$q_x \approx \frac{\mu_{x+\frac{1}{2}}}{1 + \frac{\mu_{x+\frac{1}{2}}}{2}}. \quad (14)$$

För detaljer kring ovan samband hänvisas till Andersson (2013)

4 Ansättning av modeller

4.1 Makeham-modellen

Makehams modell är den modell som nästan uteslutande använts i Sverige för utjämning av observerad dödlighet. Modellen presenterades första gången år 1860, och kan med sina relativt få parametrar väl beskriva dödlighet. Dock observerar vi sämre anpassning vid låga såväl som mycket höga åldrar. Modellen ges på följande form:

$$\mu_x = \alpha + \beta \cdot e^{\gamma x}. \quad (15)$$

För att i någon mening minimera felskattningen vid höga åldrar används ibland en linearisering (Andersson, 2013). Vi definierar därför i enlighet med Pensionsmyndigheten (2012), den linearisering som görs gällande vid 97 års ålder. Makeham-fördelningen ges då på följande format:

$$\mu_x = \begin{cases} \alpha + \beta \cdot e^{\gamma x} & , \quad x \leq 97, \\ \mu_{97} + 0,03 \cdot (x - 97) & , \quad x > 97. \end{cases} \quad (16)$$

På grund av den i många fall bristfälliga data (vid mycket höga åldrar) man ofta använder, fokuseras då på ovan beskriven metod med linearisering.

4.2 Lee-Carter-modellen

Lee-Carter-modellen använder sig, i motsats till Makeham-modellen, av en trendbaserad utjämning. För att kunna beskriva trender i dödligheten använder vi således hela det datamaterial som SCB tillhandahåller. Trots att modellen tidigare visat god skattningsförmåga i det korta och medellånga perspektivet saknas möjlighet att fånga åldersspecifika förändringar i dödlighet. Som nämnt i bakgrunden till denna studie, avsnitt 1.1, så gällde under 1900-talets början nedgången i dödlighet framförallt yngre, medan de senaste decenniernas nedgång snarare förklaras av de äldre. Detta kan alltså innebära att användande av datamaterial som täcker en alltför lång tidsserie ger instabilitet. Av samma skäl gäller också det omvända. Att använda en kort och idag närliggande period kommer generera prediktioner där männens dödlighet alltmer närmar sig kvinnornas, och i somliga åldrar även att bli lägre än kvinnornas om trenderna i denna period antas verka i oförändrad takt långt in i framtiden.

I studien författad av Lundström & Qvist (2004) görs gällande att en basperiod på omkring 25 år är lagom för att fånga de båda könen nedgång i dödlighet. Då inga större trendbrott har skett sedan studies publicerades väljer vi att anpassa modellen till det datamaterial som rör de senaste 37 årens dödlighet. Trendbrott av de typer beskrivna ovan bör därför inte utgöra någon större risk för felskattningar i våra framtida prediktioner.

Metoden där man använder historisk data för att förutspå framtida dödlighet kallas allmänt för *prediktion* eller *framskrivning*.

4.2.1 Stokastisk modell i Lee-Carter-modellen

Implementering av Lee-Carter-modellen kräver en utvidgning av våra tidigare definitioner – både genom att inkludera kalenderåret och en dessutom en ansättning av fler definitioner.

Vi knyter således an den teori presenterad i avsnitt 3 till den årliga data som studien utgår från. Vi definierar någorlunda analogt med tidigare, följande då t representerar kalenderår:

$$\left\{ \begin{array}{l} T_x(t) = \text{återstående livslängd för en } x\text{-årig individ vid kalenderår } t. \\ N_x(t) = \text{antal individer som lever vid utgången av kalenderår } t \\ \qquad \qquad \qquad \text{och fyllde } x \text{ år under kalenderåret } t. \\ D_x(t) = \text{antal individer som avlider under kalenderår } t \text{ och} \\ \qquad \qquad \qquad \text{fyllde eller skulle ha fyllt } x \text{ år under kalenderår } t. \\ \mu_x(t) = \text{Dödlighetsintensiteten för åldern } x \text{ vid kalenderår } t. \\ q_x(t) = P(T_x(t) \leq 1) = \text{Dödsrisken för åldern } x \text{ vid kalenderår } t. \end{array} \right. \quad (17)$$

Vi definierar också begreppet *risktid* för den x -åriga delen av populationen under kalenderår t enligt följande:

4.2.1.1 Risktiden för den x -åriga delen av populationen under kalenderåret t

Definition: Risktiden för den x -åriga delen av populationen under kalenderår t ges av summan av den tid som de individer som är födda under kalenderåret $t - x$ har varit en del av beståndet under kalenderåret t . Således:

$$R_x(t) = \sum_{i=1}^{N_x(t)} R_{x,i}(t). \quad (18)$$

Där $R_{x,i}(t)$ representerar risktiden för person i .

4.2.1.2 Andel av ett försäkringsbestånd som lämnar beståndet på annat sätt än att avlida

För att mäta hur stor *andel av ett försäkringsbestånd som lämnar beståndet på annat sätt än att avlida* anger vi kvoten:

$$\xi_x(t) = \frac{N_{x-1}(t-1) - D_x(t)}{N_x(t)}, \quad (19)$$

vilket, om ingen migration in och ut ur befolkning sker, innebär att $N_{x-1}(t-1) = N_x(t) + D_x(t)$.

Eftersom Lee-Carter-modellen använder sig av trendbaserad utjämning ges vi också anledning att definiera *dödlighetsintensiteten för en viss ålder x med födelseår F* . Vi inför begreppet *födelsekohort*, det vill säga en grupp människor födda under samma kalenderår. En individ som uppnår ålder x under kalenderår t är född kalenderår $F = t - x$.

Vi ges således följande samband:

$$\mu_x^*(F) = \mu_x(x + F). \quad (20)$$

Vi antar vidare att de i populationen ingående livslängderna är av varandra oberoende. Invändningar till det antagandet finns givetvis – det finns till exempel viss evidens för att om ett sammanlevande par tillsammans uppnår hög ålder och den ena plötsligt avlider, tenderar den efterlevande att själv få en högre sannolikhet att avlida inom en snar framtid (Andersson, 2013). På populationsnivå bör den typen av beroende ha en försumbar inverkan på prediktion av framtida livslängder. Således gäller att $D_x(t)$ är binomialfördelad enligt:

$$D_x(t) \sim \text{Bin}(N_{x-1}(t-1), q_x(t)). \quad (21)$$

Vidare gäller enligt grundläggande sannolikhetsteoretisk teori att binomialfördelningen asymptotiskt kan approximeras med en Poisson-fördelning. Dock kan det finnas skäl att uttrycka approximationen i termer av dödlighetsintensitet (som ju är primär i Lee-Carter-modellen) och inte dödsrisk som är fallet i (21). Brouhns et al (2005) diskuterar i sin studie nedan Poisson-approximation, med dödlighetsintensitet som ingående parameter:

$$\begin{cases} D_x(t) \sim Po(R_x(t) \cdot \mu_x(t)), \\ E[D_x(t)] = R_x(t) \cdot \mu_x(t). \end{cases}$$

Lee-Carter-modellen modellerar slutligen dödlighetsintensiteten enligt följande:

$$\ln \left[\hat{\mu}_x(t) \right] = \ln [\mu_x(t)] + \varepsilon_x(t) = \alpha_x + \kappa(t)\beta_x + \varepsilon_x(t), \quad (22)$$

där vi alltså beräknar dödlighetsintensitet enligt nedan då $d_x(t)$ representerar observerat värde av den stokastiska variabeln $D_x(t)$:

$$\mu_x(t) = \frac{d_x(t)}{R_x(t)}. \quad (23)$$

Vi analyserar ovan närmare i avsnitt 5.5.

4.3 CH-funktionen

Som nämnt i inledningen presenteras i studien *Forecasting life expectancy: Evidence from a new survival function* en ny överlevelsefunktion (Wong & Tsui, 2015). Välkänt är att Makeham-modellen som traditionellt används för att beskriva dödlighetsintensiteter ofta har svårt att beskriva såväl mycket låga som mycket höga åldrar. CH-funktionen består av en summa av två funktioner. Den första är tänkt att kunna bidra till att fånga dödlighetsmönster i framförallt låga åldrar, och den andra framförallt i högre åldrar med förhoppningen om att bättre kunna beskriva dödlighetsintensiteter där Makeham-modellen brister.

Vi ansätter:

$$l(x) = \alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}, \quad (24)$$

med följande exponenter

$$\begin{cases} \theta_1(x) = \exp\{(x/\beta_1)^{\gamma_1}\}, \\ \theta_2(x) = \cosh\{(x/\beta_2)^{\gamma_2}\}. \end{cases}$$

Vi påminner oss om att den hyperboliska cosinus-funktionen definieras genom följande:

$$\cosh(y) = \frac{e^y + e^{-y}}{2} = \frac{e^{2y} + 1}{2e^y} = \frac{1 + e^{-2y}}{2e^{-y}}.$$

Vi avser nu skriva överlevelsefunktionen (24) som en funktion av μ_x för att kunna jämföra med (15). Detta görs enklast med användande av (7):

$$\begin{aligned} \mu_x &= -\frac{l'(x)}{l(x)} = -\frac{d(\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)})}{dx} \cdot \frac{1}{\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}}. \\ &\Rightarrow \\ \mu_x &= -\frac{l'(x)}{l(x)} = \frac{\omega_1(x)}{1 + (\alpha_2/\alpha_1)e^{\theta_1(x) - \theta_2(x)}} + \frac{\omega_2(x)}{1 + (\alpha_1/\alpha_2)e^{\theta_2(x) - \theta_1(x)}}, \end{aligned} \tag{25}$$

där

$$\begin{cases} \omega_1(x) = (\gamma_1/\beta_1)(x/\beta_1)^{\gamma_1-1} \theta_1(x), \\ \omega_2(x) = (\gamma_2/\beta_2)(x/\beta_2)^{\gamma_2-1} \sinh\{(x/\beta_2)^{\gamma_2}\}, \\ \alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2 \in \mathbb{R}_{>0} \end{cases}$$

Vi kan också konstatera att följande, för överlevelsefunktioner, fundamentala samband är uppfyllda:

$$\begin{aligned} i) \quad & l(0) = 1, \\ ii) \quad & l(\infty) = 0. \end{aligned}$$

För detaljer kring ovan beräkning hänvisas till bilaga, avsnitt 8.1.

Vi avser i nästa avsnitt applicera de olika modellerna på de data över befolkningsdödlighet som tillhandahålls av SCB (2017), för att kunna studera hur utfallen skiljer sig i praktiken.

5 Matematisk modellering

Det datamaterial vi utgår från ges av SCB (2017) på följande format för män respektive kvinnor:

		1980	...	2016
0 år	Risktid	49860	...	60199
	Antal döda	402	...	157
	Antal döda efter födelsedagen	364	...	131
	Dödsrisker (promille)	8,06	...	2,61
	Kvarlevande per 100 000 levande födda	100000	...	100000
	Återstående medellivslängd	72,76	...	80,56
...				
110 år	Risktid			
	...			

Vi anpassar samtliga våra modeller med avseende på dödlighetsintensiteten – som alltså ej anges explicit i SCB:s tabeller. Enligt den teori presenterad i ekvation (23), använder vi således, för varje ålder och kalenderår, kvoten nedan för att beräkna dödlighetsintensitet.

$$\frac{\text{Antal döda}}{\text{Risktid}} = \frac{d_x(t)}{R_x(t)} = \mu_x(t),$$

Allmänt förekommer, som nämnt i avsnitt 2, stora brister i datamaterialet vid mycket höga åldrar, 100 år och uppåt. SCB (2017) anger i många fall inte ens någon risktid, och i de fall det görs är antalet observationer så pass få att vi väljer att anpassa våra modeller till åldrarna 0-99. Ytterligare skäl till den inskränkningen utöver den statistiska osäkerheten är att framtida dödsrisker vid särskilt höga åldrar inte antas förändras så mycket. I SCB:s publikation *Sveriges framtida befolkning 2016-2060* (2016) görs till exempel gällande att predikterade dödsrisker för både män och kvinnor antas förändras i mycket begränsad omfattning fram till år 2060 vid åldrarna 100 och uppåt. Nedgångarna i dödsrisk förväntas däremot vara av betydligt mer omfattande betydelse i yngre åldrar.

För kvinnor förväntas till exempel årlig procentuell förändring i dödsrisk (jämfört med året innan) år 2017-2060 vara i genomsnitt nästan 1,7 % för 75-åringar medan förändring för 100-åringar samma tidsperiod beräknas till endast 0,11 % (SCB, 2016).

Eftersom Lee-Carter-modellen använder en logaritmering av observerade dödlighetsintensiteter kan vi inte använda $\mu_x(t) = 0$, för någon kombination x, t . I de enstaka fallen, som i vårt datamaterial endast gäller för kvinnor ansätter vi istället följande approximation:

$$\hat{\mu}_x(t) = \frac{\mu_x(t-1) + \mu_x(t+1)}{2}, \quad t_{\min} < t < t_{\max}.$$

År 1994 avled till exempel inte en enda 8-årig kvinna varför vi då istället antar dödlighetsintensiteten det året enligt följande:

$$\hat{\mu}_8(1994) = \frac{\mu_8(1993) + \mu_8(1995)}{2}.$$

Ovan manipulering av data förekommer i så pass begränsad utsträckning att det inte kan antas ha någon väsentlig påverkan på vår modellering.

5.1 Tillvägagångssätt

Modelleringsavsnittet syftar i första steget till att anpassa modeller för att kunna beskriva observerade dödlighetsintensiteter åren 1980-2016. Det andra steget syftar till att med hjälp av de anpassade modellerna från det första steget, beskriva framskrivning av dödlighetsintensiteter för åren 2017-2050. Det tredje steget beskriver slutligen omvandling av dödlighetsintensiteter till det denna rapport avser mynna ut i – nämligen prediktioner av förväntade, framtida livslängder.

Vi följer följande procedur för det första steget:

1. Anpassar Makeham-modeller till observerade dödlighetsintensiteter för varje år 1980-2016.
2. Anpassar CH-modeller till observerade dödlighetsintensiteter för varje år 1980-2016.
3. Jämför med statistiska mått skillnader i anpassningsförmåga mellan Makeham-modellen och CH-modellen.

Vi följer efter ovan följande procedur för det andra steget – framskrivning av dödlighetsintensiteter:

4. Anpassar Lee-Carter-simulering för framskrivning av dödlighetsintensiteter för åren 2017-2050.
5. Anpassar autoregressiva modeller till CH-modellens parametrar för framskrivning av dödlighetsintensiteter för åren 2017-2050.
6. Jämför med statistiska mått skillnader mellan Lee-Carter-framskrivning och CH-framskrivning.

Vi följer slutligen följande procedur för det tredje steget:

7. Beräknar utifrån våra framskrivna dödlighetsintensiteter från det andra steget, överlevelsefunktioner och förväntade återstående livslängder för åren 2017-2050.

5.2 Makeham-modellen

I detta avsnitt anpassar vi för varje år 1980-2016 en makeham-modell för att så väl som möjligt beskriva dödlighetsintensiteter för respektive år. Som tidigare skiljer vi på män och kvinnor. Vi väljer att inte använda en linearisering vid höga åldrar utan anpassar modellen enligt nedan – detta då vi gjort inskränkningen till 0-99 års ålder.

$$\mu_x(t) = \alpha(t) + \beta(t) \cdot e^{\gamma(t)x}.$$

För att anpassa Makeham-modellen till observerade dödlighetsintensiteter använder vi en minstakvadrat-skattning. Vi avser alltså minimera följande kvadratsumma:

$$Q = \sum_{x=0}^{x_{\max}} [\mu_x(t) - \alpha(t) - \beta \cdot e^{\gamma(t)x}]^2. \quad (26)$$

För att inte oproportionerligt mycket vikt ska läggas vid statistiskt osäkra åldrar med få observationer adderar man ibland en viktning till residualkvadratsumman. Viktningen adderar en parameter som tar hänsyn till antalet avlidna, medan absolutavvikelse leder till att för stor vikt riskerar att ges till åldrar med få observationer. Vi bildar således följande:

$$Q_{viktad} = \sum_{x=0}^{x_{\max}} \omega_x(t) \cdot [\mu_x(t) - \alpha(t) - \beta(t) \cdot e^{\gamma(t)x}]^2, \quad \omega_x(t) = R_x(t) / \mu_x(t), \quad (27)$$

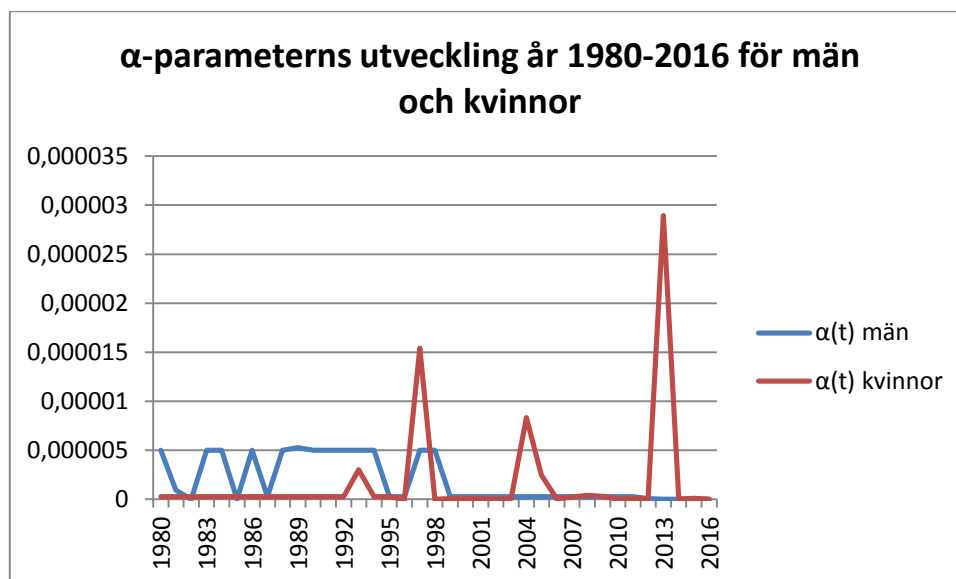
där $\omega_x(t)$ således representerar vikten.

I denna studie görs dock ingen manipulering av residualkvadratsumman enligt ovan – att analysera statistiskt osäkra åldrars påverkan på Makeham-modellens parametrar, och i senare avsnitt CH-modellen, är en del av den svårighet som kan vara intressant att studera.

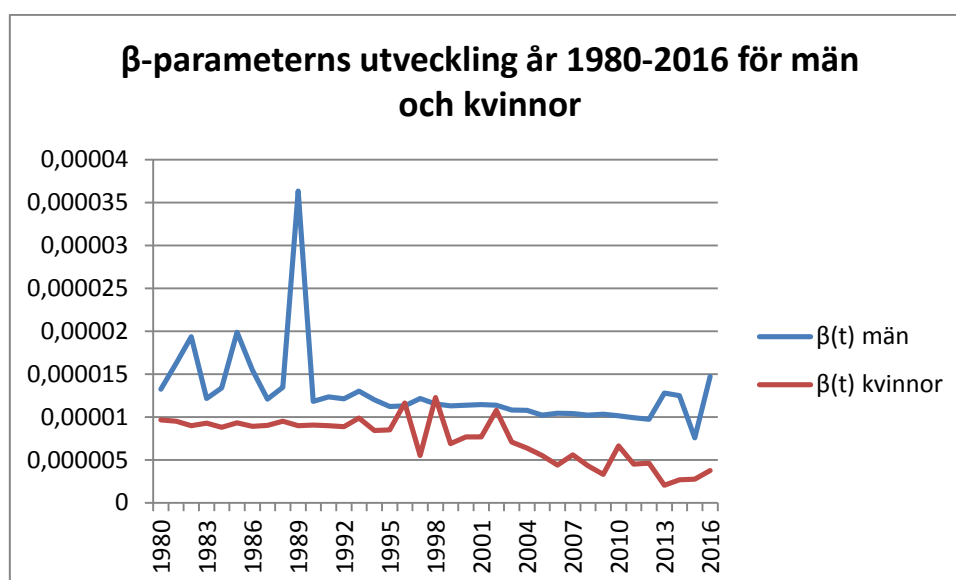
5.2.1 Resultat av Makeham-anpassning

Vi redovisar i nedan plottar estimat av de tre olika parametrarna i Makeham-modellen.

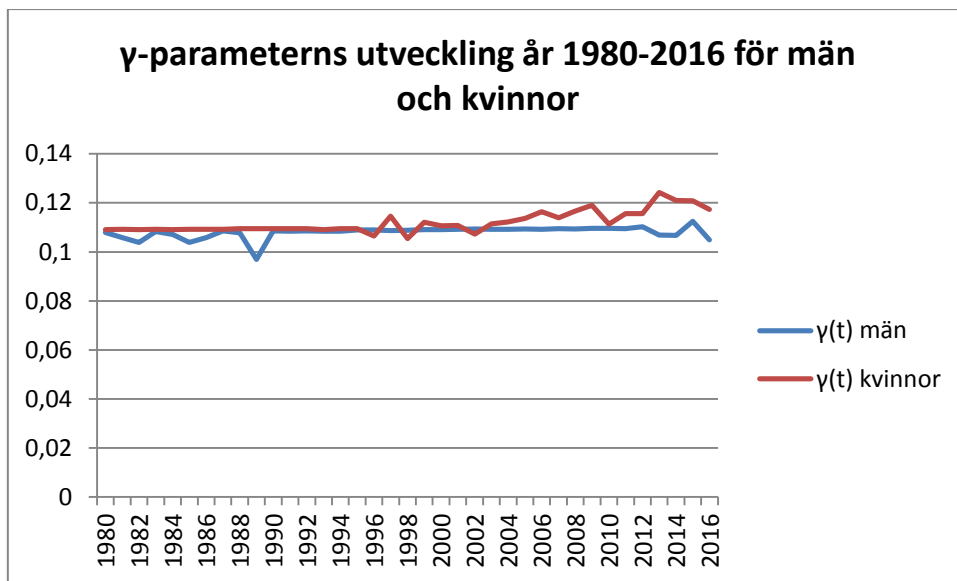
Figur 2



Figur 3

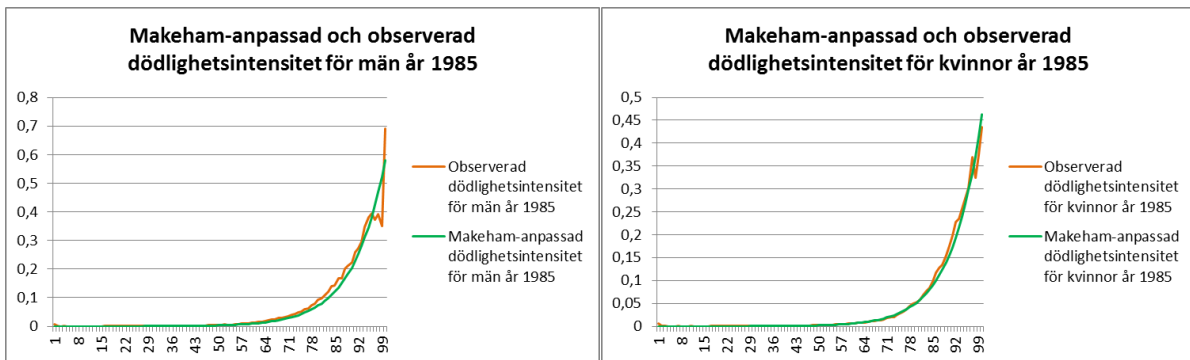


Figur 4

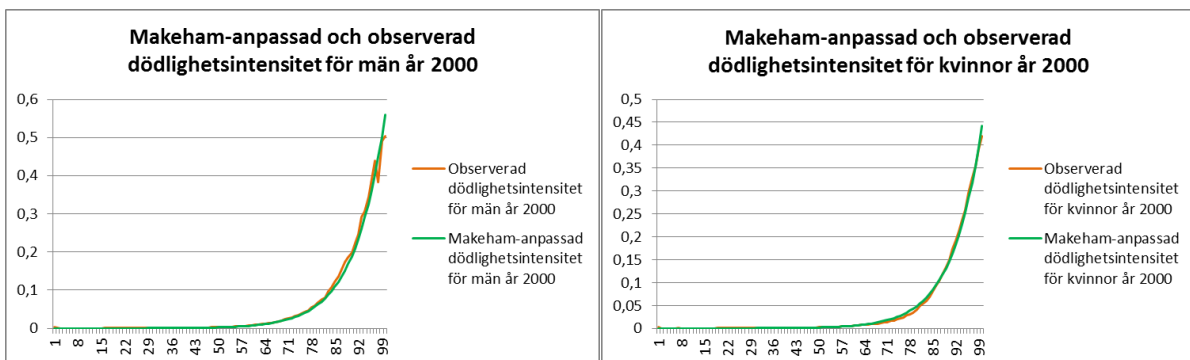


Nedan redovisar vi ett axplock av de estimerade Makeham-funktionerna plottade med observerade dödlighetsintensiteter. Vi redovisar av tydlighetsskäl män och kvinnor separat. Vi väljer att redovisa för åren 1985, 2000 & 2016.

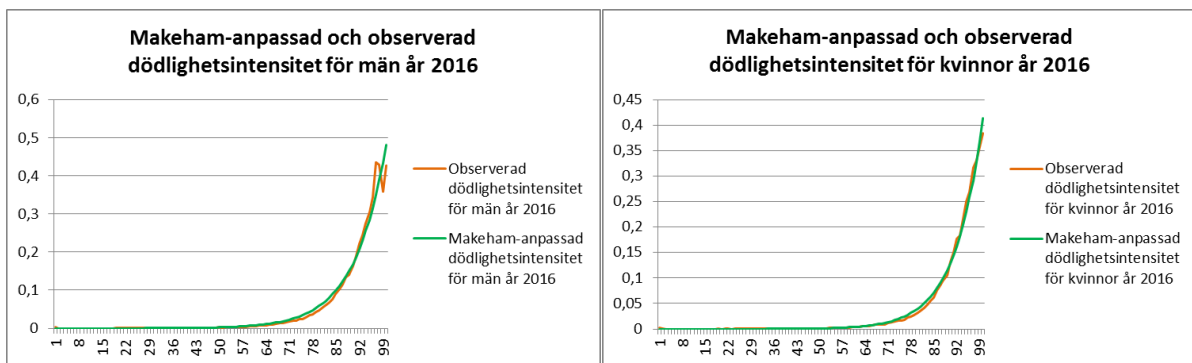
Figur 5



Figur 6



Figur 7



Vi observerar som väntat god men inte perfekt anpassning genom ansättning av Makeham-modeller för respektive år 1980-2016. Svårigheten att anpassa en funktion som fångar både unga och gamlas dödlighetsintensiteter gör sig påmind. Vidare analys görs i avsnitt 5.4.

5.3 CH-modellen

Vi avser nu applicera den teori som presenteras i inledningen till denna studie. Vi använder i likhet med Makeham-fallet klassisk teknik för minstakvadrat-skattning. Anpassningen av CH-modellens parametrar till vår observerade data är väsentligt mycket mer krävande än i Makeham-fallet. CH-modellen har 6 olika parametrar och är av betydligt mer avancerad natur än Makeham-modellen. Vi möter vissa svårigheter att hantera såväl mycket små, som mycket stora tal. En viss inskränkning görs därför i de fall Excel & Matlab ej klarar av att beräkna mycket stora exponenter. Vi behöver i somliga fall hantera tal i storleksordningen e^a , $|a| > 700$. Inskränkningen i de fall den behövs har dock försumbar inverkan på resultatet, och ger inte utslag ens inom 20e decimalen vid vår modellering av dödlighetsintensiteter.

5.3.1 Resultat av CH-anpassning

Vi kan inledningsvis konstatera att vårt datamaterial i särskilt höga åldrar i somliga fall uppvisar mycket ologiska mönster – som ställer till skattningarna av våra CH-parametrar. Detta är en konsekvens av det bristande datamaterial Sveriges, i sammanhanget, ringa befolkningsmängd utgör. För att vi i huvud taget ska ha möjlighet att grafiskt observera parametrarnas utveckling väljer vi att exkludera männens år 1985 i Figur 8 nedan. Detta då parametern α_1 i samtliga övriga år varierar mellan 0,024354 & 0,273606, men år 1985 estimeras till 1213,110158. Vi betraktar inledningsvis nedan tabell för att få en uppfattning om parametrarnas spridning. För fullständig tabell med varje års parametrar, residualkvadratsummor & R^2 -värden för modellen hänvisas till bilaga, avsnitt 8.5.

Tabell 1

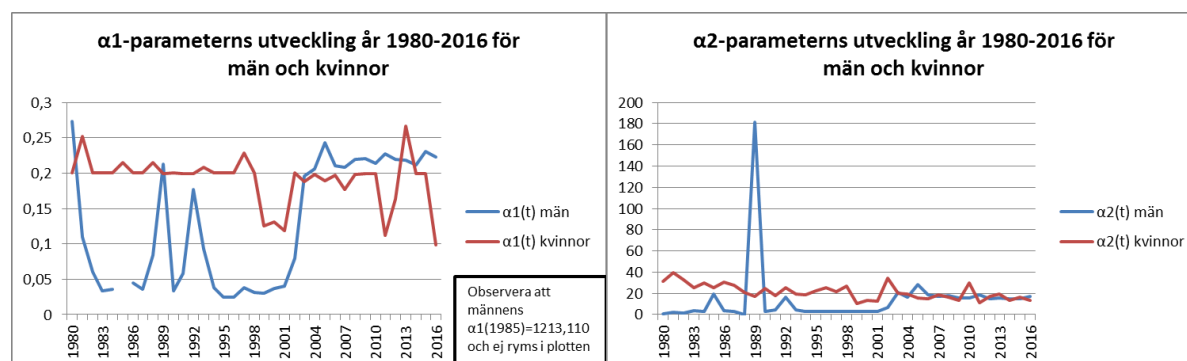
Sammanfattning av CH-parametrarna för män år 1980-2016						
	$\alpha 1$	$\beta 1$	$\gamma 1$	$\alpha 2$	$\beta 2$	$\gamma 2$
Medelvärde	32,912	40,649	3,422	14,174	74,793	3,562
Maximum	1213,110	97,487	11,530	181,373	94,724	12,156
Minimum	0,024	24,854	2,117	0,107	68,796	2,667
Standardavvikelse	196,700	23,445	1,881	28,867	4,166	1,472
Skevhet	6,083	1,689	3,056	5,447	2,901	5,574
Kurtosis	37,000	1,107	9,752	31,719	13,416	32,820

Tabell 2

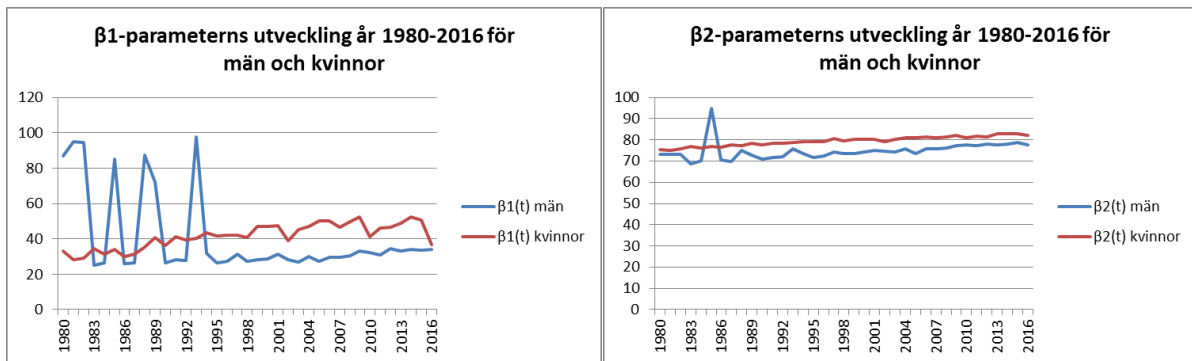
Sammanfattning av CH-parametrarna för kvinnor år 1980-2016						
	$\alpha 1$	$\beta 1$	$\gamma 1$	$\alpha 2$	$\beta 2$	$\gamma 2$
Medelvärde	0,191	41,616	3,111	21,393	79,424	3,834
Maximum	0,266	52,232	4,744	39,363	82,939	4,502
Minimum	0,098	28,399	2,105	10,628	75,176	3,213
Standardavvikelse	0,034	6,868	0,628	7,047	2,209	0,352
Skevhet	-1,057	-0,311	0,720	0,568	-0,290	-0,101
Kurtosis	1,910	-0,973	-0,286	-0,331	-0,924	-0,896

Vi redovisar i nedan plottar estimat av de tre olika parametrarna i CH-modellen.

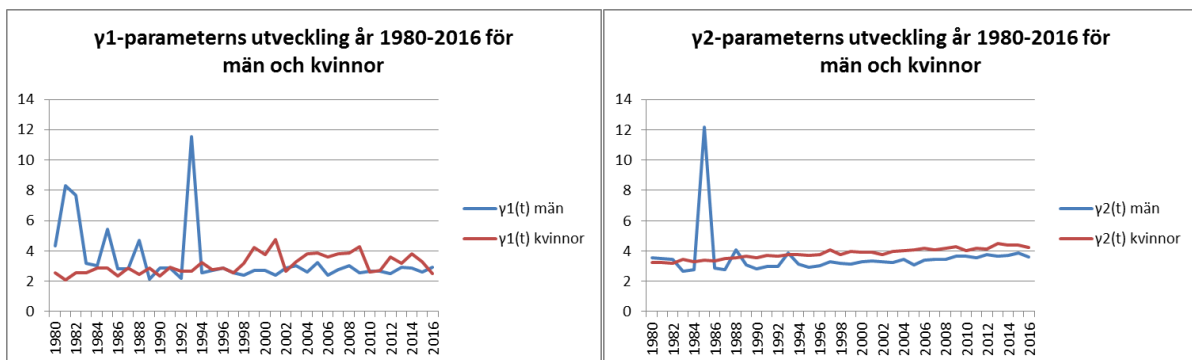
Figur 8



Figur 9

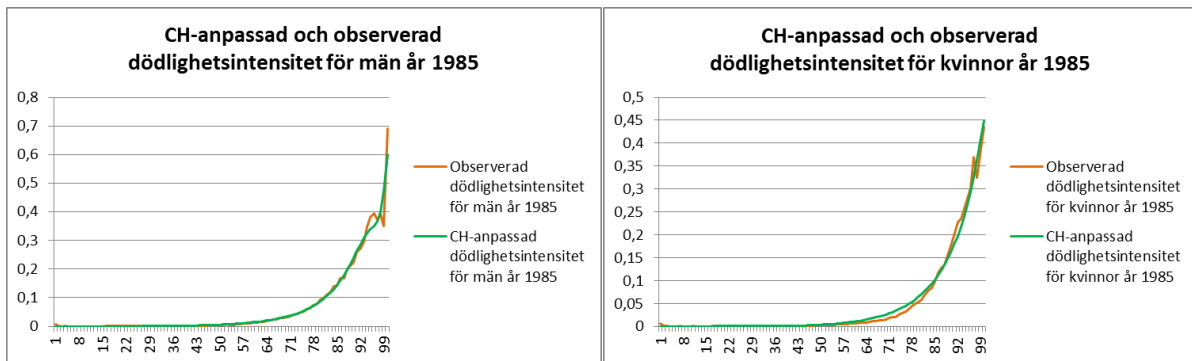


Figur 10

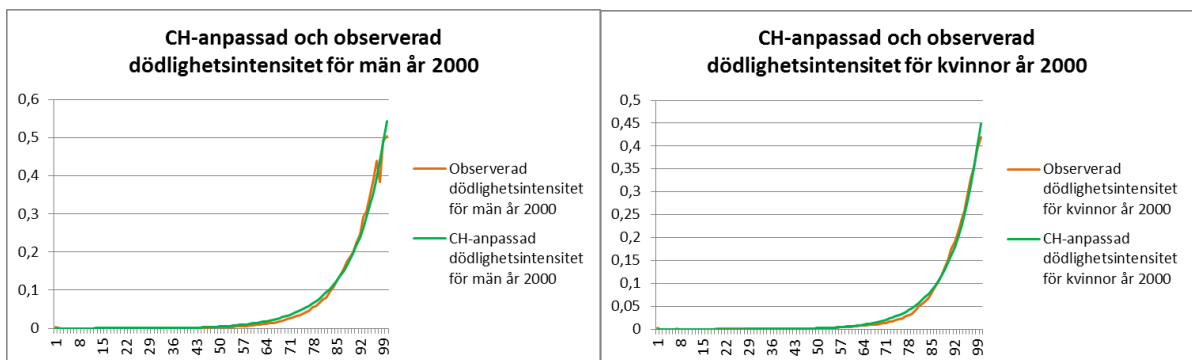


Vi observerar att vi bortsett från de uppenbara ”utstickarna” – mer vedertaget ”outliers” – kan identifiera en tydlig trend i de flesta av de sex parametrarnas utveckling. Att vi anpassat en funktion med sex parametrar ger också en förklaring till varför små avvikelser i observerade dödlighetsintensiteter kan ge ett sådant enormt utslag. Flexibiliteten i en fler-parametrig modell är givetvis större än i en få-parametrig – antal sätt att kombinera parametrarna för att få en så bra skattning, ökar givetvis när antalet parametrar ökar. I likhet med i Makeham-fallet redovisas nedan de estimerade CH-funktionerna tillsammans med observerade dödlighetsintensiteter. Vi väljer även här åren 1985, 2000 & 2016.

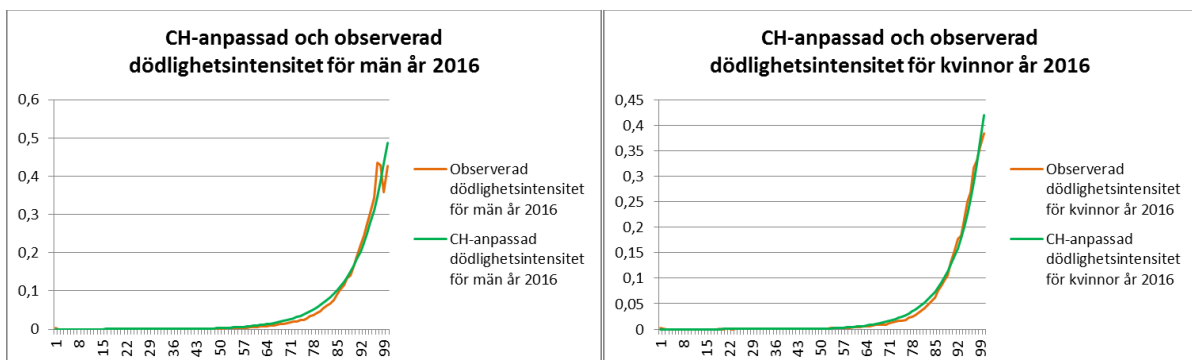
Figur 11



Figur 12



Figur 13

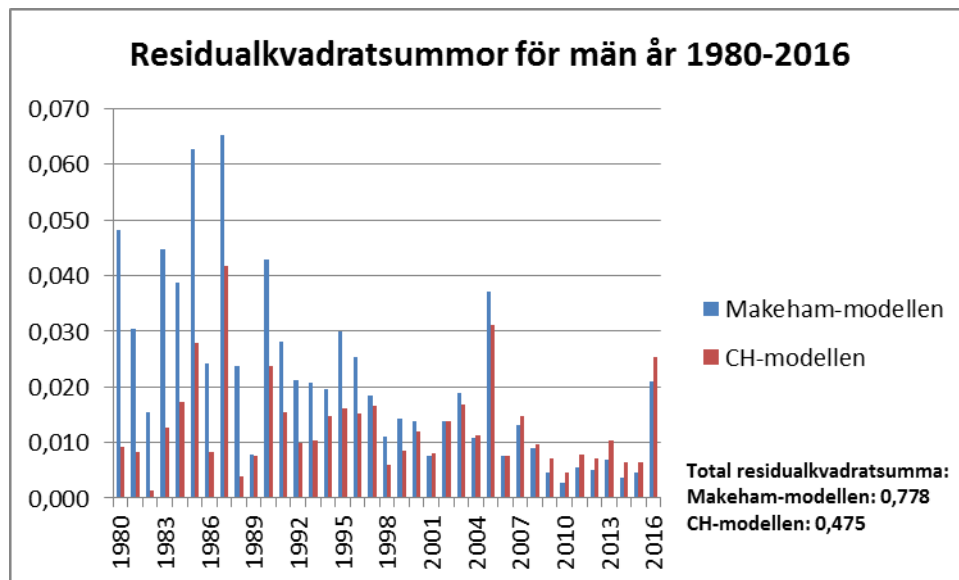


I ovan plottar, tydliggörs varför männens α_1 -parameter år 1985 givits ett så pass avvikande värde mot övriga år – nämligen den mycket ologiska upp- och nedgången i dödlighetsintensitet vid åldrarna 95-99. För att fånga förändringarna ges α_1 ett skyhögt värde vilket ställer till och ger en mycket missvisande bild av medelvärde, standardavvikelse, skevhet och kurtosis för just männens α_1 -parameter. Om vi helt enkelt stryker 1985 års α_1 -värde ges vi direkt mer rättvisande värden i samma storleksordning som kvinnornas.

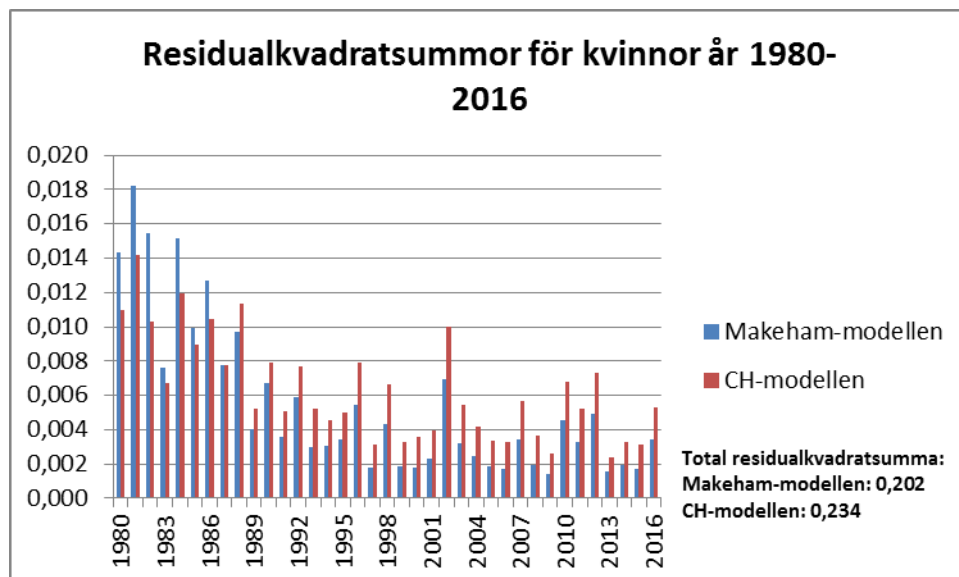
5.4 Jämförelse mellan Makeham- & CH-anpassning

Vi i analyserar i detta avsnitt skillnader i precision för modellerna. Vi presenterar i nedan plottar residualkvadratsummor för både män och kvinnor för åren 1980-2016.

Figur 14

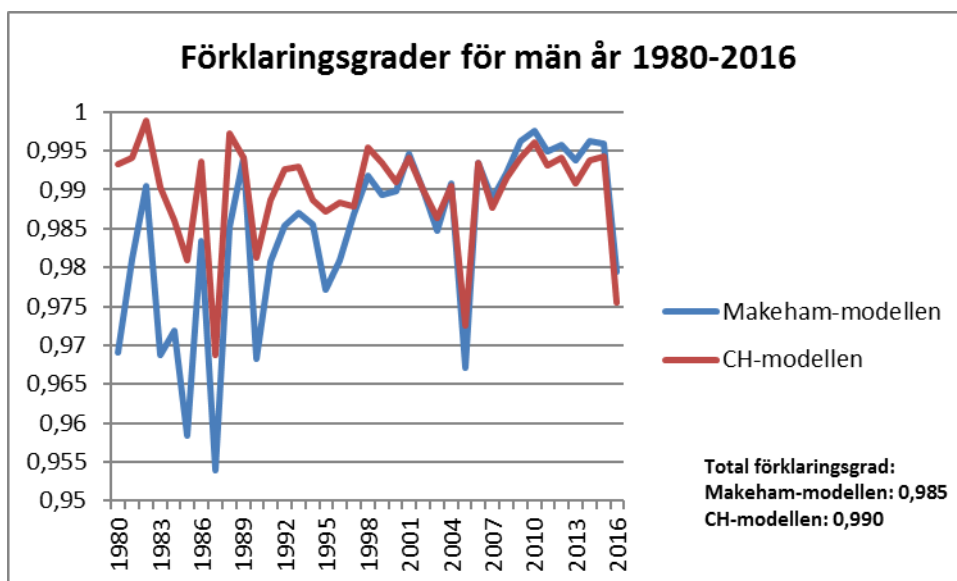


Figur 15

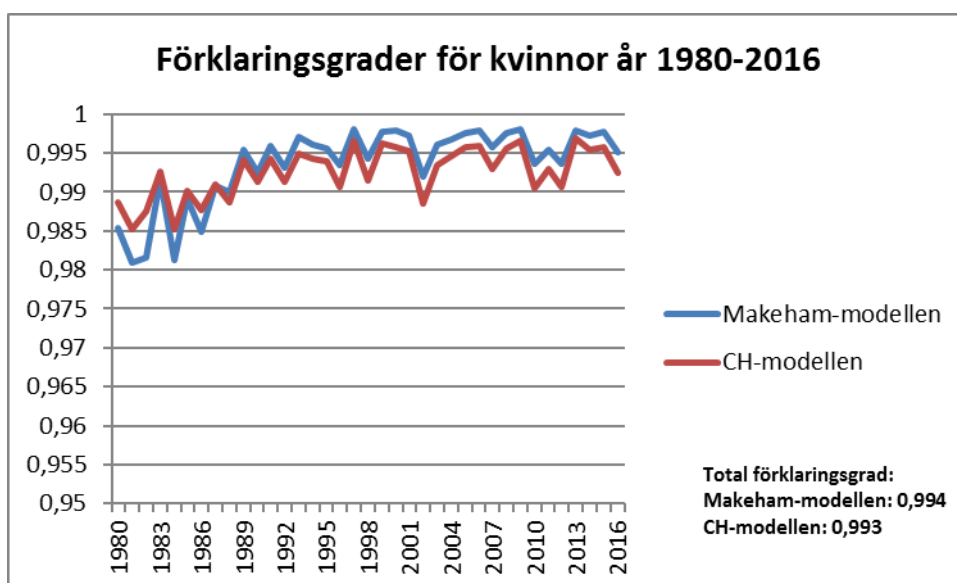


Vi ser i ovan histogram att totalt sett över alla år, 1980-2016, så har CH-modellen beskrivit observerade dödlighetsintensiteter bättre än Makeham-modellen för män. För kvinnor gäller det omvända även om skillnaden i det fallet är påfallande liten. Vi analyserar också förklaringsgraden, R^2 , i nedan plottar:

Figur 16



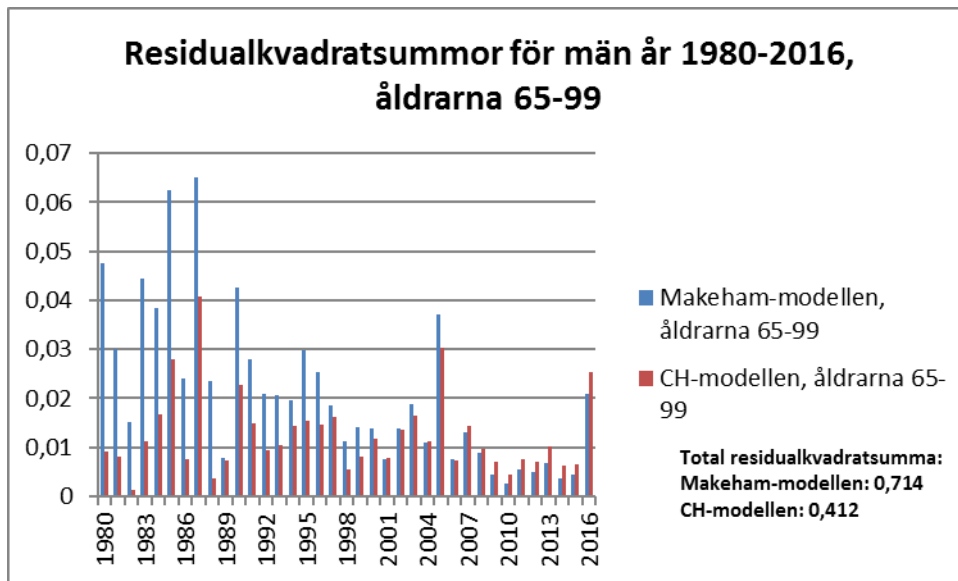
Figur 17



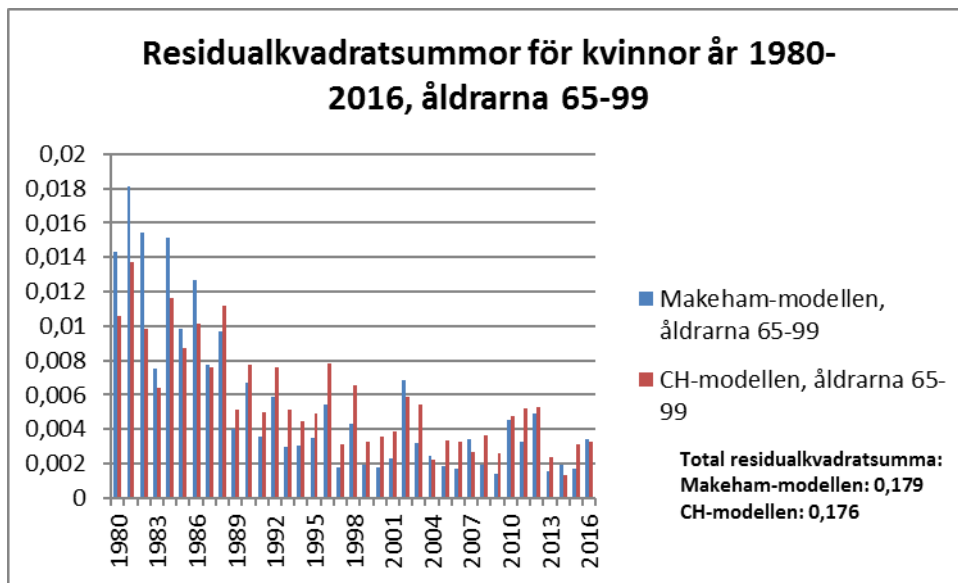
Även förklaringsgraderna visar att CH-modellen bättre än Makeham-modellen beskriver dödlighetsintensiteter för män. Det omvända gäller även här för kvinnor.

Vi analyserar också residualkvadratsummor och förklaringsgrader inom åldersspannet 65-99, vilket nog kan antas vara de viktigaste åldrarna för livförsäkringsbolagen.

Figur 18

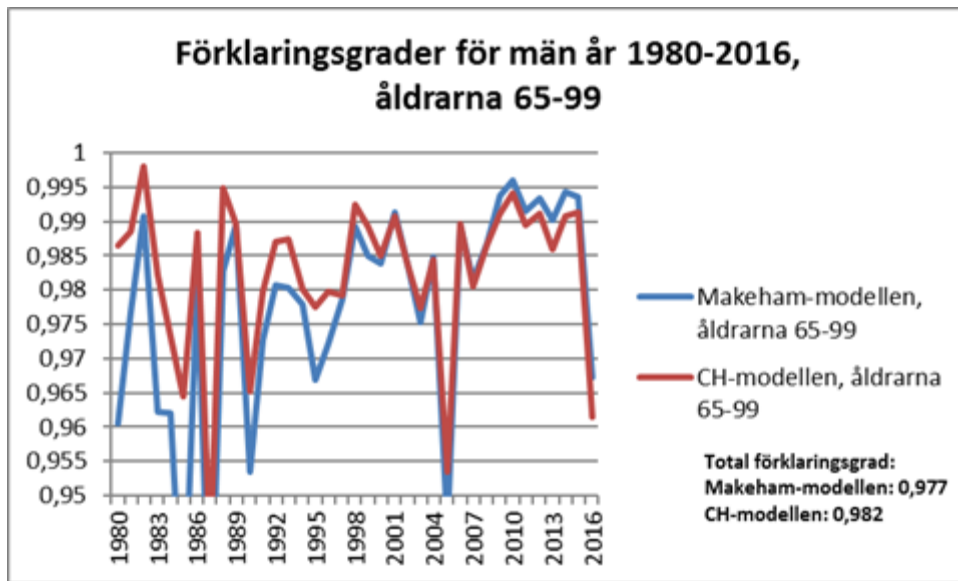


Figur 19

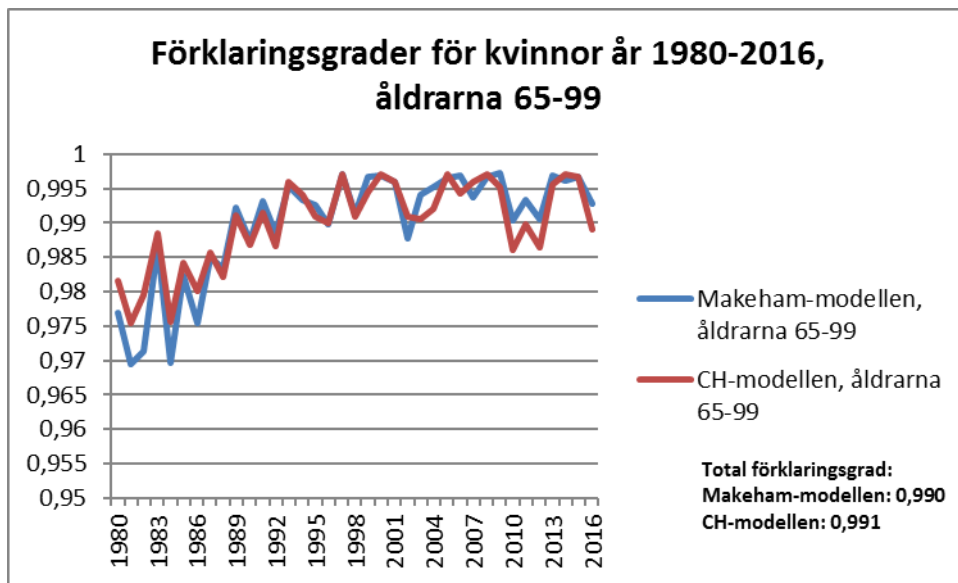


Vi ser ovan att CH-modellen här bättre beskriver observerade dödlighetsintensiteter även för kvinnor, även om skillnaden är mycket liten.

Figur 20



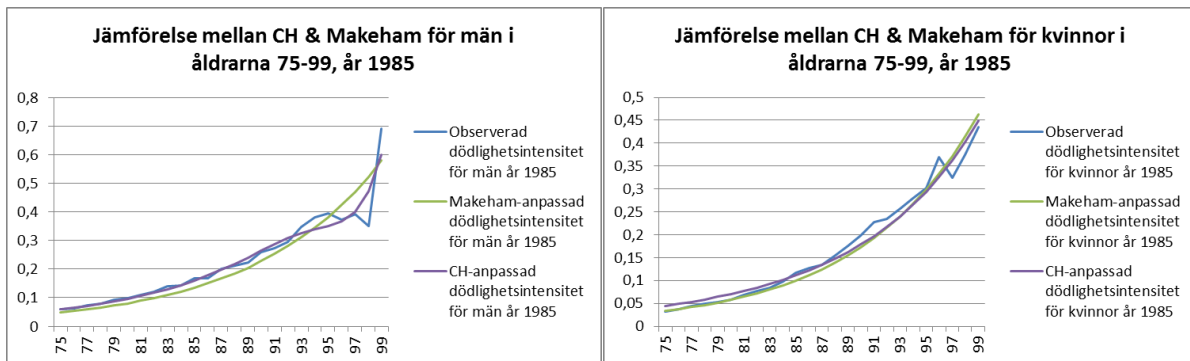
Figur 21



Vi ser ovan att förklaringsgraderna för både män och kvinnor sjunker något jämfört med fallet då alla åldrar var inkluderade, men att kvinnornas förklaringsgrad för CH-modellen nu är marginellt högre än för Makeham-modellen.

Vi presenterar i nedan plottar en grafisk jämförelse mellan Makeham-modellen & CH-modellen för åldrarna 75-99. Det vill säga de åldrar då vi identifierat ologiska mönster i våra observerade dödlighetsintensiteter. Vi väljer här att presentera jämförelse för år 1985.

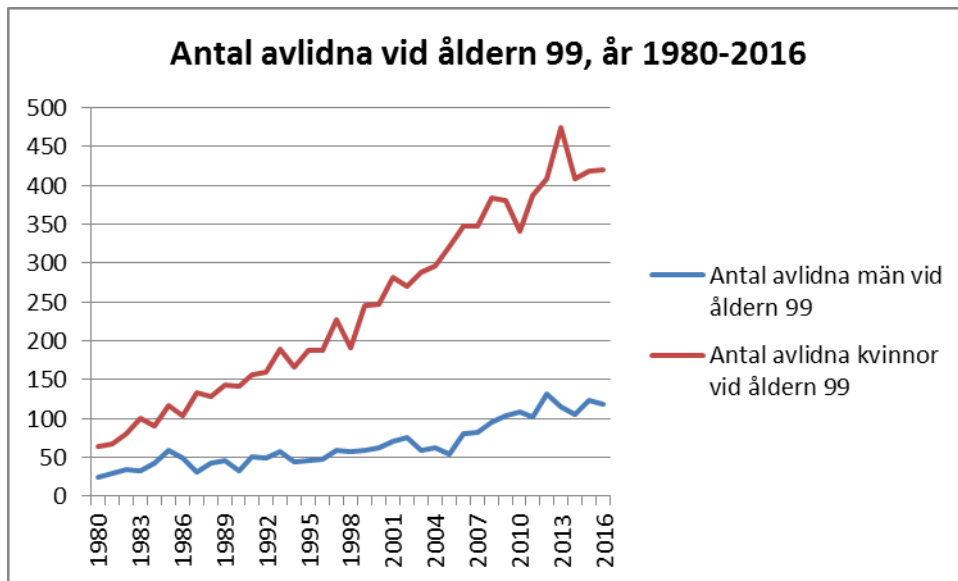
Figur 22



Vi ser i ovan plottar att CH-modellen bättre tycks fånga de ologiska mönster som tidigare beskrivits. I ovan görs det extra tydligt i männens nedgång i observerad dödlighetsintensitet där CH-modellen relativt väl fångar förändringen.

I fallen för både män och kvinnor så är det tydligt i våra histogram över residualkvadratsummor och förklaringsgrader att modellerna i allt större utsträckning liknar varandra i precisionen. Åren 1980-1995 ser vi väsentligt större skillnader mellan modellerna, samtidigt som de båda också sämre beskrev observerade dödlighetsintensiteter. Ett sådant mönster kan förklaras i termer av det som denna studie avser knyta an till – nämligen förändringar i livslängd. Svårigheter att anpassa särskilt Makeham-modeller till vår observerade data åren 1980-1995 beror troligen till stor del på att vi då haft en större statistisk osäkerhet vid särskilt höga åldrar. Att vi idag har ett mer omfattande datamaterial minskar givetvis risken för att observera den typen av märkliga mönster vi beskrivit ovan. För att åskådliggöra denna förändring visar vi nedan antal avlidna vid åldern 99 för både män och kvinnor. Tydligt är att det vid tidsaxelns början rörde sig om ett fåtal avlidna per år, varför en upp- eller nedgång med bara ett fåtal dödsfall kunde ge stort utslag i skattningarna.

Figur 23



Med bakgrund i vår jämförelse mellan den mycket väletablerade skattningstekniken Makeham, och den nu beprövade CH-modellen finns nu viss evidens för att även framgent använda oss av CH-modellen. För att kunna använda modellen för att prediktera framtida dödlighetsintensiteter – och därmed kunna konkurrera med den idag etablerade Lee-Carter-modellen presenterad i avsnitt 5.5, introducerar vi i avsnitt 5.6 av denna studie teknik för att prediktera modellens parametrar för kommande år.

5.5 Lee-Carter-modellen för framskrivning av dödlighetsintensitet

Som nämnt i avsnitt 4.2.1 modellerar vi dödlighetsintensiteten enligt följande vid användande av Lee-Carter-modellen:

$$\mu_x(t) = \exp\{\alpha_x + \kappa(t) \cdot \beta_x\}.$$

Vi logaritmerar och inför en felterm och modellerar sedermera dödlighetsintensiteter enligt:

$$\ln \left[\hat{\mu}_x(t) \right] = \ln [\mu_x(t)] + \varepsilon_x(t) = \alpha_x + \kappa(t)\beta_x + \varepsilon_x(t), \quad (28)$$

där $\varepsilon_x(t)$ representerar feltermen. Vi tolkar respektive variabel enligt nedan:

- α_x : genomsnittlig dödlighetstrend vid åldern x
- β_x : ett mått på känslighet i dödlighetsintensitet vid varje ålder x , då $\kappa(t)$ varierar

- $\kappa(t)$: den åldersoberoende dödlighetsintensiteten vid tiden t .
- $\varepsilon_x(t)$: en felterm som ska fånga variationer som ej förklaras av övriga variabler i modellen. Vidare antas att $\varepsilon_x(t) \sim N(0, \sigma_\varepsilon^2)$.

Vi använder också följande antaganden vid anpassning av modellens parametrar:

$$\begin{cases} \sum_t \kappa(t) = 0, \\ \sum_x \beta_x = 1. \end{cases} \quad (29)$$

Viktigt att uppmärksamma är att ovan antaganden, (29), i sig inte är någon inskränkning av modellen vilket diskuteras och bevisas i Andersson (2013).

Vi använder klassisk teori för minstakvadratanpassning. Vi anpassar den teoretiska modellen till observerad, historisk data och bildar följande kvadratsumma:

$$Q = \sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \varepsilon_x(t)^2,$$

vilket enligt uttryck (28) ger:

$$\sum_{t=t_{\min}}^{t_{\max}} \sum_{x=x_{\min}}^{x_{\max}} \left[\ln \left(\hat{\mu}_x(t) \right) - \alpha_x + \kappa(t) \beta_x \right]^2. \quad (30)$$

Flertalet olika tekniker för att skatta parametrarna enligt ovan skrivsätt har genom åren presenterats. Vi använder en teknik som kallas *Singular Value Decomposition* - *Singulärvärdesuppdelning* (SVD). Det är den teknik Lee & Carter (1992) själva använder i sin studie. Detaljer kring SVD-metoden beskrivs för den intresserade läsaren ingående i bilaga, avsnitt 8.2.

5.5.1 Resultat av Lee-Carter-simulering

Nedan presenterar vi ovan beskrivna parametrar α_x , β_x & $\kappa(t)$ för både män och kvinnor. Vi presenterar här α_x & β_x för åldrarna 0, 5, 10, ..., 95 och $\kappa(t)$ för åren 1980, 1985, ..., 2050. Vi avrundar till 3 decimalers precision. För fullständig tabell hänvisas till bilaga, se avsnitt 8.4.

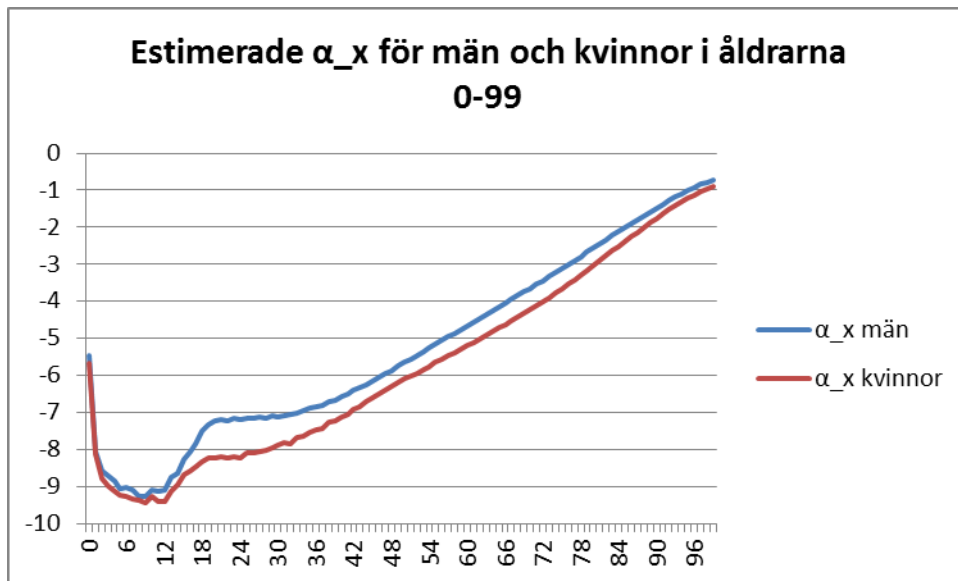
Tabell över α_x & β_x :

Tabell 3

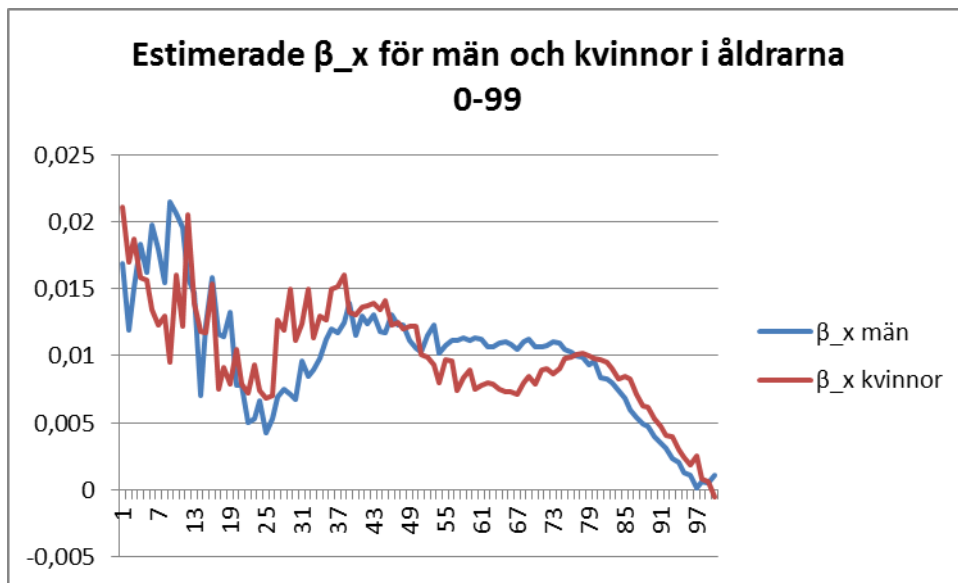
Ålder	α_x män	α_x kvinnor	β_x män	β_x kvinnor
0	-5,455	-5,656	0,017	0,021
5	-9,063	-9,252	0,020	0,013
10	-9,106	-9,255	0,020	0,012
15	-8,265	-8,682	0,016	0,015
20	-7,242	-8,230	0,008	0,008
25	-7,143	-8,101	0,005	0,007
30	-7,108	-7,899	0,010	0,012
35	-6,896	-7,549	0,012	0,015
40	-6,569	-7,134	0,013	0,014
45	-6,154	-6,589	0,013	0,012
50	-5,648	-6,090	0,010	0,010
55	-5,153	-5,649	0,011	0,010
60	-4,657	-5,194	0,011	0,008
65	-4,138	-4,711	0,011	0,007
70	-3,652	-4,230	0,011	0,009
75	-3,116	-3,675	0,010	0,010
80	-2,560	-3,028	0,008	0,010
85	-2,002	-2,376	0,006	0,008
90	-1,486	-1,761	0,004	0,005
95	-1,016	-1,223	0,001	0,002

Plottar över α_x & β_x :

Figur 24



Figur 25



Tabell över $\kappa(t)$:

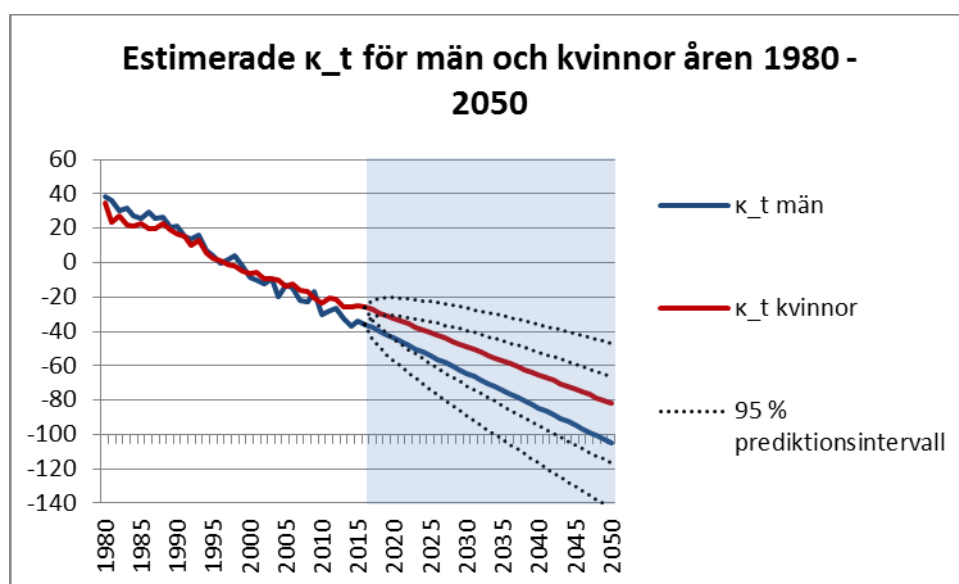
Tabell 4

År	$\kappa(t)$ män	$\kappa(t)$ kvinnor
1980	38,787	34,967
1985	25,999	22,955
1990	21,222	16,911

1995	4,137	2,211
2000	-8,844	-6,040
2005	-13,086	-14,115
2010	-30,150	-23,641
2015	-33,943	-24,944
2020	-44,051	-32,461
2025	-54,153	-40,684
2030	-64,256	-48,907
2035	-74,358	-57,129
2040	-84,460	-65,352
2045	-94,562	-73,575
2050	-104,664	-81,798

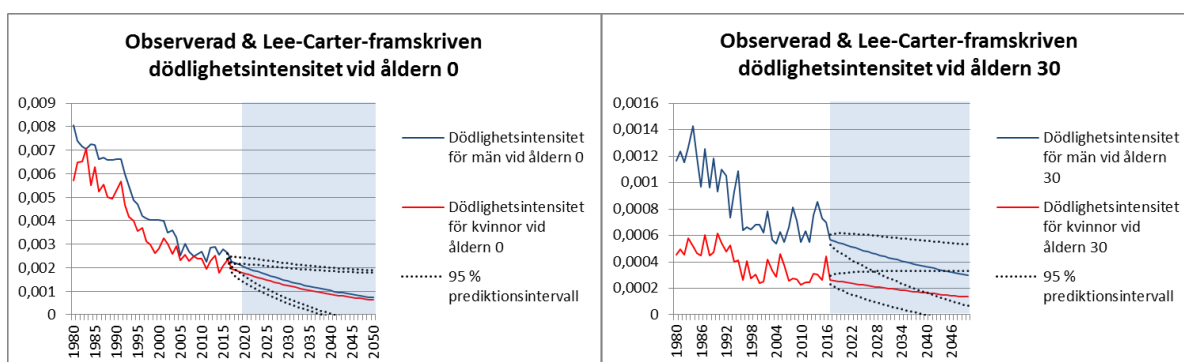
Plot över $\kappa(t)$ där skuggning avser predikerade värden:

Figur 26

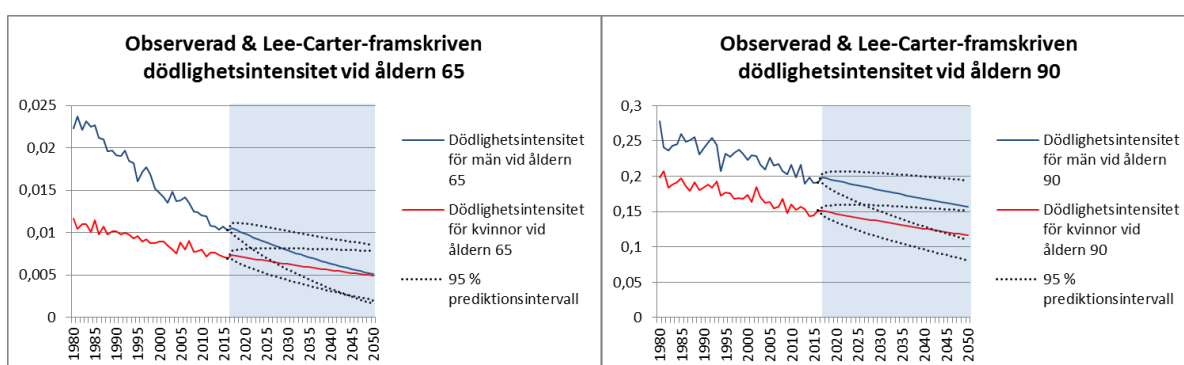


Vi redovisar nedan plottade resultat av Lee-Carter-predikerade dödlighetsintensiteter för män såväl som kvinnor. Vi redovisar också simulerade prediktionsintervall på 95 %-nivån. Vi har här valt åldrarna 0, 30, 65 och 90. Här gäller liksom ovan att skuggning avser predikerade värden.

Figur 27



Figur 28



Vi ser ovan att prediktionerna för särskilt åldern 65 visar på den i avsnitt 4.2 beskrivna nackdel som Lee-Carter-metoden för med sig – nämligen att användande av idag en idag närliggande period kan komma att generera prediktioner där människors dödlighet alltmer närmar sig eller till och med blir lägre än kvinnornas. Det föreligger också en stor statistisk osäkerhet vilket tydliggörs i prediktionsintervallen.

5.6 Framskrivning av parametrar i CH-modellen

Som nämnt i jämförelsen mellan Makeham- och CH-modellen finns viss evidens för att fortsatt använda CH-modellen för att beskriva den svenska befolkningens dödlighetsintensiteter. Vi har identifierat en anpassningsförmåga som i många fall är bättre än Makeham-modellens, med överlag låga residualkvadratsummor och höga R^2 -värden.

Vi avser i detta avsnitt, medelst tidsserieanalys i form av autoregressiva modeller prediktera framtida dödlighetsintensiteter genom att anpassa modeller till de 6 olika parametrarna som ingår i CH-modellen. Vi inskränker detta avsnitt till att enbart innefatta kvinnor.

Parameterskattningarna för kvinnor har väsentligt lägre varians än motsvarande för män – och ett mer logiskt mönster klart bättre lämpat för att kunna prediktera framtida

dödlighetsintensiteter. Mer ingående analys kring denna studies inskränkningar och begränsningar görs i avsnitt 7.

Vid analys av tidsserierna för respektive parameter stöter vi genomgående på problem med bristande stationäritet – det vill säga att vi observerar att strukturer för väntevärde, varians och autokorrelation ändras över tiden. Vi definierar därför istället förändringsgraden för varje parameter enligt nedan:

$$d(c_t) = \frac{c_t - c_{t-1}}{c_{t-1}}, \quad c_t \in \{\alpha_1(t), \beta_1(t), \gamma_1(t), \alpha_2(t), \beta_2(t), \gamma_2(t)\}. \quad (31)$$

Vi anpassar därefter autoregressiva modeller, $ar(p)$ enligt beräknat AIC (Akaike Information Criterion). Vi landar i att $ar(1)$ -modeller väl kan beskriva parametrarnas värden för respektive år och anpassar således modeller enligt nedan:

$$d(c_t) = C + \varphi \cdot d(c_{t-1}) + \varepsilon_t. \quad (32)$$

Vi utför på varje tidsserie test för stationäritet, ADF-tester och Phillips-Perron-tester som till skillnad mot den bristande stationäritet vi observerat tidigare, samtliga är signifikanta på 5 %-nivån. Vi kan därför förkasta nollhypotesen om enhetsrotens existens – det vill säga att vi kan förkasta följande hypotes:

$$d(c_t) = d(c_{t-1}) + \varepsilon_t.$$

Vi anpassar också för varje tidsserie test för autokorrelation i residualerna, Ljung-Box-test, där vi för varje parameter bortsett från α_1 kan förkasta nollhypotesen om autokorrelation. Vi använder sedermera ovan definierade förändringsgrader till nedan för att modellera tidsseriernas utveckling för övriga 5 parametrar:

$$c_t = c_{t-1} + d(c_t) \cdot c_{t-1} + \varepsilon_t = c_{t-1} \cdot [1 + d(c_t)] + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2), \quad 1981 \leq t \leq 2016. \quad (33)$$

För α_1 har vi svårt att anpassa en modell på grund av autokorrelation och svårigheter att upptäcka någon tydlig trend. Efter viss manuell hantering av så kallade ”outliers”, där vi stryker det värde som avviker mest från medelvärdet, lyckas vi dock anpassa en modell med en något ökande trend. Vi jämför också med de resultat anpassning av CH-parametrar på

amerikansk dödlighetsdata ger för värden till α_1 , och konstaterar att vår manuella hantering verkar vara fullt rimlig.

I allmänhet görs vår framskrivning enligt nedan procedur, där vi på grund av enkelhet modellerar för en parameter i taget utan hänsyn tagen till eventuell samvariation mellan de olika parametrarna:

1 steg fram:

$$d(c_{t+1}) = C + \varphi \cdot d(c_t) + \varepsilon_{t+1},$$

2 steg fram:

$$\begin{aligned} d(c_{t+2}) &= C + \varphi \cdot d(c_{t+1}) + \varepsilon_{t+1} \\ &= C + \varphi \cdot (C + \varphi \cdot d(c_t) + \varepsilon_{t+1}) + \varepsilon_{t+2} \\ &= (C + \varphi \cdot C) + \varphi^2 \cdot d(c_t) + \varphi \cdot \varepsilon_{t+1} + \varepsilon_{t+2}, \end{aligned}$$

och h steg fram ger analogt:

$$\begin{aligned} d(c_{t+h}) &= C + \varphi \cdot d(c_{t+h-1}) + \varepsilon_{t+h} \\ &= (C + \varphi \cdot C + \dots + \varphi^{h-1}) + \varphi^h \cdot d(c_t) + \varphi^{h-1} \cdot \varepsilon_{t+1} + \dots + \varepsilon_{t+h}. \end{aligned}$$

Om vi låter I_t : information vid tid t , ges vi följande väntevärden:

$$\begin{aligned} E[d(c_{t+1}) | I_t] &= C + \varphi \cdot d(c_t), \\ E[d(c_{t+2}) | I_t] &= (C + \varphi \cdot C) + \varphi^2 \cdot d(c_t), \\ E[d(c_{t+h}) | I_t] &= (C + \varphi \cdot C + \dots + \varphi^{h-1}) + \varphi^h \cdot d(c_t). \end{aligned}$$

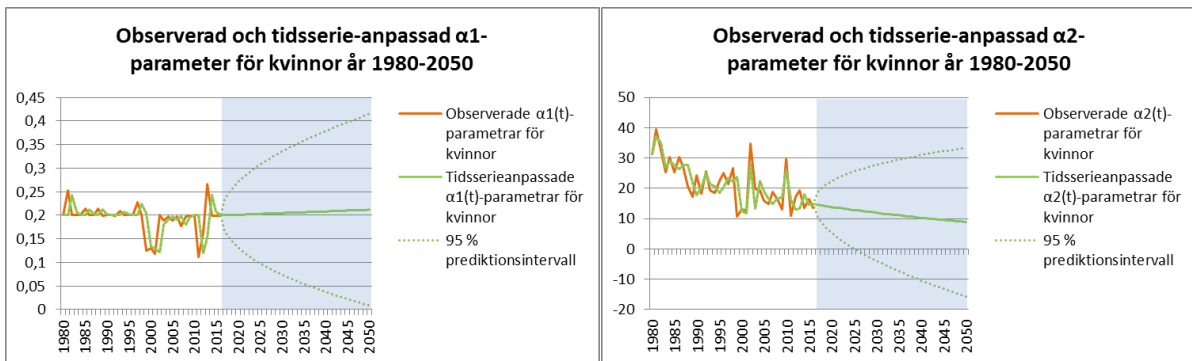
Nedan redovisas tabell för koefficienter i tidsserierna för förändringsgraden, och plottar över anpassade tidsserier till observerade parameter-värden. Vi redovisar i plottarna varje parameter tillsammans med det för varje parameter tillhörande framskrivningsintervallet på 95 %-nivån.

Tabell 5

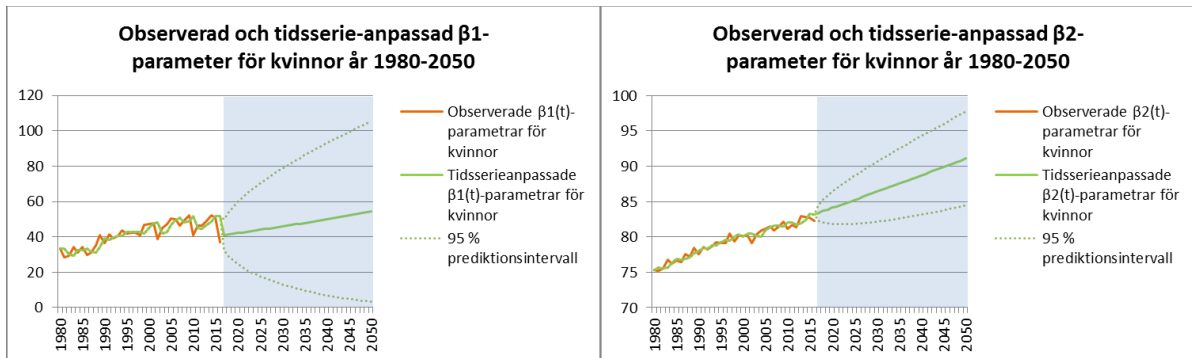
Koefficienter för förändringsgradernas ar(1)-tidsserie för kvinnor år 1981-2016

	$\alpha 1$	$\beta 1$	$\gamma 1$	$\alpha 2$	$\beta 2$	$\gamma 2$
C	0,002994	0,01165	0,01103	-0,02	0,00425	0,01258
ϕ	-0,14824	-0,36712	-0,43870	-0,34672	-0,60046	-0,6334

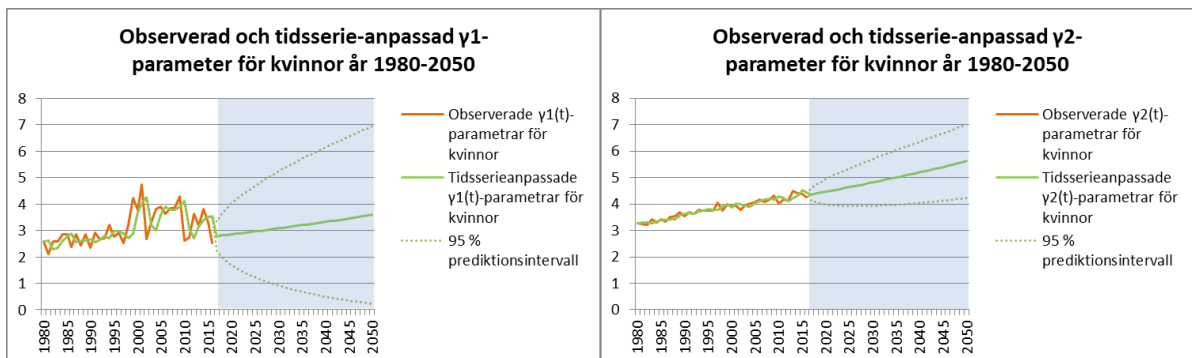
Figur 29



Figur 30

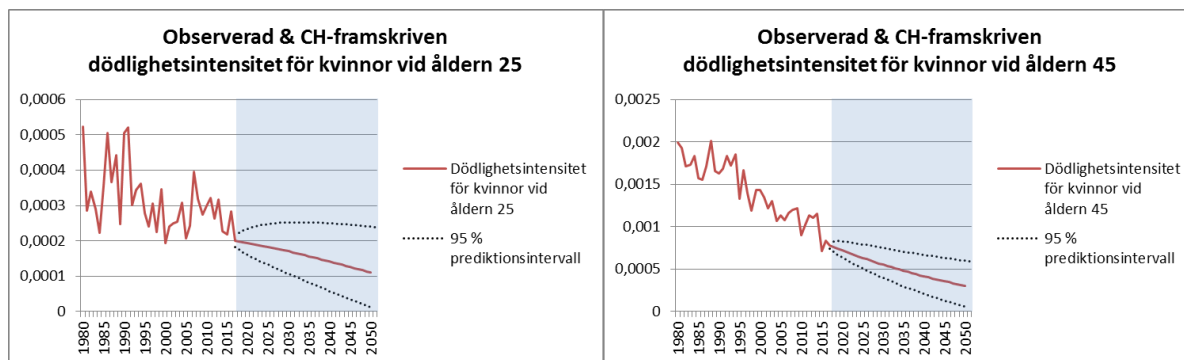


Figur 31

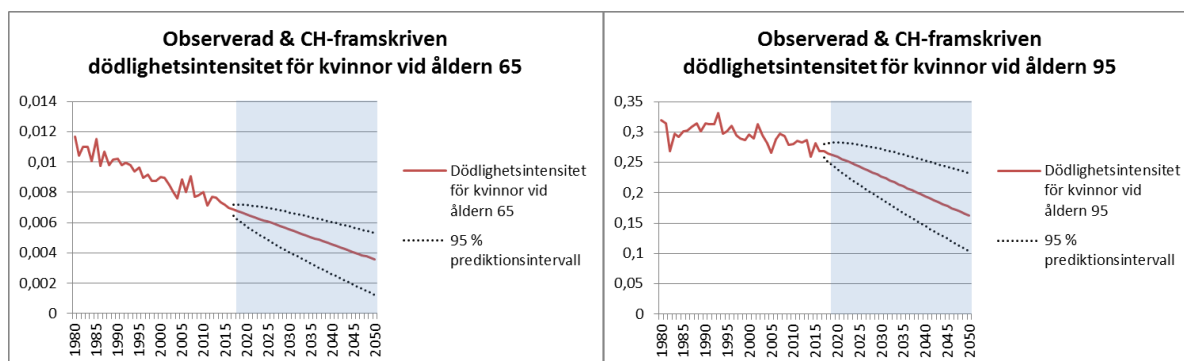


Vi beräknar nu, för varje år 2017-2050, funktioner för dödlighetsintensiteter enligt de parametervärden vi med hjälp av tidsserieanalys predikterat enligt ovan. Vi väljer i nedan plottar att redovisa observerade och CH-predikterade dödlighetsintensiteter för åldrarna 25, 45, 65 och 95.

Figur 32



Figur 33

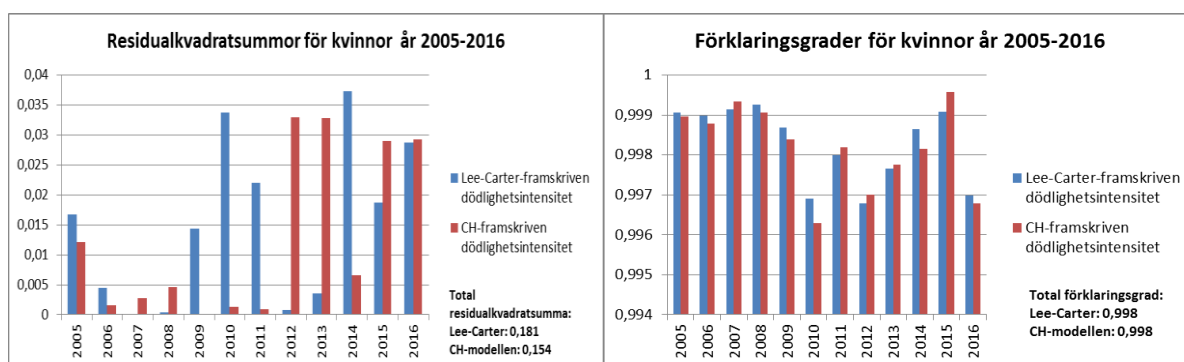


5.7 Jämförelse mellan Lee-Carter- & CH-modellens framskrivna dödlighetsintensiteter

Vi avser nu presentera en precisions-jämförelse mellan CH-modellens predikterade dödlighetsintensiteter och de motsvarande resultat Lee-Carter-modellen genererat. Vi gör detta genom att anpassa modeller enligt samma metodik som i ovan avsnitt, men här enbart utifrån åren 1980-2004. Vi använder därefter åren 2005-2016 som utvärderingsperiod.

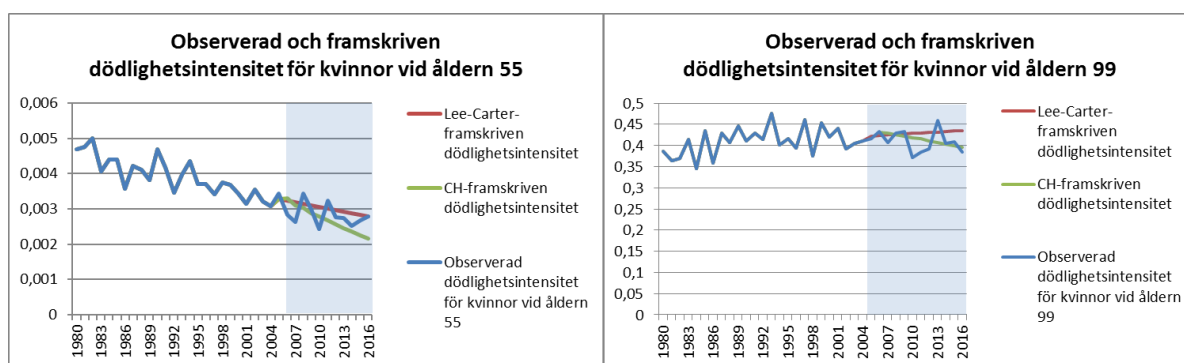
Vi går nu vidare med analys av förklaringsgrader och residualkvadratsummor för utvärderingsperioden 2005-2016. Vi presenterar residualkvadratsummor med avseende på framskrivna dödlighetsintensiteter för de båda modellerna, mot observerade dödlighetsintensiteter. Vi presenterar också förklaringsgrad mellan de framskrivna dödlighetsintensiteterna och de observerade dödlighetsintensiteterna för de båda modellerna.

Figur 34



Vi redovisar i nedan plottar ett axplock av framskrivna dödlighetsintensiteter tillsammans med observerade för perioden. Här för åldrarna 55 och 99.

Figur 35



Ovan ser vi en återkommande trend för så gott som samtliga åldrar – nämligen att CH-modellen ger en lägre dödlighetsintensitet än Lee-Carter-modellen, även om de ofta är snarlika vad gäller precision. Lee-Carter överskattar dödlighetsintensitet oftare än den underskattar, och vice versa för CH-modellen.

5.8 Omvandling av dödlighetsintensiteter till förväntade livslängder

Som nämnt under det inledande teori-avsnittet så kan vi utifrån våra framskrivna dödlighetsintensiteter enkelt approximera den ettåriga dödsrisken, överlevelsefunktionen och i förlängningen den förväntade livslängden. Vi hänvisar till ekvationerna (9), (12) och (14) och beräknar förväntad återstående livslängd enligt följande:

$$e_x \approx \left[\sum_{i=0}^{\infty} \frac{l(x+i)}{l(x)} \right] - \frac{1}{2} - \frac{1}{12} \cdot \mu_x, \quad x=1, 2, \dots$$

där

$$l(x) = l(x-1) \cdot (1 - q_{x-1}) = \prod_{i=0}^{x-1} (1 - q_i), \quad x = 1, 2, \dots$$

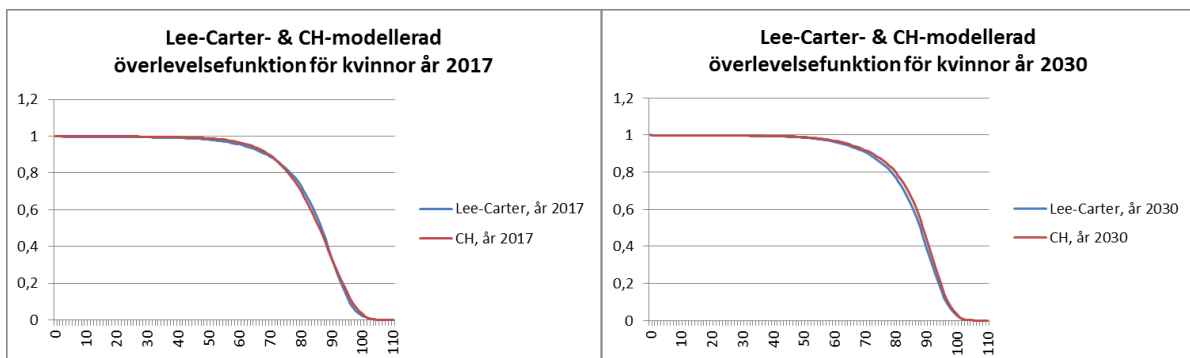
och dödsrisken approximeras enligt

$$q_x \approx \frac{\mu_{x+\frac{1}{2}}}{1 + \frac{\mu_{x+\frac{1}{2}}}{2}}$$

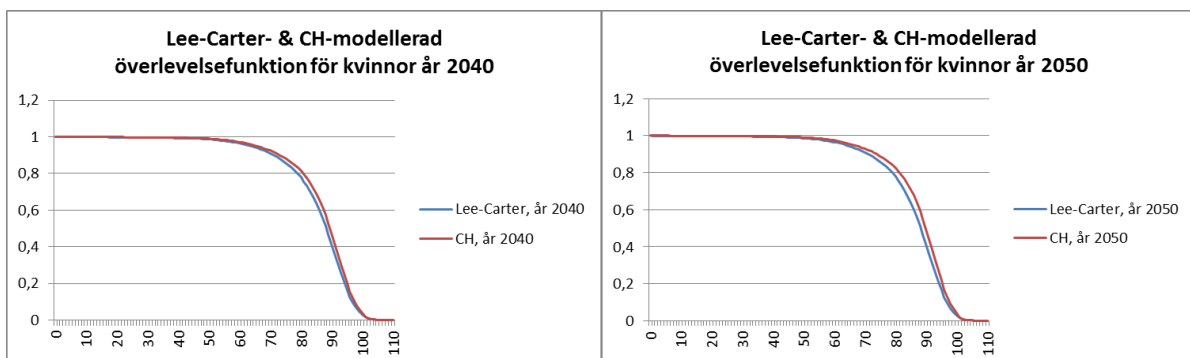
För åldrarna 100 och uppåt beräknar vi de framtida dödsriskerna enligt den årliga reduktion av dödsrisker som SCB (2016) gör gällande i rapporten *Sveriges framtida befolkning 2016-2060*. Detta då vi, som tidigare nämnt, ej har kunnat göra egna skattningar vid så pass höga åldrar på grund av stora brister i datamaterialet.

I nedan plottar åskådliggörs överlevelsefunktioner genererade från både Lee-Carter-modellen och CH-modellen för åren 2017, 2030, 2040 & 2050.

Figur 36

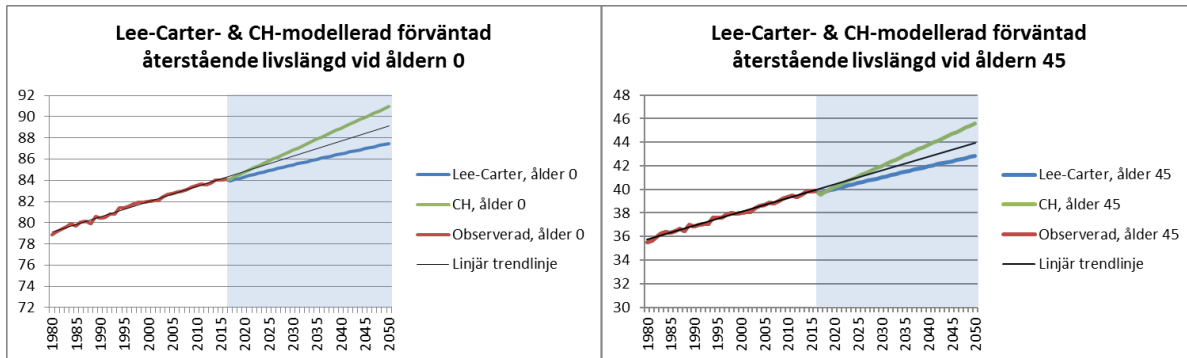


Figur 37

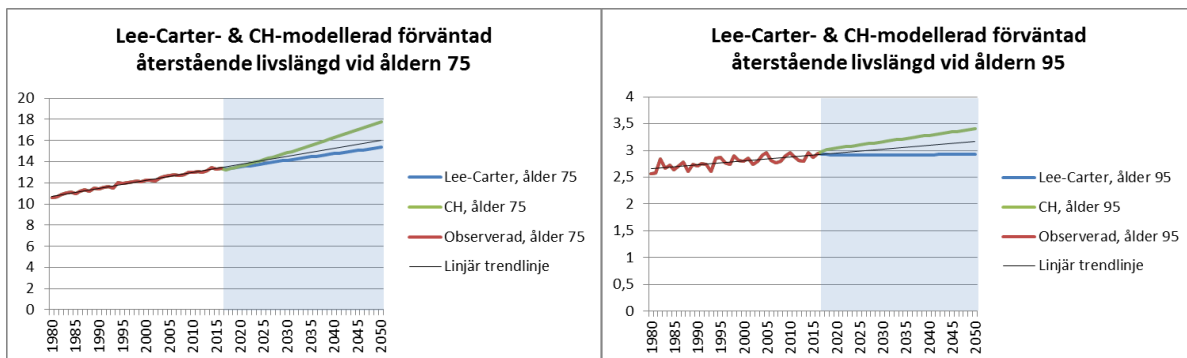


I nedan plottar åskådliggör vi nu förväntad återstående livslängd, även här för både Lee-Carter- & CH-modellen. Vi redovisar också en linjär trendlinje. Här för åldrarna 0, 45, 75 & 95.

Figur 38



Figur 39



Vi ser tydligt i ovan plottar att Lee-Carter-modellen beskriver framtida återstående livslängder som lägre, och CH-modellen som högre än den linjära trenden.

För att avslutningsvis få en bild av CH-modellens framskrivningar av återstående livslängder i termer av spridningsmått vill vi, liksom i parameter-plottarna, åskådliggöra prediktionsintervall för livslängdsskattningarna. Detta görs genom att betrakta de sex olika parametrarna som en multipel tidsserie, dock som tidigare nämnt utan hänsyn tagen till eventuell samvariation parametrarna emellan.

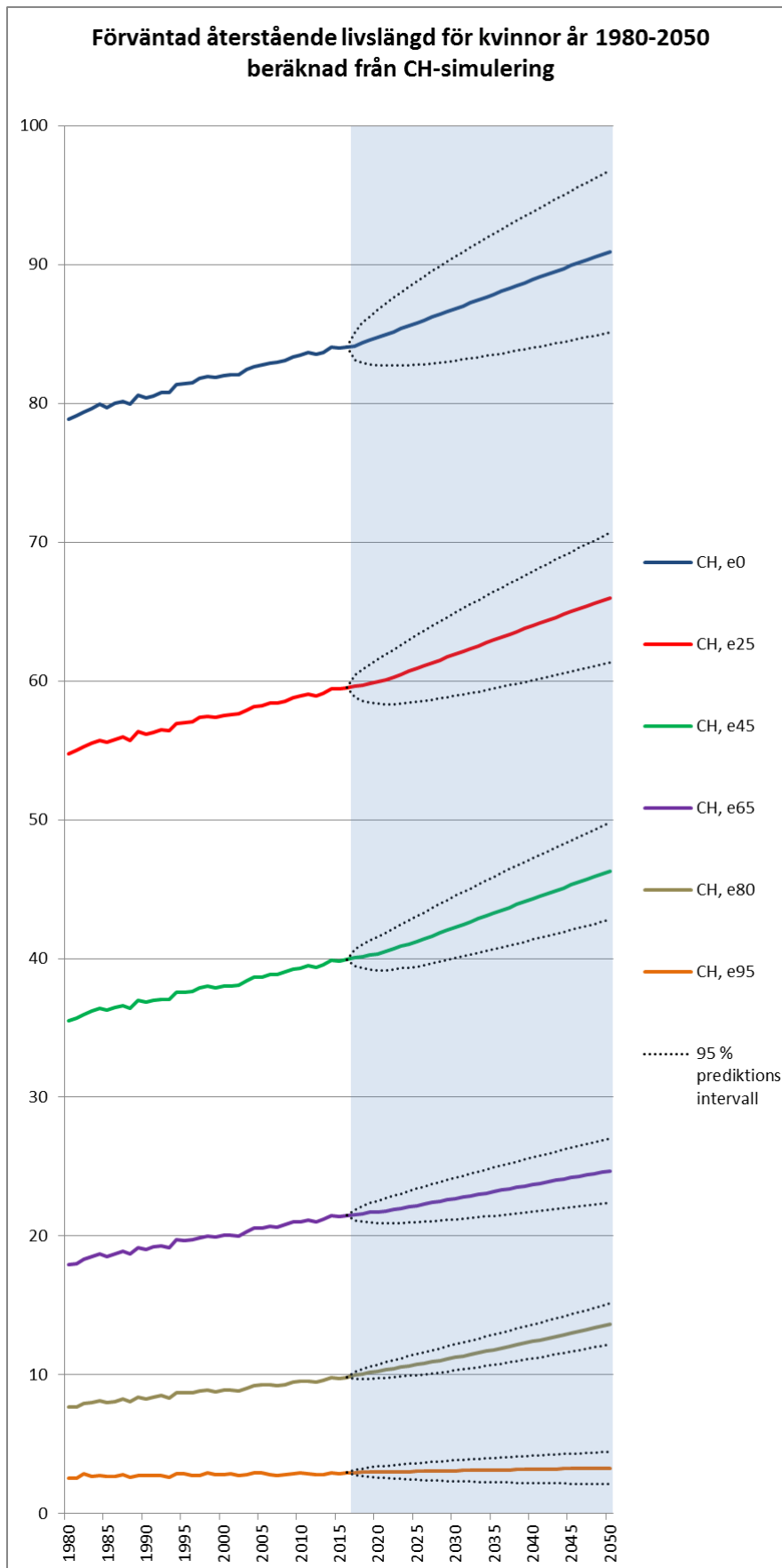
Vi har alltså följt följande procedur:

1. Anpassar autoregressiva modeller till respektive parameters förändringsgrader och applicerar test för stationaritet och autokorrelation på respektive tidsserie.
2. Beräknar respektive parameter utifrån estimerade förändringsgrader i steg 1.

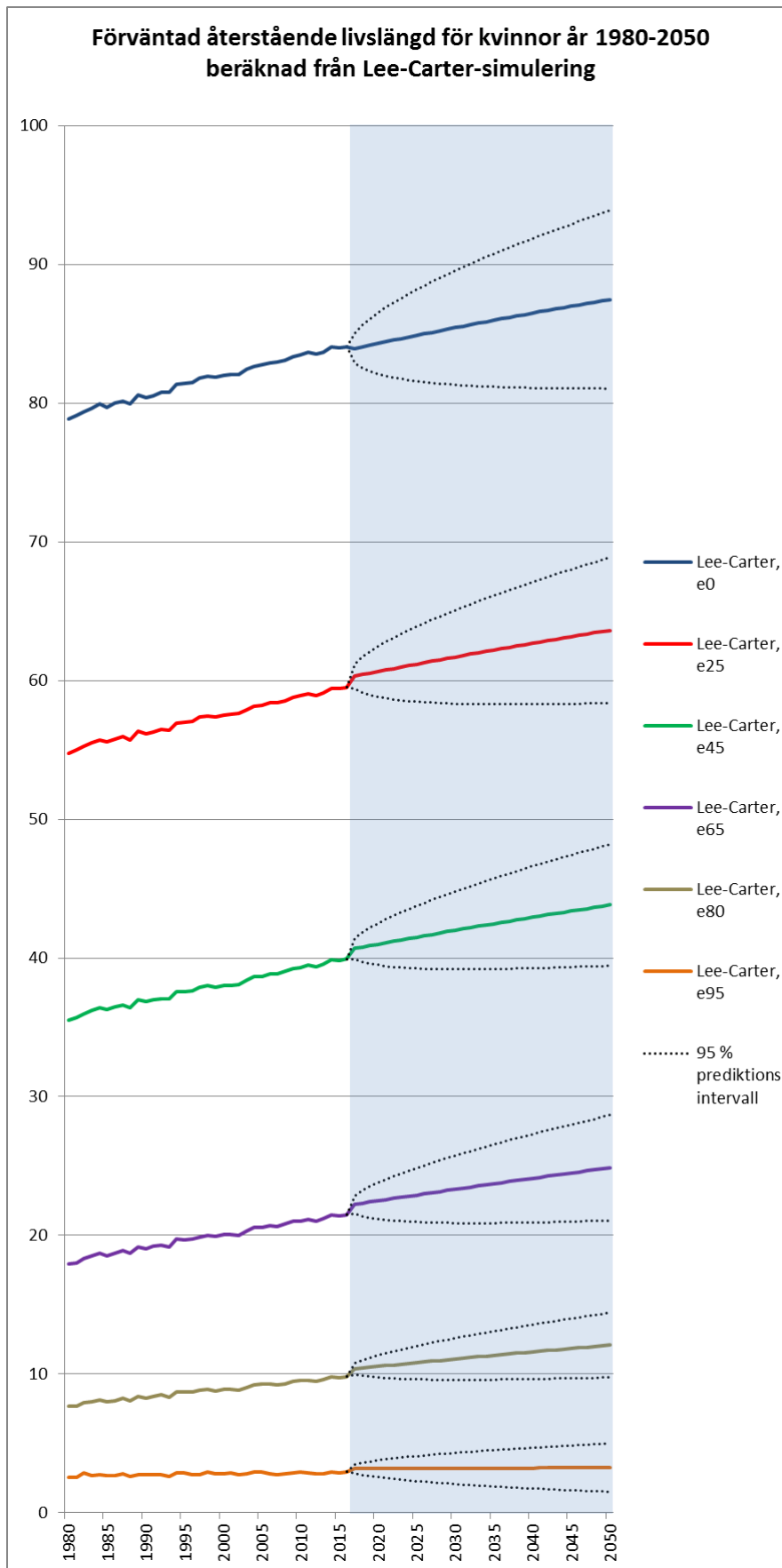
3. Simulerar framskrivna parametervärden med respektive felterm 1 000 000 gånger och beräknar utifrån resultaten konfidensintervall på 95 %-nivån.
4. Pluggar in resultaten från steg 3 i CH-funktionen för dödlighetsintensiteter.
5. Beräknar estimat av återstående livslängder.

Resultatet är således 1 000 000 skattningar av framskrivna återstående livslängder för respektive år och ålder – utifrån vilka vi enkelt kan redovisa konfidensintervall för att få en bild av den statistiska säkerheten i våra skattningar. Vi redovisar i nedan plot skattningar med tillhörande 95 % prediktionsintervall för åldrarna 0, 25, 45, 65, 80, 95. För att få en bild av skillnaden mellan CH-simulerade, framtida återstående livslängder, med dylika för Lee-Carter-modellen, visar vi i plotten efter även resultat för Lee-Carter-modellen.

Figur 40



Figur 41



Vi redovisar slutligen nedan tabell med värden från ovan plottar, med respektive prediktionsintervall för de båda modellerna.

Tabell 6

	CH-modellen			Lee-Carter-modellen		
	Prediktion 2050	Nedre 95 %	Övre 95 %	Prediktion 2050	Nedre 95 %	Övre 95 %
e0	90,946	85,302	96,717	87,459	81,147	93,873
e25	65,998	61,553	70,462	63,624	58,476	68,822
e45	46,283	42,681	49,782	43,816	39,143	48,389
e65	24,700	22,368	27,132	24,861	21,071	28,821
e80	13,659	12,201	15,017	12,086	9,793	14,508
e95	3,284	2,118	4,440	3,255	1,510	5,104

6 Slutord

Vi kan utifrån denna studies samlade innehåll konstatera att CH-modellen producerar intressanta resultat med en i många fall god anpassningsförmåga till svenskarnas dödlighetsintensitet. CH-modellen har med sin komplexitet visat sig vara annorlunda mot de väletablerade Makeham- & Lee-Carter-modellerna.

Det finns idag flertalet spännande studier som presenterar olika förslag på modeller för att beskriva olika typer av dödlighetsdata. Att utmana de väletablerade modellerna bör ses som ett naturligt och viktigt åtagande inom den aktuariatiska sfären.

7 Förslag på förbättringar & framtida studier

Sveriges i sammanhanget begränsade befolkningsmängd ger tyvärr utrymme för stor statistisk osäkerhet i särskilt höga åldrar. Som tidigare nämnt så har ett mycket litet antal dödsfall i vissa åldrar kunnat ge en stor inverkan när det gäller skattningar av dödlighetsintensitet. Vi ser också återkommande i höga åldrar fall av en, icke strikt ökande dödlighetsintensitet vilket givetvis också är en konsekvens av ett för litet statistiskt underlag. I särskilt låga åldrar har vi också påvisat exempel på hur skattningar av dödlighetsintensitet ibland är 0, då inte en enda i gällande ålder avlidit under året.

Den statistiska osäkerheten är särskilt tydlig för män, då det varit mycket ovanligt med män som når höga åldrar. Detta har resulterat i en stor spridning i skattningar av parametrarna i CH-funktionen – ett problem som är långt mycket mindre omfattande för kvinnor och också skälet till att vi inskränkt framskrivningen av dödlighetsintensiteter till att bara gälla kvinnorna. Samma analys för männen resulterar i en mycket stor osäkerhet i våra skattningar.

En idé för framtida studier av det alternativa sättet (CH-modellen) att beskriva populationens dödlighet, kan vara att ej undersöka män och kvinnor separat utan istället som helhet. Som tidigare nämnt finns evidens för att mäns och kvinnors dödlighet närmar sig varandra i allt större utsträckning, vilket kan tala för att analys av populationen som helhet kan bli mer rimlig framgent.

Istället för att utvärdera varje år 1980-2016 separat kan en idé också vara att exempelvis sammanfoga på en 2, 3 eller 5-årsbasis. Detta skulle troligen resultera i ett tydligare mönster i CH-modellens parametrar, men samtidigt skulle vi ges färre mätpunkter till vår framskrivning vilket i sin tur mynnar ut i osäkerhet. För att minska prediktionsintervallen och därmed risken för felskattningar, bör man med fördel också kunna betrakta CH-modellens parametrar som en multipel tidsserie, där hänsyn till samvariation parametrarna emellan tas i beaktning. Detta skulle högst troligen ge en mer rimlig och fördelaktig prediktionsförmåga med snävare prediktionsintervall.

Vid eventuell omvandling av våra framtida, beräknade dödlighetsintensiteter och livslängder till motsvarande inom ett specifikt försäkringsbestånd möter vi en ännu större statistisk osäkerhet då dataunderlaget blir ännu mindre. Välkänt är att försäkringsdödligheten är lägre

än befolkningsdödligheten, varför vi i det fallet sannolikt möter ett större behov av till exempel sammanfogning på köns- eller tidsnivå.

Jämförelse med i det närmaste ekvivalent dödlighetsdata, från kanske framförallt våra grannländer Norge och Island, skulle också kunna tillföra viktiga mätpunkter för att säkra vår analys i särskilt höga åldrar.

Viktigt är också att framhålla att det finns olika syften och målsättningar med att beskriva svenskarnas dödlighet. Att som i denna studie ge en demografisk överblick över de nedgångar i dödsrisk befolkningen mött, och förutspå möta i framtiden, har krävt att vi inkluderat så gott som samtliga åldrar (0-99 år) i analysen. För ett livförsäkringsbolag med syfte att göra så stora vinster som möjligt, ligger det primära syftet förmodligen istället i att så bra som möjligt beskriva dödlighet i de åldrar där deras livförsäkringsprodukter har stor inverkan på livförsäkringsbolagets resultat. Det kan därför i den typen av fall vara intressant att begränsa analysen och anpassningen av modellerna till låt säga enbart åldrarna 61-90, där syftet alltså primärt inte behöver vara att beskriva dödligheten som helhet, utan snarare den finansiella belastningen.

Den vikt $\omega_x(t)$, som vi introducerade i avsnitt 5.2 gällande minstakvadrat-skattning, skulle vid behov kunna styra vilka åldrar som ska ges högst prioritet vid anpassning av de olika modellerna. Definitionen av vikten kan givetvis också bestämmas någorlunda godtyckligt beroende på syftet. Är till exempel åldrarna 61-90 de mest intressanta och viktigaste, tillskrivs vikten höga värden i just det intervallet – och de mindre intressanta åldrarnas inverkan på minstakvadrat-skattningen begränsas.

8 Bilagor

8.1 CH-funktion för dödlighetsintensitet

Vi betraktar överlevelsefunktionen $l(x)$:

$$\begin{aligned}l(x) &= \alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)} \\ \Rightarrow \\ l'(x) &= -\theta_1'(x) \cdot \alpha_1 e^{-\theta_1(x)} - \theta_2'(x) \cdot \alpha_2 e^{-\theta_2(x)},\end{aligned}$$

som vi deriverar enligt:

$$\begin{aligned}\theta_1(x) &= \exp\left\{\left(x / \beta_1\right)^{\gamma_1}\right\} \\ \Rightarrow \\ \theta_1'(x) &= \exp\left\{\left(x / \beta_1\right)^{\gamma_1}\right\} \cdot \gamma_1 \frac{x^{\gamma_1-1}}{\beta_1^{\gamma_1}},\end{aligned}$$

respektive:

$$\begin{aligned}\theta_2(x) &= \cosh\left\{\left(x / \beta_2\right)^{\gamma_2}\right\} \\ \Rightarrow \\ \theta_2'(x) &= \frac{\gamma_2 \left(\frac{x}{\beta_2}\right)^{\gamma_2} \sinh\left\{\left(\frac{x}{\beta_2}\right)^{\gamma_2}\right\}}{x}.\end{aligned}$$

Insättning ger således:

$$\begin{aligned}l'(x) &= -\exp\left\{\left(x / \beta_1\right)^{\gamma_1}\right\} \cdot \gamma_1 \frac{x^{\gamma_1-1}}{\beta_1^{\gamma_1}} \cdot \alpha_1 e^{-\theta_1(x)} - \frac{\gamma_2 \left(\frac{x}{\beta_2}\right)^{\gamma_2} \sinh\left\{\left(\frac{x}{\beta_2}\right)^{\gamma_2}\right\}}{x} \cdot \alpha_2 e^{-\theta_2(x)} \\ \Rightarrow \\ -\frac{l'(x)}{l(x)} &= \frac{\exp\left\{\left(x / \beta_1\right)^{\gamma_1}\right\} \cdot \gamma_1 \frac{x^{\gamma_1-1}}{\beta_1^{\gamma_1}} \cdot \alpha_1 e^{-\theta_1(x)} + \frac{\gamma_2 \left(\frac{x}{\beta_2}\right)^{\gamma_2} \sinh\left\{\left(\frac{x}{\beta_2}\right)^{\gamma_2}\right\}}{x} \cdot \alpha_2 e^{-\theta_2(x)}}{\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}}.\end{aligned}$$

Vi delar nu upp och skriver om ovan enligt:

$$-\frac{l'(x)}{l(x)} = \frac{\exp\{(x/\beta_1)^{\gamma_1}\} \cdot (\gamma_1/\beta_1)(x/\beta_1)^{\gamma_1-1} \cdot \alpha_1 e^{-\theta_1(x)}}{\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}} + \frac{(\gamma_2/\beta_2)(x/\beta_2)^{\gamma_2-1} \sinh\{(x/\beta_2)^{\gamma_2}\} \cdot \alpha_2 e^{-\theta_2(x)}}{\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}}.$$

Vi förlänger respektive kvot enligt:

$$\begin{aligned} & \frac{\exp\{(x/\beta_1)^{\gamma_1}\} \cdot (\gamma_1/\beta_1)(x/\beta_1)^{\gamma_1-1} \cdot \alpha_1 e^{-\theta_1(x)}}{\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}} \cdot \frac{e^{\theta_1(x)}/\alpha_1}{e^{\theta_1(x)}/\alpha_1} \\ & + \\ & \frac{(\gamma_2/\beta_2)(x/\beta_2)^{\gamma_2-1} \cdot \sinh\{(x/\beta_2)^{\gamma_2}\} \cdot \alpha_2 e^{-\theta_2(x)}}{\alpha_1 e^{-\theta_1(x)} + \alpha_2 e^{-\theta_2(x)}} \cdot \frac{e^{\theta_2(x)}/\alpha_2}{e^{\theta_2(x)}/\alpha_2}, \end{aligned}$$

vilket vi nu kan skriva om på nedan form:

$$= \frac{\exp\{(x/\beta_1)^{\gamma_1}\} \cdot (\gamma_1/\beta_1)(x/\beta_1)^{\gamma_1-1}}{1 + (\alpha_2/\alpha_1)e^{\theta_1(x)-\theta_2(x)}} + \frac{(\gamma_2/\beta_2)(x/\beta_2)^{\gamma_2-1} \cdot \sinh\{(x/\beta_2)^{\gamma_2}\}}{1 + (\alpha_1/\alpha_2)e^{\theta_2(x)-\theta_1(x)}}.$$

Slutligen skriver vi om uttrycket enligt nedan för bättre överskådlighet:

$$\mu_x = -\frac{l'(x)}{l(x)} = \frac{\omega_1(x)}{1 + (\alpha_2/\alpha_1)e^{\theta_1(x)-\theta_2(x)}} + \frac{\omega_2(x)}{1 + (\alpha_1/\alpha_2)e^{\theta_2(x)-\theta_1(x)}},$$

där

$$\begin{cases} \omega_1(x) = (\gamma_1/\beta_1)(x/\beta_1)^{\gamma_1-1} \theta_1(x), \\ \omega_2(x) = (\gamma_2/\beta_2)(x/\beta_2)^{\gamma_2-1} \sinh\{(x/\beta_2)^{\gamma_2}\}, \\ \alpha_1, \beta_1, \gamma_1, \alpha_2, \beta_2, \gamma_2 \in \mathbb{R}_{>0}. \end{cases}$$

■

8.2 Singular Value Decomposition (SVD)

Enligt Holst & Ufnarovski (2012) gäller för metoden med singularvärdesuppdelning att alla $m \times n$ -matriser kan skrivas enligt:

$$A = USV^T,$$

där U och V är kvadratiska, unitära matriser av storlek $m \times m$ och $n \times n$ respektive. S är en diagonal, icke-negativ $m \times n$ -matris med singularvärdena till A i sin diagonal enligt:

$$S = \left(\begin{array}{cccc|c} \lambda_1 & 0 & \dots & 0 & \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & \lambda_r & \\ \hline & & & 0 & 0 \end{array} \right).$$

8.3 Matlab-kod för Lee-Carter-modellering

```
%obs använder bara upp till ålder 99 pga saknas risktid efter det%
muman=xlsread('Livslangdsdata 1980-2016',2,'D4:AN103');
t=2017;
T=2051;

deltaT=1:(T-t);
[m,n]=size(muman);
I=ones(m,n);
Ia=ones(1,n);
Ip=ones(1,(T-t));

lnmuman=log(muman);

a_x1=mean(lnmuman,2);

Z=lnmuman-(a_x1*Ia);
[U,S,V]=svd(Z);
u1=U(:,1);
v1=V(:,1);
sum_u1=sum(u1);
s1=S(1,1);

b_x1=u1/sum_u1;
k_t1=s1*sum_u1*v1';

sumb_x1=sum(b_x1);
while sumb_x1~=1;
    disp('Uppfyller ej villkoret sum(b_x1=1.)')
    disp('Rekalibrering krävs')
    break
end

sumk_t1=sum(k_t1);
while sumk_t1~=0;
    disp('Uppfyller ej villkoret sum(k_t1=0.)')
    disp('Rekalibrering krävs.')
    break
end

meank_t1=mean(k_t1);
a_x=a_x1+b_x1*meank_t1;
b_x=b_x1/sumb_x1;
k_t=(k_t1-meank_t1*Ia)*sum(b_x1);

sumb_x=sum(b_x);
sumk_t=sum(k_t);

kT=k_t(1,length(k_t));
```

```
k1=k_t(1,1);
d=(kT-k1)/length(k_t);
k_projected=kT*Ip+deltaT*d;
lnmuman_proj=a_x*Ip+b_x*k_projected;
muman_proj=exp(lnmuman_proj);

%file=input('Namn på excel-fil');
xlswrite('PrediktionMan.xlsx','Z','D4:AN103');
xlswrite('PrediktionMan.xlsx','a_x,2','D4:D103');
xlswrite('PrediktionMan.xlsx','b_x,2','E4:E103');
xlswrite('PrediktionMan.xlsx','k_t,3','D4:AN4');
xlswrite('PrediktionMan.xlsx','k_projected,4','D4:AW4');
xlswrite('PrediktionMan.xlsx','lnmuman_proj,5','D4:AW103');
xlswrite('PrediktionMan.xlsx','muman_proj,6','D4:AW103');
```


8.4 Resultat från Lee-Carter-modellering

Ålder	α_x män	α_x kvinnor	β_x män	β_x kvinnor
0	-5,4548671	0,016915	-5,656162299	0,02111459
1	-8,0651664	0,011893	-8,131279757	0,016931012
2	-8,5889035	0,015093	-8,780570906	0,018665822
3	-8,7201636	0,01828	-9,0087921	0,015843949
4	-8,8556703	0,01624	-9,138454228	0,015677102
5	-9,0628303	0,019766	-9,251911584	0,013409096
6	-9,0283645	0,017981	-9,256264398	0,012292039
7	-9,0822476	0,01542	-9,332363875	0,012998924
8	-9,2582804	0,021502	-9,359654198	0,009488239
9	-9,2547699	0,020637	-9,434897468	0,016031073
10	-9,1064534	0,019579	-9,255269441	0,012240578
11	-9,1263566	0,015869	-9,40198275	0,020546107
12	-9,1114024	0,014695	-9,391446696	0,013890048
13	-8,7614597	0,007034	-9,122119133	0,011802858
14	-8,6589809	0,012307	-8,951011889	0,011749176
15	-8,2651384	0,015826	-8,681777403	0,015377416
16	-8,047358	0,011662	-8,571923502	0,007524364
17	-7,8574293	0,011449	-8,483439857	0,009113763
18	-7,5050716	0,013265	-8,333710272	0,007931387
19	-7,3179245	0,007779	-8,224457635	0,010457692
20	-7,2422958	0,00782	-8,23001001	0,007934477
21	-7,2097842	0,004981	-8,198529212	0,007272518
22	-7,2326232	0,005312	-8,228119815	0,009301916
23	-7,1560268	0,006696	-8,189046516	0,007382788
24	-7,1876449	0,004235	-8,24394011	0,006834211
25	-7,1428207	0,005349	-8,101029477	0,006989674
26	-7,1622431	0,006921	-8,11011037	0,012719088
27	-7,1388744	0,007524	-8,072038786	0,011867875
28	-7,1510401	0,007156	-8,027385649	0,014927069
29	-7,1032548	0,006789	-7,96152719	0,011150619
30	-7,1084514	0,009611	-7,89943636	0,012396454
31	-7,0951671	0,008443	-7,823956267	0,014932678

32	-7,0584202	0,008909	-7,838297475	0,011313687
33	-7,0275938	0,009797	-7,692507118	0,012971908
34	-6,9656503	0,011273	-7,643525154	0,012679237
35	-6,8959356	0,011978	-7,548928324	0,014979404
36	-6,8481183	0,011766	-7,462674963	0,015162824
37	-6,8271506	0,01252	-7,433981836	0,016040329
38	-6,7246279	0,013918	-7,277724656	0,013218879
39	-6,6774786	0,011547	-7,214229019	0,013035099
40	-6,5690137	0,012991	-7,133946949	0,013591009
41	-6,5137508	0,012346	-7,062974106	0,013716033
42	-6,3949904	0,013017	-6,903361533	0,01389677
43	-6,3166704	0,011854	-6,857562897	0,013471232
44	-6,2421599	0,011693	-6,724366483	0,014138845
45	-6,1540657	0,013062	-6,588810863	0,012305723
46	-6,0430046	0,012339	-6,51555725	0,012484677
47	-5,9331003	0,012256	-6,409026499	0,012029559
48	-5,8752381	0,0111	-6,309206273	0,012203229
49	-5,7389586	0,010531	-6,198448963	0,012175692
50	-5,6481399	0,010288	-6,090013602	0,010110367
51	-5,5530234	0,011571	-6,021038239	0,009918655
52	-5,4612427	0,012247	-5,940102523	0,009340356
53	-5,3436995	0,010144	-5,841575871	0,008008566
54	-5,2402245	0,010717	-5,77411348	0,009733687
55	-5,1529079	0,01119	-5,648503825	0,009578286
56	-5,0467548	0,011178	-5,580260615	0,007397973
57	-4,9617913	0,011325	-5,474682712	0,008338314
58	-4,8628943	0,011174	-5,384663049	0,008928022
59	-4,7613374	0,011375	-5,28883951	0,007557381
60	-4,6567548	0,011214	-5,193503529	0,007841874
61	-4,5499818	0,010623	-5,108409358	0,007971232
62	-4,4632468	0,010641	-5,016003269	0,007921029
63	-4,3543016	0,01098	-4,922203137	0,007528218
64	-4,2430445	0,011086	-4,811196804	0,007280574
65	-4,1375562	0,010879	-4,710906634	0,007278009
66	-4,0370517	0,010444	-4,630366264	0,007111155
67	-3,9524898	0,01109	-4,537959331	0,008037792

68	-3,8489754	0,011233	-4,434165784	0,008477259
69	-3,7459539	0,010701	-4,334823326	0,007898283
70	-3,6517709	0,010646	-4,229922262	0,008921094
71	-3,5370994	0,010803	-4,118994341	0,0090715
72	-3,4429946	0,01103	-4,014918569	0,008682576
73	-3,3305489	0,01092	-3,898812446	0,009088161
74	-3,2198206	0,010505	-3,776232692	0,009822665
75	-3,116094	0,010248	-3,675491777	0,009872313
76	-3,0115282	0,009995	-3,540994875	0,010049443
77	-2,8863651	0,009939	-3,420849557	0,01014976
78	-2,7839189	0,009367	-3,298239438	0,009991784
79	-2,6749578	0,009587	-3,164329238	0,009831654
80	-2,5600413	0,00839	-3,027517	0,009694676
81	-2,4444961	0,008291	-2,903716026	0,009547716
82	-2,3425878	0,00801	-2,774205904	0,009074194
83	-2,2272298	0,007372	-2,643613647	0,008314315
84	-2,1130931	0,006822	-2,506487636	0,00848541
85	-2,0024892	0,006022	-2,376363291	0,008293144
86	-1,8976708	0,005456	-2,251943717	0,007093893
87	-1,7914582	0,004925	-2,128264263	0,006264546
88	-1,6872612	0,004744	-1,996416376	0,006168017
89	-1,5826332	0,003968	-1,869745085	0,005299778
90	-1,4855754	0,003525	-1,761125094	0,004769228
91	-1,377768	0,003149	-1,635764775	0,004108089
92	-1,2819562	0,002362	-1,525475322	0,004014297
93	-1,1763353	0,002057	-1,421503896	0,003037633
94	-1,1112303	0,001306	-1,321771609	0,002419887
95	-1,0160681	0,001064	-1,222569947	0,001900722
96	-0,942653	0,000188	-1,122483327	0,002584952
97	-0,8369591	0,000626	-1,039737977	0,000862683
98	-0,8016265	0,000522	-0,960598017	0,000651675
99	-0,7197547	0,001137	-0,893594391	-0,000531574

År	$\kappa(t)$ män	$\kappa(t)$ kvinnor	År	$\kappa(t)$ män	$\kappa(t)$ kvinnor
1980	38,787	34,967	2016	-35,969	-25,883
1981	36,453	23,841	2017	-37,990	-27,527
1982	30,493	27,083	2018	-40,010	-29,172
1983	31,405	21,861	2019	-42,031	-30,816
1984	27,451	21,025	2020	-44,051	-32,461
1985	25,999	22,955	2021	-46,072	-34,105
1986	29,204	19,636	2022	-48,092	-35,750
1987	25,610	19,681	2023	-50,113	-37,395
1988	26,359	22,434	2024	-52,133	-39,039
1989	20,834	18,959	2025	-54,153	-40,684
1990	21,222	16,911	2026	-56,174	-42,328
1991	15,288	15,426	2027	-58,194	-43,973
1992	13,904	10,390	2028	-60,215	-45,617
1993	16,120	12,955	2029	-62,235	-47,262
1994	6,712	5,330	2030	-64,256	-48,907
1995	4,137	2,211	2031	-66,276	-50,551
1996	-0,195	1,078	2032	-68,296	-52,196
1997	1,529	-1,364	2033	-70,317	-53,840
1998	3,997	-2,252	2034	-72,337	-55,485
1999	-1,766	-4,970	2035	-74,358	-57,129
2000	-8,844	-6,040	2036	-76,378	-58,774
2001	-9,864	-5,662	2037	-78,399	-60,419
2002	-12,274	-9,236	2038	-80,419	-62,063
2003	-9,664	-9,504	2039	-82,439	-63,708
2004	-19,811	-10,284	2040	-84,460	-65,352
2005	-13,086	-14,115	2041	-86,480	-66,997
2006	-14,588	-12,239	2042	-88,501	-68,641
2007	-22,392	-15,793	2043	-90,521	-70,286
2008	-22,513	-16,658	2044	-92,542	-71,931
2009	-17,122	-20,570	2045	-94,562	-73,575
2010	-30,150	-23,641	2046	-96,583	-75,220
2011	-28,052	-20,470	2047	-98,603	-76,864
2012	-26,319	-21,479	2048	-100,623	-78,509
2013	-32,280	-25,574	2049	-102,644	-80,153
2014	-36,673	-26,064	2050	-104,664	-81,798
2015	-33,943	-24,944			
2016	-35,969	-25,883			
2017	-37,990	-27,527			
2018	-40,010	-29,172			
2019	-42,031	-30,816			
2020	-44,051	-32,461			

8.5 CH-modellens parametrar

Tabell över CH-modellens parametrar med residualkvadratsummor & förklaringsgrad per respektive år, män:

MÄN	$\alpha_1(t)$	$\beta_1(t)$	$\gamma_1(t)$	$\alpha_2(t)$	$\beta_2(t)$	$\gamma_2(t)$	RSS	R2
1980	0,27361	86,97398	4,32808	0,8423	73,24731	3,56103	0,00912	0,99324
1981	0,11012	95,12451	8,28433	2,55081	73,23457	3,4979	0,00819	0,99413
1982	0,06049	94,62556	7,67921	1,2046	73,32684	3,45863	0,00149	0,99891
1983	0,03377	24,8543	3,17153	3,39402	68,79589	2,66749	0,01257	0,99038
1984	0,03622	26,40167	3,03731	2,99977	70,20113	2,79239	0,0174	0,98598
1985	1213,11	85,06368	5,43617	19,18313	94,72414	12,15563	0,02801	0,98093
1986	0,04438	26,16718	2,80281	3,46586	70,72908	2,8607	0,00831	0,99360
1987	0,03562	26,20081	2,84232	2,99989	69,81529	2,77588	0,04182	0,96877
1988	0,08425	87,22559	4,69495	0,10702	75,13226	4,09635	0,00383	0,99725
1989	0,21285	72,00249	2,11666	181,3726	72,65438	3,07818	0,0077	0,99412
1990	0,03343	26,23528	2,88545	2,99998	70,73393	2,84725	0,02367	0,98124
1991	0,05771	28,12081	2,89336	4,68764	71,59303	2,96509	0,01535	0,98867
1992	0,1766	27,60353	2,20197	16,57651	71,95719	2,99654	0,00987	0,99263
1993	0,09258	97,48689	11,52953	4,19894	75,86918	3,85943	0,01047	0,99299
1994	0,03786	31,85699	2,57145	2,90317	73,39721	3,15398	0,01472	0,98862
1995	0,02507	26,30769	2,72528	2,99998	71,70789	2,93887	0,01607	0,98716
1996	0,02435	27,52313	2,84927	2,99998	72,2421	3,01024	0,01518	0,98827
1997	0,0386	31,48179	2,58178	3,28857	74,20814	3,28126	0,01654	0,98794
1998	0,03105	27,23895	2,426	2,97847	73,69474	3,18129	0,00593	0,99545
1999	0,02992	28,01344	2,70165	3	73,37781	3,1434	0,0086	0,99344
2000	0,03668	28,66311	2,72628	3,02933	74,25806	3,26849	0,01206	0,99099
2001	0,04037	31,3174	2,42741	3,00001	74,88197	3,36525	0,00807	0,99419
2002	0,07926	28,37809	2,89653	7,05561	74,48849	3,31442	0,01395	0,98998
2003	0,19609	27,07221	3,01614	20,95879	74,30323	3,23887	0,0169	0,98637
2004	0,20657	29,784	2,59897	16,24913	75,90877	3,44567	0,01137	0,99051
2005	0,24267	27,1516	3,26316	28,02258	73,54365	3,0975	0,03106	0,97246
2006	0,21027	29,76621	2,38517	18,88812	75,72101	3,39765	0,00764	0,99346
2007	0,20829	29,35666	2,75216	17,34718	75,83764	3,42826	0,01469	0,98774
2008	0,219	30,26491	3,05034	17,84705	76,00011	3,42711	0,00973	0,99151
2009	0,22022	33,11237	2,54175	15,39984	77,13463	3,63341	0,00714	0,99418
2010	0,21395	32,37883	2,67835	15,84802	77,47515	3,66304	0,00457	0,99610
2011	0,22784	30,98442	2,67597	18,98714	77,10231	3,57528	0,00773	0,99306

2012	0,219	34,56205	2,52955	14,99918	77,94326	3,76995	0,00718	0,99416
2013	0,21871	33,17967	2,93192	15,60531	77,61292	3,66334	0,01034	0,99082
2014	0,21141	34,02296	2,84966	14,60691	78,08783	3,69277	0,00638	0,99382
2015	0,23031	33,4289	2,63275	15,00354	78,83015	3,88413	0,00644	0,99435
2016	0,22333	34,09351	2,91606	16,84274	77,56112	3,60809	0,02535	0,97553

Tabell över CH-modellens parametrar med residualkvadratsumma & förklaringsgrad per respektive år, kvinnor:

KVINNOR	$\alpha_1(t)$	$\beta_1(t)$	$\gamma_1(t)$	$\alpha_2(t)$	$\beta_2(t)$	$\gamma_2(t)$	RSS	R2
1980	0,19996	33,18536	2,57148	31,38729	75,32551	3,25734	0,01098	0,98870
1981	0,25206	28,39856	2,10543	39,36263	75,17564	3,22721	0,01417	0,98513
1982	0,20000	29,29930	2,56988	32,58345	75,58621	3,21286	0,01027	0,98759
1983	0,19993	34,26186	2,57384	25,30647	76,77605	3,42428	0,00671	0,99263
1984	0,20000	31,16499	2,86407	30,22376	76,26576	3,29148	0,01192	0,98522
1985	0,21471	34,18339	2,84864	25,35690	76,70687	3,41947	0,00899	0,99019
1986	0,20000	29,94818	2,36338	30,23596	76,39382	3,33014	0,01044	0,98765
1987	0,20000	31,33997	2,85617	27,26920	77,50248	3,50609	0,00778	0,99095
1988	0,21461	35,34624	2,44294	20,63327	77,19134	3,53658	0,01136	0,98864
1989	0,19979	40,87332	2,85529	17,22826	78,46135	3,67790	0,00524	0,99413
1990	0,19991	36,29305	2,35292	24,32915	77,57158	3,54177	0,00790	0,99133
1991	0,19987	41,27516	2,92002	18,16153	78,54395	3,69745	0,00511	0,99434
1992	0,19955	39,33107	2,66397	25,66146	78,25734	3,63477	0,00771	0,99124
1993	0,20854	40,33805	2,68780	19,29145	78,67945	3,78546	0,00524	0,99485
1994	0,19990	43,66612	3,22077	18,55584	79,21403	3,73872	0,00455	0,99431
1995	0,19990	41,67552	2,76773	22,55770	79,13446	3,73575	0,00497	0,99389
1996	0,19990	42,29465	2,89818	25,05463	79,10324	3,75304	0,00791	0,99064
1997	0,22820	42,12599	2,53430	21,51735	80,50013	4,05698	0,00313	0,99653
1998	0,19988	40,78635	3,20166	26,50971	79,37291	3,74679	0,00664	0,99144
1999	0,12555	46,88622	4,23681	10,62817	80,17276	3,97540	0,00331	0,99625
2000	0,13076	47,06422	3,78626	13,16569	80,12551	3,93263	0,00361	0,99571
2001	0,11908	47,65451	4,74390	12,71708	80,13699	3,93869	0,00395	0,99532
2002	0,19989	38,79704	2,68101	34,62081	79,17345	3,77560	0,00998	0,98850
2003	0,18859	45,14717	3,29465	19,85061	80,34531	3,96416	0,00543	0,99344
2004	0,19774	47,06552	3,81956	19,31703	80,86510	4,02282	0,00421	0,99454
2005	0,18958	50,32895	3,88570	15,88940	81,17085	4,08660	0,00332	0,99570
2006	0,19717	50,00892	3,62066	14,97519	81,44796	4,17322	0,00329	0,99593

2007	0,17718	46,52774	3,83235	18,83200	80,88940	4,06171	0,00569	0,99303
2008	0,19804	49,54298	3,85275	16,49708	81,44708	4,17901	0,00363	0,99557
2009	0,19955	52,23208	4,28204	13,14733	82,16616	4,30210	0,00262	0,99661
2010	0,19970	41,09261	2,61796	29,73336	81,09677	4,01373	0,00676	0,99052
2011	0,11169	46,29850	2,73520	10,91263	81,73439	4,17166	0,00525	0,99290
2012	0,16298	46,55863	3,62539	17,12922	81,41710	4,14208	0,00734	0,99062
2013	0,26637	49,00646	3,16522	19,24467	82,93873	4,50166	0,00238	0,99695
2014	0,19961	52,18291	3,80322	13,55556	82,81412	4,39873	0,00332	0,99538
2015	0,19957	50,66496	3,31270	16,47767	82,69342	4,38502	0,00314	0,99572
2016	0,09845	36,92766	2,51123	13,60466	82,28007	4,25511	0,00532	0,99242

8.6 Dödlighetsintensitet för svenska män

Dödlighetsintensitet för svenska män i åldrarna 0, 5, 10, ..., 95, år 1980, 1985, ..., 2015

	1980	1985	1990	1995	2000	2005	2010	2015
0	0,008063	0,007212	0,006622	0,004698	0,00404	0,002517	0,002694	0,00279
5	0,000252	0,000301	0,000137	0,000142	5,41E-05	0,000211	1,84E-05	4,88E-05
10	0,000161	0,000108	0,000158	7,59E-05	9,41E-05	7,08E-05	4,07E-05	5,28E-05
15	0,000332	0,000376	0,000356	0,000192	0,000281	0,000277	0,000171	0,000134
20	0,000969	0,000956	0,000804	0,000816	0,000874	0,000587	0,000507	0,00064
25	0,001226	0,001129	0,000702	0,000641	0,000769	0,000797	0,000767	0,000891
30	0,001162	0,001188	0,000929	0,001087	0,000681	0,000626	0,00055	0,000727
35	0,001813	0,001223	0,001019	0,001044	0,000912	0,000911	0,000661	0,000936
40	0,002491	0,002093	0,001676	0,001667	0,001153	0,000994	0,001076	0,001024
45	0,003794	0,002928	0,002745	0,0025	0,002041	0,001795	0,00133	0,001243
50	0,006007	0,004982	0,003713	0,003785	0,002909	0,002693	0,002844	0,002266
55	0,008816	0,008027	0,007207	0,006056	0,005065	0,005113	0,004342	0,004023
60	0,015962	0,013124	0,011393	0,010074	0,008826	0,008123	0,00728	0,006537
65	0,022261	0,022664	0,019151	0,016087	0,014744	0,013852	0,012006	0,010726
70	0,037745	0,033999	0,034113	0,028102	0,02591	0,022278	0,020162	0,017034
75	0,061984	0,059035	0,05671	0,04994	0,043043	0,038633	0,034582	0,030824
80	0,106774	0,098161	0,089388	0,082127	0,076694	0,069092	0,062516	0,057604
85	0,169016	0,166407	0,147974	0,143741	0,135611	0,130823	0,112466	0,108701
90	0,277555	0,259331	0,238187	0,232278	0,222503	0,225788	0,216056	0,190146
95	0,36971	0,393878	0,355357	0,394698	0,390995	0,364292	0,327791	0,371975

8.7 Dödlighetsintensitet för svenska kvinnor

Dödlighetsintensitet för svenska kvinnor i åldrarna 0, 5, 10, ..., 95, år 1980, 1985, ..., 2015

	1980	1985	1990	1995	2000	2005	2010	2015
0	0,005699	0,006287	0,005268	0,003565	0,002827	0,002332	0,002382	0,002095
5	0,000189	0,000126	0,000145	0,000133	7,56E-05	4,45E-05	5,82E-05	0
10	0,000113	0,000208	0,000207	8,05E-05	9,92E-05	7,43E-05	4,29E-05	0,000204
15	0,000348	0,000113	0,00028	0,000181	0,000139	0,000179	0,000108	6,12E-05
20	0,000532	0,000328	0,000346	0,000326	0,000237	0,000309	0,000219	0,000256
25	0,000523	0,000355	0,000505	0,00028	0,000195	0,000206	0,0003	0,000218
30	0,000455	0,000468	0,000617	0,000413	0,000241	0,000461	0,000227	0,000263
35	0,000711	0,000833	0,00046	0,000477	0,000533	0,000403	0,000278	0,000249
40	0,001564	0,000958	0,000845	0,000837	0,000617	0,000632	0,000603	0,000521
45	0,001993	0,001574	0,001628	0,001329	0,001432	0,001137	0,000905	0,000718
50	0,003001	0,003009	0,002624	0,002358	0,002419	0,002342	0,001683	0,001544
55	0,00469	0,004397	0,004691	0,003697	0,003446	0,003432	0,002422	0,002674
60	0,007398	0,006414	0,006135	0,00561	0,005831	0,005301	0,004938	0,004653
65	0,01167	0,011512	0,01021	0,009644	0,00899	0,008864	0,008043	0,007192
70	0,0201	0,018063	0,016871	0,014898	0,01338	0,013552	0,012028	0,011015
75	0,036686	0,032797	0,029262	0,0269	0,023778	0,021787	0,020848	0,021109
80	0,06845	0,058487	0,05758	0,051107	0,049758	0,042355	0,037732	0,037409
85	0,125946	0,117243	0,101122	0,095873	0,092246	0,083154	0,077303	0,074752
90	0,197855	0,197545	0,18414	0,176669	0,173505	0,164021	0,160185	0,144019
95	0,319664	0,301554	0,314494	0,300514	0,295898	0,265394	0,280607	0,282158

9 Litteraturförteckning

- Andersson, G., 2013. *Livförsäkringsmatematik*. 2 red. Stockholm: Svenska Försäkringsföreningen.
- Brouhns, N., Denuit, M. & van Keilegom, I., 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, Volym III, pp. 212-224.
- Europeiska kommissionen, 2012. *Factsheet: EU rules on gender-neutral pricing in insurance*. [Online] Available at: http://europa.eu/rapid/press-release_MEMO-12-1012_en.htm [Använd 15 03 2017].
- Folkhälsomyndigheten, 2016. *Folkhälsan i Sverige 2016*, Stockholm: Folkhälsomyndigheten.
- Holst, A. & Ufnarowski, V., 2012. *Matrix Theory*. LUND: KFS AB.
- Lee, R. D. & Carter, L. R., 1992. Modeling and Forecasting U.S Mortality. *Journal of the American Statistical Association*, Volym 87, pp. 659-671.
- Lundström, H. & Qvist, J., 2004. Mortality Forecasting and Trend Shifts: an application of the Lee-Carter model to Swedish Mortality Data. *International Statistical Review*, Volym 72, pp. 37-50.
- Pensionsmyndigheten, 2012. *Underlag för Standard för pensionsprognoser*. [Online] Available at: <https://secure.pensionsmyndigheten.se/download/18.361b851614f842fc4f51fod9/1495785612351/Underlag+till+Standard+f%C3%B6r+pensionsprognoser+2+o.pdf> [Använd 27 03 2017].
- SCB Statistik databasen, 2017. *Ettårig livlängdstabell för hela riket efter kön och ålder. År 1980-2017*. [Online] Available at: http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START_BE_BE0101_BE0101I/LivslangdEttariga/?rxid=df800e4e-3e96-4doc-a79d-do661b6afbce# [Använd 02 03 2017].
- Statistiska Centralbyrån (SCB), 2012. Medellivslängden ökar stadigt. *Välfärd*, 03 10, pp. 13-17.
- Statistiska Centralbyrån (SCB), 2016. *Sveriges framtida befolkning 2016-2060*, Stockholm: SCB.
- Wong, C. H. & Tsui, A. K., 2015. Forecasting life expectancy: Evidence from a new survival function. *Insurance: Mathematics and Economics*, Volym 65, pp. 208-226.