

SCB:s modell för befolkningsprognoser

En dokumentation

SCB's model for population projections

A documentation

Statistics Sweden
2005

Producent
Producer SCB, Enheten för demografisk analys och jämställdhet
Box 24300, 104 51 Stockholm
Tfn +46 8 506 940 00
E-post: demografi@scb.se

Förfrågningar
Inquiries Jan Qvist, tfn +46 8 506 943 73
E-post: jan.qvist@scb.se

Hans Lundström, tfn +46 8 506 943 70
E-post: hans.lundstrom@scb.se

Torbjörn Israelsson, tfn +46 8 506 948 93, ger ytterligare upplysningar
om verksamheten inom statistikenheten.

Om du citerar ur denna publikation, var god uppge källan:
SCB, SCB:s modell för befolkningsframskrivningar. Bakgrundsmaterial om demografi, barn och familj
2005:1.

©2005, Statistiska centralbyrån

Enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk är det förbjudet att helt eller
delvis mångfaldiga innehållet i denna publikation utan medgivande från Statistiska centralbyrån.

ISSN 1104-4047

Printed in Sweden
SCB-Tryck, Örebro 2005.04

Förord

Vart tredje år gör Statistiska centralbyrån, SCB, en befolkningsprognos för riket som publiceras i serien *Demografiska rapporter* (senast år 2003). I prognoser för mellanliggande år, som publiceras i serien *Statistiska meddelanden*, görs endast mindre justeringar vid behov.

Föreliggande rapport ger en detaljerad beskrivning av den grundläggande beräkningstekniken. Redogörelsen gäller den bakomliggande modellen och en redovisning av beräkningsprogrammet.

Vidare ges en summarisk beskrivning av beräkningsmodulerna som ligger till grund för antagandena om prognosens förändringsfaktorer – fruktsamhet, dödlighet och migration. Dessa skattningar ingår alltid som en del i de redovisade prognospublikationerna.

Rapporten har utarbetats av Jan Qvist och Hans Lundström.

Innehåll

A separate text in English is provided at the end of the publication, on page 29

Förord	3
Sammanfattning	7
1. Inledning	8
2. Allmänt om prognosmodellen	9
Förändringsfaktorernas beräkningsmoduler.....	10
3. Formell beskrivning av prognosen	11
Framskrivningsmodellen	11
4. Översiktlig beskrivning av prognosprogrammet	15
Centrala prognosrutinen	15
Prognoskörning.....	17
Programmet startas i fliken Pyramid	18
Indata till prognosen ska arrangeras enligt en given mall.....	19
Utdata från prognosen	20
Layout på del olika pivot-tabell med prognosresultat.....	21
5. Beräkningsmoduler för förändringsfaktorerna	23
Dödlighet	23
Migration	25
Fruksamhet.....	26
Referenser	28
In English	29
Summary	29
List of terms	30

Sammanfattning

Beräkningen av den framtida befolkningsstorleken benämns ofta framskrivning eller projektion. Ordet prognos har en vidare betydelse. Ordet prognos används för ett uttalande om framtiden som bedöms vara realistiskt och ha en acceptabel trolighet. Vid en prognoshorisont på 50 år eller mer, som numera förekommer, är prognosbegreppet naturligtvis mest adekvat för den närmaste framtiden. Framskrivningen som helhet brukar ändå kallas för en prognos.

Inför befolkningsprognosen för riket 2003-2050 gjordes en rad viktiga metodologiska förändringar av modellen. Utöver uppdelningen av befolkningen på kön, ålder och kalenderår tillfogades en uppdelning efter personer födda i utlandet respektive i Sverige. Dessutom ändrades beräknings sättet för utvandringen. Tidigare var utvandringen en exogen variabel. I den nya modellen hänger utvandringen samman med hur stor befolkningen är, både då det gäller den Sverigefödda befolkningen och den utlandsfödda. Genom uppdelningen efter födelseland kan mera realistiska förutsägelser göras av antalet utvandrare, eftersom de varierande utvandringsriskerna efter födelseland kan appliceras.

Antaganden om parametervärden under prognosperioden för förändringsfaktorerna fruktsamhet, dödlighet och migration utgör förutsättningarna för prediktionen tillsammans med utgångsfolkmängden. Vid beräkningen av förändringsfaktorerna används olika metoder. För fruktsamheten används en kohort-periodmodell, för dödligheten en modell utarbetad av Lee-Carter och för migrationen trendanalys med uppdelning av befolkningen efter födelseland.

I prognosmodellen beräknas endast ett prediktionsvärde för folkmängdsstorleken i respektive befolkningsgrupp. Detta beräkningsätt kallas deterministiskt. Numera används i vissa länder även stokastiska prognosmodeller där man låter folkmängdsstorlekarna anta en stokastisk form för i första hand beräkning av variationsmått i framskrivningen (prediktionsintervall). Vid SCB pågår för närvarande ett utvecklingsarbete inom detta område. Den statistiska komplexiteten ökar emellertid i detta fall eftersom det krävs en strikt sammanhållen modellstruktur för variansbestämningen. För uppskattning av variationsgränser i den deterministiska modellen används alternativa prognosantaganden (scenarier). Dessa är dock inte helt kongruenta med variationen i stokastiska prognoser, eftersom scenarierna representerar alternativa modellstrukturer i en vid mening.

1. Inledning

Beräkningen av den framtida befolkningsstorleken benämns ofta framskrivning eller projektion. Detta är tekniska termer som åsyftar en kalkyl av en framtida utveckling under vissa antaganden om fruktsamhet, dödlighet och migration. Ordet prognos har en vidare betydelse. Prognosuttrycket används för ett uttalande om framtiden som bedöms vara realistisk och ha en acceptabel trolighet¹. Vid en prognoshorisont på 50 år eller mer, som numera förekommer, är prognosbegreppet naturligtvis mest adekvat för den närmaste framtiden. Framskrivningen som helhet brukar ändå kallas för en prognos.

Vart tredje år gör Statistiska centralbyrån, SCB, en befolkningsprognos för riket som publiceras i serien "Demografiska rapporter" (senast år 2003). I den rapportserien visas hur förändringsfaktorerna fruktsamhet, dödlighet och migration estimeras och extrapoleras. Extrapoleringen mynnar ut i fixa antaganden om utvecklingen. I prognoser för mellanliggande år anges endast eventuella korrigeringar av förändringsfaktorerna (publiceras i serien "Statistiska meddelanden", senaste 2004).

Föreliggande rapport innehåller en mera detaljerad beskrivning av den grundläggande beräkningstekniken. Redogörelsen består av den bakomliggande modellen och en redovisning av beräkningsprogrammet. Rapporten tjänar också ett annat syfte. Den utgör en särskild dokumentation av framskrivningarna gjorda i prognosomgångarna åren 2003 och 2004. Inför prognosen 2003 gjordes en rad viktiga metodologiska förändringar av modellen. Utöver uppdelningen av befolkningen på kön, ålder och kalenderår tillfogades en uppdelning av befolkningen efter personer födda i utlandet respektive i Sverige. Dessutom ändrades beräknings sättet för utvandringen. Tidigare var utvandringen en exogen variabel. I den nya modellen hänger utvandringen samman med hur stor befolkningen är, både då det gäller den Sverigefödda befolkningen och den utlandsfödda. Genom uppdelningen av befolkningen efter födelseland kan mera realistiska förutsägelser göras av antalet utvandrare, eftersom de varierande utvandringsriskerna efter födelseland kan appliceras.

Denna rapport är disponerad på följande sätt. I kapitel 2 beskrivs modellen allmänt. Kapitel 3 innehåller den formella prognosmodellen och kapitel 4 datorprogrammet för prognosmodellen. Slutligen redovisas i kapitel 5 en summarisk beskrivning av beräkningsmodulerna för prognosens förändringsfaktorer, vilka ligger vid sidan av själva prognosformuleringen.

¹ Såväl projektion som prognos kommer från latin. Medan projektion betyder "kasta framåt" betyder prognos "se framåt". Medan en projektion är en mekanisk beräkning av framtida utveckling, är en prognos ett uttalande om framtiden.

2. Allmänt om prognosmodellen

Under ett framskrivningsår förekommer det ett flöde av personer in och ut ur befolkningen. Födda och invandrare tillkommer, döda och utvandrare utgår. I modellen kalkyleras storleken av dessa variabelflöden. Utifrån folkmängden vid ett kalenderårs början erhålls folkmängden vid årets slut genom addition och subtraktion av flödena. Processen som helhet är rekursiv. Det första årets predicerade befolkning utgör i sin tur utgångspunkt för beräkningen av folkmängden i slutet av nästföljande år, osv.

Antaganden om parametervärden under prognosperioden för förändringsfaktorerna fruktsamhet, dödlighet och migration utgör förutsättningarna för prediktionen tillsammans med utgångsfolkmängden. Då det gäller förändringsfaktorerna i modellen består dödligheten och utvandringen av övergångssannolikheter/risker i olika åldrar. Antalet invandrare är dock bestämt i form av en exogen variabel. Efter beräkning av avgångar ur befolkningen (döda och utvandrare) och tillskottet av invandrare kan antal födda beräknas med hjälp av uppgifterna om medelfolkmängden kvinnor i olika åldrar och fruktsamhetstal (centralkvoter). Förändringsfaktorernas parametrar estimeras och extrapoleras med hjälp av olika statistiska modeller (se beräkningsmoduler i kap. 5).

I prognosmodellen beräknas endast ett prediktionsvärde för folkmängdsstorleken i respektive befolkningsgruppering. Detta beräkningsätt kallas deterministiskt. För en fixerad uppsättning parametrar (antaganden) representerar prediktionsvärdena formellt väntevärden för folkmängden i modellen. Väntevärden för folkmängdsutvecklingen ger den minsta felmarginalen jämfört med observerade utfallsvärden. Notera att det ur statistisk synpunkt då är väsentligt att skattningsarna av parametrarna som ligger till grund för antagandena är stabila och tillförlitliga för att uppnå önskade centrala prediktionsvärden².

Numera används i vissa länder även stokastiska prognosmodeller där man låter folkmängdsstorlekarna anta en stokastisk form för i första hand beräkning av variationsmått i framskrivningen (prediktionsintervall)³. Den statistiska komplexiteten ökar emellertid i detta fall eftersom det krävs en strikt sammanhållen modellstruktur för variansbestämningen⁴. Vanligtvis används simuleringsteknik för själva beräkningen av variationsmåten för den framskrivna folkmängden. För uppskattning av variationsgränser i den deterministiska modellen används alternativa prognosantaganden

² Om parameterestimaten är behäftade med stor slumpmässig variation innehåller beräkningarna av folkmängdsuppgifterna för varje enskild prognosomgång en märkbar avvikelse från ett centralt mått. Mer väsentligt är dock att eventuella strukturella brister i beskrivningen av den framtida utvecklingen av parametrarna skapar "bias", särskilt på lång sikt.

³ Vid SCB pågår för närvarande ett utvecklingsarbete inom detta område (Hartmann 2003).

⁴ En annan svårighet hänger samman med den långa prognoshorisonten i befolkningsprognoser som används nuförtiden. En strikt sammanhållen modellstruktur för parameterestimationen kan innebära att relativt blygsamma trender kan få konsekvenser på lång sikt.

(scenarier). Dessa är dock inte helt kongruenta med variationen i stokastiska prognoser, eftersom scenarierna representerar alternativa modellstrukturer i vid mening.

Förändringsfaktorernas beräkningsmoduler

Denna rapport är i första hand en helhetsbeskrivning av framskrivningsmodellen. Estimationen, extrapoleringen och fixeringen av förändringsfaktorerna görs i särskilda beräkningsmoduler. Redovisningen av dessa beräkningar sker i respektive prognospublikation. Vi kommer dock här att översiktligt redovisa metodiken för att ge en sammanhängande metodbeskrivning.

3. Formell beskrivning av prognosen

Utgångspunkten för framskrivningen är den senast observerade folkmängden den 31/12 efter kön, i ettåriga ålderklasser och efter födelseland⁵. Prognosberäkningarna rullar fram från årsskifte till årsskifte (1-årsframskrivning). Alla beräkningar görs med åldersspecifika parametrar.

I grunden består modellen av fyra populationer. Män respektive kvinnor födda i Sverige, män respektive kvinnor födda utomlands. Det finns dock länkar mellan populationerna. Antalet nyfödda bestäms med utgångspunkt från antalet kvinnor i barnafödande åldrar ur båda kvinnopopulationerna. Antalet nyfödda pojkar och flickor beräknas sedan med en könkvot och fördelas på populationerna Sverigefödda. Hela Sveriges befolkning består av de sammanslagna populationerna.

Formlerna i det följande för befolkningen från 1 års ålder är generella för de fyra nämnda populationerna. Vi anger därför inga index för kön och födelseland. I prognosprogrammet sker dock alla beräkningarna för respektive population.

Framskrivningsmodellen

Framskrivning under prognosåret av antalet personer som är 1 år eller äldre vid årets slut

Genom en rekursiv prognosmodell erhålls folkmängden i åldrarna $x = 1, 2, 3, \dots, w$ år vid slutet av kalenderår t ,

$$P_x^t = P_{x-1}^{t-1} - D_x^t + I_x^t - U_x^t \quad (3.1)$$

där

P_x^t = antal män/kvinnor efter födelseland vid slutet av år t , vilka var födda år $t-x$

D_x^t = antal döda män/kvinnor efter födelseland under år t , vilka var födda år $t-x$

I_x^t, U_x^t = antal in- resp. utvandrare (män/kvinnor) efter födelseland under år t , vilka var födda år $t-x$.

Som utgångspopulation används i högerledet den senast observerade befolkningen den 31/12. Notera att w är en öppen åldersklass (105+ år). Denna åldersklass utgörs i slutet av år t av summan av två framskrivna grupper av personer vid årets början (104 år respektive 105+ år).

⁵ Uppdelning på födda i Sverige och födda utomlands.

Beräkning av befolkningsförändringar under ett givet prognosår

I det följande visas beräkningsrutinerna för befolkningsförändringarna som utgörs av döda, immigranter och emigranter i den tidigare redovisade rekursiva formeln (3.1).

Antalet döda (D)

Antal döda män/kvinnor efter födelseland år t beräknas genom

$$D_x^t = P_{x-1}^{t-1} \cdot q_x^t \quad (3.2)$$

där

$x = 1, 2, 3, \dots, w$ vid årets slut

q_x^t = antagen ettårig dödsrisk, d.v.s. prognostiserad sannolikhet för person född år $t-x$ att dö under år t .

Det kan noteras att antalet döda ovan beräknas med hjälp av könsspecifika dödsrisker men någon differentiering av dödlighet mellan utlandsfödda och Sverigefödda har inte varit nödvändig. I stora drag överensstämmer dödligheten mellan grupperna.

Beräkningen av antalet döda baseras enbart på dem som fanns i landet vid årets början. Migrationsströmmarna under ett givet prognosår påverkar alltså inte beräkningen av antalet döda under detta år. Bakgrunden till förfaringsättet är att de flesta in- och utvandrare är unga, vilket innebär att skillnaden i antalet dödsfall blir försumbar.

Definitionen av en dödsrisk och dess estimator överensstämmer med den teknik som används för "livslängdstabellerna" i den officiella statistiken. Detta innebär att vi skattar en nettosannolikhet för dödligheten (baserad på en riskfolkmängd)⁶.

De antagna dödsriskerna för prognosperioden utgörs av extrapolerade dödsrisker för en vald observationsperiod. Hur dessa beräkningar görs visas i kapitel 5.

Vi kan också notera att antalet döda enligt formeln (3.2) motsvarar väntevärdet av en binomialfördelad stokastisk variabel.

Antalet in- och utvandrare (I,U)**Invandring**

Invandringen av utrikes födda är exogen i prognosmodellen. Återinvandringen av Sverigefödda bestäms dock på ett något annat sätt. Denna invandringsström ställs i relation till tidigare utvandring av Sverigefödda (beräknas separat, se kap. 5). Sambandet innebär att återinvandringen av Sverigefödda inte är exogen i vanlig bemärkelse men införs exogent formellt i prognosmodellen.

⁶ Net probability, Chiang (1968). För skattning se t.ex. Befolkningsstatistik 2003. Del 4. SCB, sid. 59. Notera att prognosen har ett något annorlunda åldersbegrepp än i livslängdstabellen.

Invandringsströmmen i modellen utgörs av kalkylerade antalsuppgifter. Detta gäller såväl utrikesfödda invandrare som Sverigefödda invandrare. Invandringen som sker under året läggs till folkmängden vid årets slut.

I_x^t = antal invandrare (män/kvinnor) efter födelseland under år t , vilka var födda år $t-x$ (exogen).

Utvandring

Beräkningen av antal utvandrare sker med hjälp av utvandringsrisker, vilket är ny rutin i prognosmodellen. Tidigare bestämdes antalet exogen i likhet med immigrationen. Prognosmodellen innehåller specifika utflyttningsrisker uppdelade efter kön, ålder och födelseland.

Antalet utvandrare (U) år t ges av

$$U_x^t = P_{x-1}^{t-1} \cdot u_x^t \quad (3.3)$$

där

$x=1, 2, 3 \dots w$ vid årets slut

u_x^t = antagen 1-årig risk att utvandra, d.v.s. prognostiserad sannolikhet för en man/kvinna efter födelseland född år $t-x$ att utvandra under år t .

Utvandringssannolikheten skattas som en bruttosannolikhet, dvs. den observerade andelen utvandrade bland antalet personer som fanns i Sverige vid årets början. Antalet dödsfall kan i detta fall påverka skattningen av utflyttningssannolikheten⁷.

Vi kan också notera att antalet utvandrare enligt formeln ovan motsvarar väntevärdet av en binomialfördelad stokastisk variabel.

Framskrivning under prognosåret av antalet födda barn som är 0 år vid årets slut

Det som saknas i den framskrivna befolkningen är 0-åringarna. Dessa består av nytillskott till folkmängden i form av förväntat antal födda i Sverige samt invandrade 0-åringar under år t . Då det gäller antalet födda i Sverige antas att fruktsamhetstalen är gemensamma för Sverigefödda och utlandsfödda kvinnor⁸. Vi kan därför enkelt formulera antalet födda.

⁷ Sannolikheten är skattad i populationen som en med dödligheten konkurrerande risk (crude probability). Se också kap. 5.

⁸ Enligt Andersson (2004) anpassar sig invandrade kvinnor på sikt till ett svenskt födelsemönster.

Totala antalet födda i Sverige år t ges av

$${}_{Sv}F^t = \sum_{x=15}^{49} f_x^t ({}^{Kv}P_{x-1}^{t-1} + {}^{Kv}P_x^t) / 2 \quad (3.4)$$

där

${}_{Sv}F^t$ = födda i Sverige under år t

f_x^t = är antagna åldersspecifika fruktsamhetstal under år t

${}^{Kv}P_x^t$ = totalt antal kvinnor vid slutet av år t , vilka var födda år $t-x$

Folkmängdsuppgifterna i formeln bildar medelantalet kvinnor år t . Produkttermerna i summan av fruktsamhetstalet och medelfolkmängden kvinnor anger i sin tur förväntat antal födda per kvinna i åldern x år⁹. Notera att både antalet döda och antalet flyttare under prognosåret påverkar kalkylen av antalet födda genom medelfolkmängdsberäkningen¹⁰.

Antalet födda ${}_{Sv}F^t$ delas upp på flickor och pojkar med hjälp av könskvoten, som satts till 5 145 födda pojkar per 10 000 levande födda barn¹¹. Förväntat antal nollåriga pojkar respektive flickor vid slutet av år t erhålls slutligen genom

$${}_{Sv}P_0^t = {}_{Sv}F^t \cdot k - {}_{Sv}D_0^t + {}_{Sv}I_0^t - {}_{Sv}U_0^t \quad (3.5)$$

där $k=0,5145$ om beräkningen avser pojkar respektive $0,4855$ för flickor. Genomgående anger index Sv i formeln födda i Sverige.

Antalet döda bland de under året födda är,

$${}_{Sv}D_0^t = {}_{Sv}F^t \cdot k \cdot {}_{Sv}q_0^t \quad (3.6)$$

och

${}_{Sv}I_0^t$ är exogent bestämd

(betingas av under året utvandrade Sverigefödda 0-åringar)¹².

Utvandringen fås genom

$${}_{Sv}U_0^t = {}_{Sv}F^t \cdot k \cdot {}_{Sv}u_0^t \quad (3.7)$$

Utrikes födda 0-åringar behandlas i modellen som invandrare i åldrarna 1 år och däröver. I detta fall beräknas ingen återutflyttning eller dödlighet.

$${}_{Ut}P_0^t = {}_{Ut}I_0^t \quad (3.8)$$

⁹ Prognosprogrammet (kap.4) medger separat kalkyl för Sverigefödda respektive utrikes födda kvinnor.

¹⁰ Vi växlar då från sannolikhetsberäkning till centralkvotsberäkning. Detta är praktiskt då ett flertal flöden behandlas.

¹¹ I prognosomgångarna 2003-2050 och 2004-3050.

¹² Mycket litet antal.

4. Översiktlig beskrivning av prognosprogrammet

Prognosmodellen består av separata beräkningsmoduler för de olika förändringsfaktorerna samt en central beräkningsrutin som på grundval av startfolkmängd och antaganden genererar själva befolkningsprognosen. I detta avsnitt behandlas den centrala beräkningsrutinen.

Prognosprogrammet är uppbyggt med Excel som bas och med utnyttjande av Visual Basic for Applications (VBA). Fördelen med denna lösning är att användningen blir enkel. All in- och utmatning av data sker i olika flikar i Excel medan själva programmet finns kopplat till Excelfilen. När Excelfilen med alla antaganden och prognosresultat sparas undan så sparas samtidigt prognosprogrammet inklusive alla eventuella justeringar. Det vill säga prognosprogrammet är självdokumenterande.

Prognosen är egentligen uppdelad i fyra delar med separat kalkyl varje kalenderår för kvinnor och män och med en uppdelning efter födda i Sverige respektive ej födda i Sverige. Totalt består input till prognosen av:

- 1) Startfolkmängd
- 2) Fruktsamhet
- 3) Dödlighet
- 4) Invandring
- 5) Utvandring

Förutom inläsning av antaganden och utmatning av resultat så arbetar VBA-programmet helt skilt från Excel i övrigt. Alla förutsättningar för körningen läses in från respektive kalkylblad och sparas undan i VBA-programmet. Själva prognosen sker sedan helt inom VBA-programmet.

Centrala prognosrutinen

VBA-koden, som är kärnan i prognosprogrammet, är strukturerad i ett antal subrutiner. Själva prognoskalkylen sker i subrutinen Prognos. Nedan finns en något redigerad version av denna subrutin. Före och efter detta kodfragment finns rutiner för dimensioneringar av variabler, matriser mm., inläsning av antaganden med kontroller samt utmatningar till pivottabeller, befolkningspyramid mm. Just delarna före och efter subrutinen "Prognos" upptar större delen av koden.

Översiktligt har prognosprogrammet följande struktur, ordnat i tidsföljd efter när respektive subrutin exekveras i programmet:

Init Initierar körningen samt styr i vilken ordning subrutiner skall anropas.
Specificering av startår respektive slutår för prognosen samt inläsning av proportionen pojkar bland födda.

Inläsning av antaganden för Sverige födda

lasfx Läser in fruktsamhetsantagandet till matrisen "fx"
laspop Läser in startfolkmängd till matrisen "pop"
lasqx Läser in dödsrisker till matrisen "qx"
lasutv Läser in utvandringsrisker till "utrisk"
lasinv Läser in antal invandrare till matrisen "inv"

Inläsning av antaganden för utrikes födda:

Blasfx Till matrisen "Bfx"
Blaspop Till matrisen "Bpop"
Blasqx Till matrisen "Bqx"
Blasutv Till matrisen "Butrisk"
Blasinv Till matrisen "Binvs"

Matriserna innehåller antaganden med uppdelning efter ålder och år. För startfolkmängd, dödsrisker, utvandringsrisker och antal invandrare finns även en uppdelning efter kön.

Subrutinen prognos i något redigerad form

For ar = start To slut
For alder = 1 To 106
For sex = 1 To 2

$dod(ar, alder, sex) = upop(ar-1, alder-1, sex) * qx(ar, alder, sex)$ antal döda
 $utv(ar, alder, sex) = upop(ar-1, alder-1, sex) * utrisk(ar, alder, sex)$ antal utv
 $Bdod(ar, alder, sex) = Bupop(ar-1, alder-1, sex) * Bqx(ar, alder, sex)$ antal döda
 $Butv(ar, alder, sex) = Bupop(ar-1, alder-1, sex) * k * Butrisk(ar, alder, sex)$ antal utv

Kalkyl folkmängd vid årets slut

If alder < 106 then
 $upop(ar, alder, sex) = upop(ar-1, alder-1, sex) - dod(ar, alder, sex) + inv(ar, alder, sex) - utv(ar, alder, sex)$
 $Bupop(ar, alder, sex) = Bupop(ar-1, alder-1, sex) - Bdod(ar, alder, sex) + Binvs(ar, alder, sex) - Butv(ar, alder, sex)$
Else
 $upop(ar, alder-1, sex) = upop(ar, alder-1, sex) + upop(ar-1, alder-1, sex) - dod(ar, alder, sex) + inv(ar, alder, sex) - utv(ar, alder, sex)$
 $Bupop(ar, alder-1, sex) = Bupop(ar, alder-1, sex) + Bupop(ar-1, alder-1, sex) - Bdod(ar, alder, sex) + Binvs(ar, alder, sex) - Butv(ar, alder, sex)$
End if
Next sex
Next alder

Kalkyl antal födda

For alder = 16 To 50
 $fod(ar, alder) = 0.5 * fx(ar, alder) * (upop(ar, alder, 2) + upop(ar-1, alder-1, 2))$
 $Bfod(ar, alder) = 0.5 * Bfx(ar, alder) * (Bupop(ar, alder, 2) +$

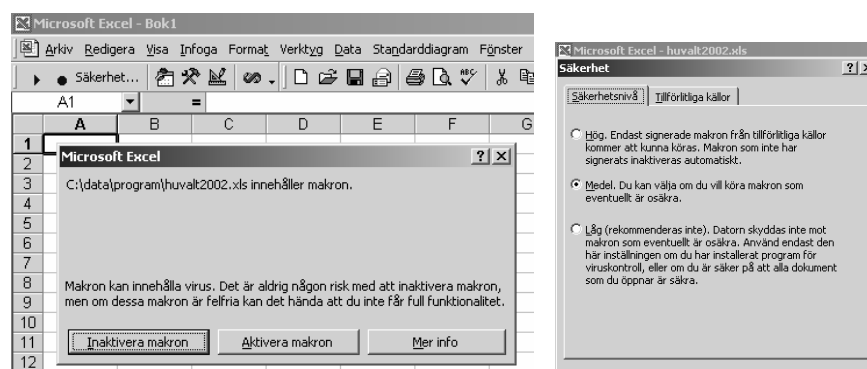

```
Bupop(ar-1,alder-1,2))
  fodda=fodda+fod(ar,alder)+Bfod(ar,alder)
Next alder
```

```
For sex=1 To 2
  f=konprop*fodda
  If sex=2 then f=(1-konprop)*fodda
  dod(ar,0,sex) = f * qx(ar,0,sex)
  utv(ar,0,sex) = f * utrisk(ar,0,sex)
  upop(ar,0,sex)=f-dod(ar,0,sex)+inv(ar,0,sex)-utv(ar,0,sex)
  Bupop(ar,0,sex)=Binv(ar,0,sex)
Next sex
```

Next Ar

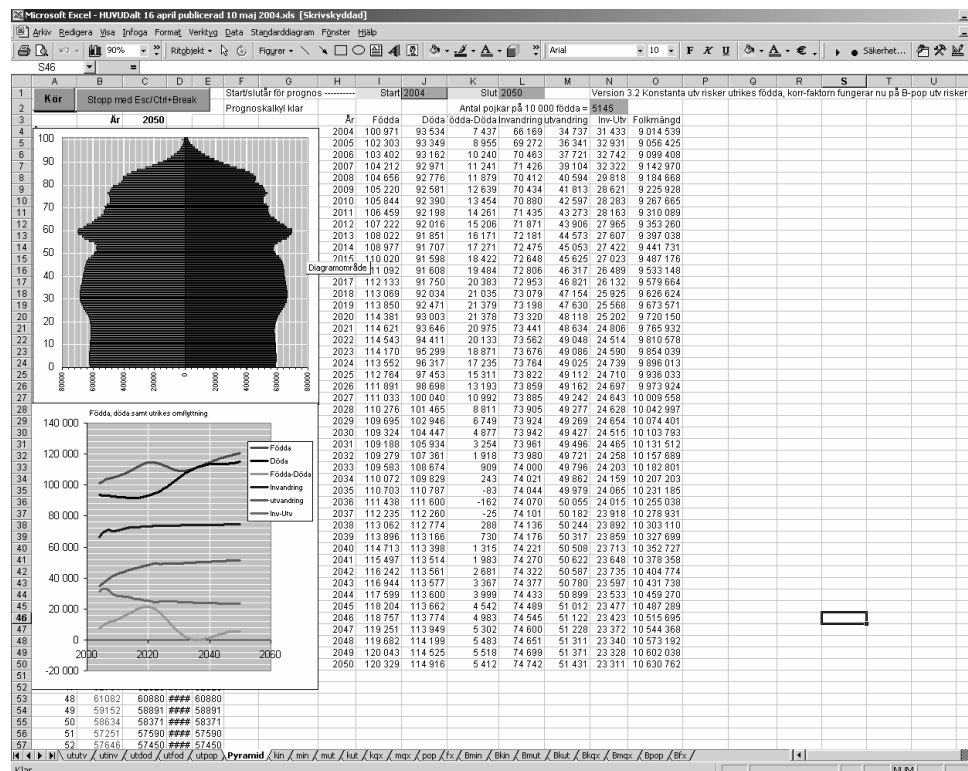
Prognoskörning

Vid en ny prognosomgång läses först en tidigare prognos in varpå antaganden mm justeras. Det är viktigt att genast spara undan Excel-filen med ett nytt unikt namn för att undvika risken att den äldre prognosen förstörs. Välj alternativet Aktivera makron.



Eventuellt kan säkerhetsnivån i Excel behöva justeras från Hög till Medium för att VBA-koden ska kunna exekveras (I Excel välj Verktyg, Makro, Säkerhet och bocka för Medel OK)

Programmet startas i fliken Pyramid



Samtliga flikar enligt skärmbilden ovan måste finnas med för att programmet ska fungera. Flikarna till höger om fliken Pyramid är specifikation av alla antaganden och startfolkmängd. Till vänster om fliken Pyramid hamnar resultatet från prognoskörningen.

Indata till prognosen ska arrangeras enligt en given mall

För att inläsningen av data ska bli korrekt måste varje flik har ett bestämt utseende. På rad två specificeras startår, slutår, startålder och slutålder för de inmatade uppgifterna i kolumnerna I till L. Vidare finns en möjlighet att enkelt skala om alla inmatade tal med en faktor. Vanligen är dock värdet i kolumn G2 lika med ett. I första kolumnen med start på rad fem specificeras åldern från och med startålder till slutålder och på rad 4 med start i kolumn två årtalen från startår till slutår. Första värdet för startålder och startår hamnar sedan i cellen B5.

The screenshot shows an Excel spreadsheet with the following structure:

- Row 1:** Fruitsamhet (Fertility rate)
- Row 2:** Alder årets slut (*1000) (Age at end of year *1000)
- Row 3:** Fruitsamhet antagande 7 april (Fertility rate assumption 7 April)
- Row 4:** Alder (Age)
- Columns:** A (Fruitsamhet), B (B-population), C (Fruitsamhetsjust. (0,5-halvering)) (Fertility rate adjusted (0.5-halving)), and columns I through S representing years from 2002 to 2019.
- Row 41:** År (Year)
- Row 42:** TFR (Total Fertility Rate)

Vid inläsningen sker sedan ett antal logiska kontroller så att ålders- och kalenderårsuppgifterna stämmer, kommer i rätt ordning samt att det finns ett numeriskt värde i varje cell.

Vidare kontrolleras att data finns från och med första prognosåret som preciserats i fliken Pyramid. Antaganden kan finnas för år före första prognosåret utan problem för programmet. Om däremot prognosantaganden inte finns med för de sista prognosåren hamnar programmet i ett okontrollerat läge. Detta kan man se genom att sammanställningen av folkmängd, födda mm i fliken Pyramid avslutas tidigare än väntat.

4. Översiktlig beskrivning av prognosprogrammet SCB:s modell för befolkningsprognoser

När det gäller startfolkmängden ska den avse 31 december året före första prognosår som specificerades i fliken Pyramid.

Själva prognosen startas genom att klicka på knappen **Start**. Under körningen uppdateras befolkningspyramiden och tabellen till höger om pyramiden. Körningen avslutas med att data sparas undan i ett antal pivot-tabeller. Hur långt prognoskörningen avancerat meddelas högst upp på fliken med pyramiden.

Utdata från prognosen

Resultatet från prognosen läggs i fem pivot-tabeller för ökad flexibilitet. Tabellerna omfattar folkmängd, födda, döda, invandring samt utvandring med en uppdelning efter kön, ålder, år och födda i Sverige respektive utrikes födda.

Därtill finns en översikt över total folkmängd, antal födda, döda och utrikes omflyttning i en tabell till höger om befolkningspyramiden.

Layout på del olika pivot-tabell med prognosresultat

Folkmängd

Folkmängd
31 dec.

År	Ålder	Kön	Åld5	Åld10	Antal	S/U-pop
2003	0	Män	0-4	0-19	50867	S
2003	0	Män	0-4	0-19	271	U
2003	0	Kvinnor	0-4	0-19	47867	S
2003	0	Kvinnor	0-4	0-19	225	U
2003	1	Män	0-4	0-19	48903	S
2003	1	Män	0-4	0-19	784	U
2003	1	Kvinnor	0-4	0-19	46382	S
2003	1	Kvinnor	0-4	0-19	1010	U
2003	2	Män	0-4	0-19	46666	S
2003	2	Män	0-4	0-19	1186	U
2003	2	Kvinnor	0-4	0-19	43951	S
2003	2	Kvinnor	0-4	0-19	1401	U
2003	3	Män	0-4	0-19	46168	S
2003	3	Män	0-4	0-19	1512	U
2003	3	Kvinnor	0-4	0-19	43419	S
..

Födda

År	Ålder	Antal	S/U-pop
2004	16	61	S
2004	16	6	U
2004	17	179	S
2004	17	20	U
2004	18	320	S
2004	18	37	U
2004	19	606	S
2004	19	79	U
2004	20	963	S
2004	20	140	U
2004	21	1364	S
2004	21	219	U
2004	22	1814	S
..

Avlidna

År	Ålder	Kön	Antal	S/U-pop
2004	0	Män	159	S
2004	0	Män	0	U
2004	0	Kvinnor	113	S
2004	0	Kvinnor	0	U
2004	1	Män	27	S
2004	1	Män	0	U
..

Invandring

År	Ålder	Kön	Antal	S/U-pop
2004	0	Män	77	S
2004	0	Män	352	U
2004	0	Kvinnor	69	S
2004	0	Kvinnor	329	U
2004	1	Män	123	S
2004	1	Män	608	U
2004	1	Kvinnor	138	S
2004	1	Kvinnor	743	U
2004	2	Män	102	S
2004	2	Män	512	U
..

Utvandring

År	Ålder	Kön	Antal	S/U-pop
2004	0	Män	116	S
2004	0	Män	0	U
2004	0	Kvinnor	112	S
2004	0	Kvinnor	0	U
2004	1	Män	257	S
2004	1	Män	38	U
2004	1	Kvinnor	235	S
2004	1	Kvinnor	28	U
2004	2	Män	224	S
..

Medellivslängd

Medellivslängd män född i Sverige

Ålder	2002	2003	2004	2005	2006	..
0	77,91	77,91	78,10	78,29	78,47	..
1	77,18	77,18	77,36	77,54	77,72	..
2	76,21	76,21	76,39	76,57	76,75	..
3	75,23	75,23	75,41	75,59	75,77	..
..

5. Beräkningsmoduler för förändringsfaktorerna

I detta avsnitt beskrivs huvudlinjerna i tekniken för skattningen av förändringsfaktorernas parametrar. Dessa beräkningsmoduler är fristående från den grundläggande framskrivningsmodellen som vi hittills redovisat. Ny prognosteknik, som oftast rör sig om estimation och extrapolation, kan då relativt enkelt inkluderas.

Det kan nämnas att i prognosrapporterna (serien "Demografiska rapporter") finns en utförlig beskrivning av olika metod- och skattningsmoment. Detta är inte minst viktigt som ett led i dokumentationen av prognoserna, där förutsättningar och resultat alltid skall hänga samman.

Dödlighet

I kapitel 3 visades hur dödligheten är representerad i framskrivningsmodellen genom en uppsättning dödsrisker i olika åldrar och tidpunkter. Dödligheten är könsspecifik men det finns inte någon differentiering efter födelseland. Generellt kan det påpekas att dödligheten präglas av en ihållande nedåtgående trend, vilket gäller de flesta i-länder i världen. Dödlighetsutvecklingen och därmed befolkningsutvecklingen på sikt saknar av detta skäl ett "känt" jämviktsläge.

Vid beräkningen av dödsriskerna under första prognosåret har vi använt oss av den senaste observerade livslängdstabellen (femårsperiod), som justerats för att motsvara första prognosåret (och åldersdefinition). Skattningen över en femårsperiod ger relativt stabila värden och en god bild av den aktuella dödlighetsprofilen efter ålder.

För att beräkna den fortsatta framtida förändringen av dödsriskerna har vi använt Lee-Carter metoden (Lee och Carter 1992). Enligt denna metod bestäms tidsparametrar som sedan extrapoleras. Lee-Cartermodellens samtliga parametrar skattas på grundval av s.k. dödstal¹³.

Dödstalen efter ålder och kalenderår (matris) kan uttryckas med hjälp av följande modell (för män respektive kvinnor),

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t} \quad (5.1)$$

$m_{x,t}$ = dödstal för ålder x och kalenderår t

a_x = åldersspecifik genomsnittlig dödlighetsnivå

k_t = tidstrend i dödstalen

¹³ Ett dödstal beräknas som antalet döda i förhållande till medelfolkmängden (risktid). Dödstalet kan transformeras till en dödsrisk (sannolikhet att dö), som vi använder oss av i prognosmodellen (se kap. 3).

b_x = åldersspecifik vikt för tidstrenden

$\varepsilon_{x,t}$ = slumpterm

Enligt modellen delas dödstalens ålders- och tidsberoende upp i två separata delar eller vektorer (faktorer i 2:a termen i högerledet). Vektorn för tidseffekten k_t antas ha en gemensam struktur i de olika åldrarna. Vektorn för ålderseffekten b_x antas spegla nivåskillnader i dödlighetsutvecklingen i olika åldrar¹⁴.

Vanligtvis används Lee-Carter metoden på data som sträcker sig över mycket långa tidsperioder. I flera fall har hela 1900-talet använts. Vi har dock funnit att ålderseffekten är instabil för så långa tidsserier. I början av seklet gällde nedgången i dödlighet i stor utsträckning yngre människor och i slutet av århundradet skedde nedgången främst bland äldre (Lundström och Qvist 2004). Vår tillämpning för estimationen har därför begränsats till den sena perioden 1990-2002¹⁵.

För att beräkna den framtida dödligheten extrapoleras tidsvektorn \hat{k}_t . Detta kan ske med någon typ av ARIMA-modell. Vi har valt en enkel variant genom att vi antar att utvecklingen är linjär¹⁶. Av modellen framgår att de årliga förändringarna visar de relativa förändringarna av dödstalen¹⁷. Den extrapolerade utvecklingen som är antaganden i framskrivningen är en centrerad utveckling utan stokastisk variabilitet.

Eftersom dödlighetsutvecklingen skall extrapoleras över en mycket lång tid har vi tagit hänsyn till heterogeniteten i förändringen efter dödsorsaker. Omkring 15-20 år in på prognosperioden används en variant på Lee-Carter modellen då vi gör skattningen av tidsutvecklingen för 4 stora dödsorsaksgrupper, som aggregeras till en total utveckling¹⁸.

I "Demografiska rapporter 2003:4" finns en utförlig redogörelse av beräkningsmetoden och resultaten (kap. 3 och 6). Där diskuteras även några metoder som används parallellt med Lee-Carter metoden och som visar dödligheten som en funktion av vissa livsstilsfaktorer. I prognosomgången 2004-2050 användes samma dödsrisker som i den föregående prognosomgången.

¹⁴ Skattningen sker med "singular value decomposition" (SVD) av ursprungsmatrisen av dödstal efter ålder och tid för män respektive kvinnor. Det första singulära värdet och första vektorn i vardera två ortogonala matriser används. Summan av b över åldrarna är lika med 1 och summan av k över tiden är lika med 0. Beräkningen har gjorts i SAS med hjälp av subrutinen SVD. Skattningarna är oviktade, dvs. vi har inte försökt kalibrera parametervärdena i relation till antalet dödsfall. En sådan oviktad skattning kan göras under antagande om att den inre strukturen i tidsutvecklingen är homogen i olika åldrar.

¹⁵ Prognosomgångarna 2003-2050 och 2004-2050.

¹⁶ Då får man den årliga tidsförändringen som skillnaden mellan högsta och lägsta värde i kt -kurvan dividerad med antalet tidpunkter minus ett (lutningskoefficient). Relativa årliga förändringsfaktorer avsedda för extrapolation av dödstalen efter ålder får man genom att lutningen viktas med åldersfaktorn. Framskrivningen av dödligheten innebär att lutningskoefficienten multipliceras med antalet år som skall prognostiseras.

¹⁷ Det kan noteras att förändringarna blir ungefär desamma för dödstalen som för motsvarande dödsrisker i föreliggande fall.

¹⁸ Cirkulationsorganens sjukdomar, tumörer, olyckor /självmod och övriga sjukdomar.

Migration

Migrationsflödena, immigration och emigration, präglas av starka upp- och nedgångar. Initialt tas hänsyn till vilken cyklisk fas flödena befinner sig i. Eftersom immigrationen är sammansatt kan även olika typer av migration påverka den periodiska variationen (flyktinginvandring, anhöriginvandring, mm).

På lång sikt görs antaganden om en jämnare form för migrationen, eftersom man i det perspektivet inte kan förutse tillfälliga fluktuationer. Som helhet kan nettoomflyttningen nära nog jämföras med en stationär process där våra antaganden alltmer närmar sig en centrerad nivå.

Under förra halvseket låg t.ex. nettoinvandringen på en genomsnittlig nivå på nära 20 000 personer per år. Tillfälliga kraftiga fluktuationer förekom under dessa år.

Invandring av utrikes födda

Invandringen som helhet till Sverige under prognosperioden har bestämts utifrån invandringmönster, trender, från nedanstående ländergrupper med uppdelning på ålder och kön. Basperioden för skattning och extrapolation har utgjorts av perioden sedan 1980. Variabeln för *invandring av utrikes födda* består av ett exogent antal personer.

De fem grupperna är:

Norden utom Sverige,
EU25 utom Norden,
länder med högt HDI (Human Development Index¹⁹) utom EU25,
länder med medel HDI (främst länder från Mellanöstern)
länder med lågt HDI.

För länderna med medel och lågt HDI består invandringen av två strömmar dels flyktingar som antas vara en stationär ström med stora variationer mellan åren, dels anhöriginvandring som en funktion av tidigare invandring.

Invandringen av personer födda i Sverige

Återinvandring av personer födda i Sverige har beräknats beroende av deras tidigare utvandring och skattas med en kohortmetod. Underlaget för estimationen är andelen återinvandrade personer födda i Sverige efter tid utomlands.

Beräknade durationsspecifika återflyttningstal appliceras sedan på antalet utvandrare födda i Sverige i prognosen. Med automatik erhålles antalet återinvandrade personer under prognosperioden.

Utvandringsrisker för utrikes födda

Beräkningen av antalet utvandrare ansågs kunna ske med estimerade utflyttningrisker (uu) både efter tid i Sverige (duration) och ålder. Skattningen gjordes dock stegvis (ej simultant), dvs. först efter duration och därefter efter ålder för perioden 1999-2002. En beräkning av antalet utvandrare

¹⁹ Human Development Index publiceras av FN. I Demografiska rapporter 2003:5 finns hela listan. Några stora invandrarländer i Sverige med högt HDI är Jugoslavien och Bosnien – Hercegovina, medelhögt Iran och Irak och lågt Etiopien och Somalia.

enligt durationsmetoden användes för en korrigerad av de åldersspecifika utflyttningriskerna, vilka blir de slutliga parametrarna. Detta kunde göras med hjälp av de relativa summorna av utvandrare mellan de båda metoderna. Vid beräkningen av antalet utvandrare med durationsmetoden gjordes beräkningen för de ovanstående ländergrupperna²⁰.

Utflyttningensrisken (korrigerad) för utrikes födda män resp. kvinnor ett givet år t fås genom,

$$u_x = uu_x \cdot \left[\frac{\sum_g \sum_d U_t P_{g,d} \cdot uu_{g,d}}{\sum_x U_t P_x \cdot uu_x} \right] \quad (5.2)$$

där

u_x = korrigerad utflyttningensrisk för utlandsfödda efter ålder

uu_x = utflyttningensrisk för utlandsfödda efter ålder

$uu_{g,d}$ = utflyttningensrisk för utlandsfödda efter ländergrupp och duration

$U_t P_{g,d}$ = utlandsfödda i ländergrupp g efter duration d

$U_t P_x$ = utlandsfödda efter ålder

Den sista faktorn som påverkar utflyttningensrisken homogent anger skillnaderna mellan durations- och åldersmetoden²¹. Den korrigerade utflyttningensrisken (5.2) efter ålder insättes i formel 3.3 för beräkning av antalet återutvandrare.

Utvandringsrisker för personer födda i Sverige

Utvandringen av personer födda i Sverige beräknas enbart med åldersspecifika utflyttningensrisker.

Den redovisade tekniken för migrationen har beskrivits utförligt i en särskild befolkningsprognos med indelning av befolkningen efter ovanstående grupper samt Sverigefödda - *Sveriges framtida befolkning 2003-2020*.

Svensk och utländsk bakgrund (Demografiska rapporter 2003:5) - men även i metodkapitlet i huvudprognosrapporten (Demografiska rapporter 2003:4).

Fruktsamhet

Även när det gäller fruktsamheten spelar det korta och långa perspektivet en stor roll, då det gäller den metodologiska utformningen. Det summera-

²⁰ Den framtida återutvandringen kalkylerades för varje framskriven invandrargrupp efter duration (täljaren i 5.2). Samma sak gjordes för grupperna sammantagna men med åldersuppdelning (nämnaren i 5.2).

²¹ I prognosomgången 2003-2050 angavs flyttningen i form av uu i olika åldrar tillsammans med korrigeringsfaktorn. I prognosomgången 2004-2050 angavs däremot flyttningen som utflyttningensrisker efter ålder. Dessa risker var justerade (produkten av de två sista faktorerna i högerledet i formeln ovan). Presentationen av förändringen av utflyttningensriskerna under prognosperioden gjordes på samma sätt som för dödligheten (Statistiska meddelanden BE 18 SM 0401).

de periodiska fruktsamhetstalet (TFR) varierar betydligt historiskt. Det är dessa periodiska TFR som i princip skall uppskattas och extrapoleras. På lång sikt kan dock den cykliska variabiliteten inte förutses och tidsbestämmas.

På kort sikt bör det aktuella cykliska läget av TFR vara ledande. Den långsiktiga *nivån* låter vi inträda asymptotiskt i den kalkylerade utvecklingen. Då används det summerade fruktsamhetstalet för kohorter som bas för skattningarna. Detta tal har i stort sett varit stabilt för 1900-talet. Som helhet kan fruktsamhetsutvecklingen därmed ses som en stationär process med TFR för kohorter som en grundnivå (knappt två barn per kvinna).

På relativt kort sikt används en särskild *kohort-periodmodell* för beräkning av den framtida fruktsamheten, vilken tar hänsyn till vilken fas den cykliska fruktsamhetsutvecklingen befinner sig i vid prognosens början²². Sedan slutet av 1980-talet har denna typ av modell över fruktsamhetsutvecklingen tillämpats i SCB:s prognosverksamhet. Det grundläggande arbetet utfördes av Sten Martinelle (Martinelle 1988a, 1988b och 1989). Modellen bygger på antaganden om en slutlig paritetsfördelning (barnantalsfördelning). Den empiriska bakgrunden är att om en kvinna fött ett barn är övergångssannolikheterna för ytterligare barn tämligen stabila, paritet för paritet. Övergångssannolikheterna varierar dock beroende på ålder. Modellen utgår från varje kohort och kan ta hänsyn till förskjutningar av nedkomsterna över ålder. Martinelle (1988a,b) införde ett nytt begrepp "relativ ålder" för att handskas med detta problem (relativ ålder bestämdes som fördelningsfunktionen av nedkomster efter ålder). Det kan noteras att svenska registerdata är en förutsättning för tillämpningen av denna modelltyp. Det bör nämnas att modellerna ovan inte angivit någon prediktionsmetod för första barnet. Denna paritet har en särskild betydelse. Förskjutningar av barnafödandet i inledningen av de fertila åldrarna har betydelse även för den slutliga barnlösheten, Martinelle (1990). I denna rapport finns en redogörelse för modellbaserade kohortvisa extrapolationer av förstabarns-fruktsamheten.

I den senaste fördjupade prognosomgången "Demografiska rapporter 2003:4" användes en liknande modell²³. Extrapolerade åldersspecifika övergångssannolikheter mellan olika pariteter tillämpades (stabila över tiden). Övergångssannolikheterna gällde övergångarna från och med paritet 1. Utgångspunkten vid prediktionen utgjordes av observerade incidenstal över barnafödandet efter paritet inom olika kohorter. Dessa incidenstal framskrevs successivt med hjälp av övergångssannolikheterna. Summation av incidenstalen över paritet ger de vanliga fruktsamhetstalen som vi söker. En viktig del i modellarbetet var att bestämma den bästa projektionen av incidenskurvan för första barnet. Numera föder kvinnor det första barnet i allt högre åldrar. Även i detta fall användes kalkylerade former för kurvan i de yngsta kohorterna för beskrivning av förskjutningarna. Redovisningen av modellen finns på sid. 101–102 i ovan nämnda rapport.

²² Det kan påpekas att den cykliska variation ofta inleds av samhällsekonomisk instabilitet (låg-respektive högkonjunktur).

²³ Gäller även prognosomgången 2004-2050.

Referenser

Andersson G. (2004) Childbearing after migration: Fertility patterns of foreign-born women in Sweden. *International Migration Review* 38(2): 747-775.

Befolkningsstatistik 2003. Del 4 – födda, döda, civilståndändringar m.m. SOS. SCB

Chiang C.L. (1968) *Introduction to stochastic processes in biostatistics*. Wiley, New York.

Hartmann M. (2003) *Primer on making stochastic population projections*. Demographic Report Series. Technical papers. No. 3.

Martinelle S. (1988a) A cohort model for fertility prediction. Bakgrundsmaterial från Demografiska funktionen 1988:1. SCB

Martinelle S. (1988b) A model for analyzing and projecting the cohort fertility by birth order. Bakgrundsmaterial från Demografiska funktionen 1988:2. SCB

Martinelle S. (1989) A cohort model for analyzing and projecting the cohort fertility by birth order. IUSSP, New Dehli

Martinelle S. (1990) The timing of first birth. Analysis and prediction of Swedish birth rates. Bakgrundsmaterial från Demografiska funktionen 1990:1. SCB

Lee R.D. and Carter L.R. (1992) Modeling and forecasting US mortality. *JASA*, 87, 419, 659-671.

Lundström H. and Qvist J. (2004) Mortality forecasting and trend shifts: an application of the Lee-Carter model to Swedish mortality data. *International Statistical Review*. 72, 1, 37-50.

Sveriges framtida befolkning. Befolkningsframskrivning för åren 2000-2050. Demografiska rapporter 2000:1. SOS. SCB

Sveriges framtida befolkning. Befolkningsframskrivning för åren 2003-2050. Demografiska rapporter 2003:4. SOS. SCB.

Sveriges framtida befolkning. Befolkningsframskrivning för åren 2003-2020. Svensk och utländsk bakgrund. Demografiska rapporter 2003:5. SOS. SCB.

Sveriges framtida befolkning 2004-2050. Reviderad befolkningsprognos från SCB. BE 18 SM 0401 (2004). SOS. SCB.

In English

Summary

Calculations of the size of future populations are often called projections. The word forecast has a further meaning. The word forecast is used as an assessment on the future in a realistic way with an acceptable credibility. Today forecast horizons are often 50 years or more, while the concept of forecasts is of course the most suitable for the near future. Projections on the whole are still usually called forecasts.

At the time of the population forecast for Sweden 2003–2050, a number of important methodological changes were made to the model. In addition to breaking down the population by sex, age and calendar year, a breakdown was made of persons born abroad and persons born in Sweden. The way to calculate emigration was also changed. Previously, emigration was an exogenous variable. In the new model, emigration is related to the size of the population, both concerning the number of Swedish-born persons and the number of those born abroad. The breakdown by country of birth allows for more realistic predictions on the number of emigrants, since the varying risks of emigration by country of birth can be applied.

Assumptions on parameter values during the forecast period for the factors of change known as fertility, mortality and migration comprise the conditions for predictions together with the initial size of the population. When calculating the factors of change, various methods are used. A cohort-period model is used for fertility, a model produced by Lee and Carter is used for mortality, and trend analysis with a breakdown of the population by country of birth is used for migration.

In the forecast model, only one prediction value is used for the size of the population in each group. This calculation method is called deterministic. Nowadays, some countries also use stochastic forecast models, which allow the size of the populations to assume a stochastic form to mainly calculate the measure of variation in projections (prediction interval). Development within this area is now underway at Statistics Sweden. However, the statistical complexity is increasing in this area, since it is essential to use a strictly consistent model structure for variance determination. To estimate the limits of variation in the deterministic model, alternative forecast assumptions (scenarios) are used. However, these are not completely congruent with the variation in the stochastic forecasts, since the scenarios represent alternative model structures in a broad sense.

List of terms

antal	number
döda	deaths
dödlighet	mortality
dödsrisk	probability of death
dödstal	mortality rate
framskrivning	projection
fruktsamhet	fertility
invandrare	immigrant
kvinnor	women
kön	sex
livslängdstabell	life table
medelfolkmängd	average population
migration	migration
modell	model
män	men
prognos	forecast
utvandrare	emigrant
utrikes födda	foreign-born
ålder	age

