

# Stokastiska befolkningsprognoser

Moderna befolkningsprognoser har inte en framtid utan oändligt många.

Ett system som påverkas av slumpmässiga effekter är en stokastisk process. Allmänt karakteriseras ett sådant system av att den framtida utvecklingen inte med säkerhet går att förutse, varken med hjälp av tidigare utvecklingsmönster eller på något annat sätt.

**Ordet stokastiskt betyder slumpmässigt och kommer av grekiskans ord för "skicklig på att gissa".**

En befolkning är ett exempel på ett system som är utsatt för stokastiska processer, eller händelseförlopp. Det går inte med säkerhet att fastställa befolkningens framtida storlek och utseende. I sina försök att säga något om framtiden utgår därför prognosmakarna mer eller mindre från den historiska utvecklingen och använder sig av olika typer av tidsserieanalyser.

**DEN STOKASTISKA PROCESSEN** har sitt intellektuella ursprung i det klassiska Grekland. Enligt Heraklit, omkring 500 f.kr, betingas existens av förändring. För sin existens kräver floden nytt vatten och

därför kan man inte stiga ner i samma flod två gånger. Aristoteles kallade denna kunskap för *Panta Rei*; allting flyter. Botanisten Robert Browns upptäckt år 1827 att pollenkorn dansar på en vattenyta var en tidig vetenskaplig insikt om en natur i ständig och svåröversäglig rörelse.

I slutet av 1920-talet visade den ryske matematikern Eugene Slutski att tidsserier kan uppfattas som en sammansättning av oberoende stokastiska impulser. Samtida med Slutski var engelsmannen George Udny Yule. Även han bidrog till den tidiga utvecklingen av tidsseriemetoder.

Många av de metoder som i dag används för analys av tidsserier härrör från Slutskis och Yules idérika arbeten på 1930-talet. I modern matematisk-statistisk bemärkelse har alltså tidsserieanalys endast använts i cirka 75 år. Detta är ett forskningsområde i ständig förnyelse.

Om en process har en historik som beskriver dess framtid med fullständig precision kallas den för deterministisk. En process som saknar denna egenskap kallas för indeterministisk. Flertalet processer innehåller både deterministiska och indeterministiska komponenter.

Processer utan trend kallas för stationära. Dessa karakteriseras av tidsberoende medelvärden och varianser, samt av att

korrelationerna mellan observationerna endast beror på tidsavståndet.

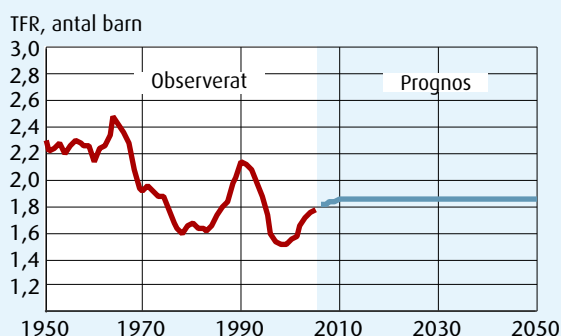
**I DETERMINISTISKA PROGNOSE** ingår antaganden om den demografiska trendutvecklingen av dödlighet, fruktsamhet och integration. Ett typexempel är prognoser över den summerade fruktsamheten (TFR). Stokastiska befolkningsprognoser tar även hänsyn till den historiska slumpmässiga variationen. Det görs med hjälp av tidsseriemodeller som innehåller en variationskomponent.

SCB:s befolkningsprognos har hittills utgått från en deterministisk ansats. Det betyder att man genom att studera utvecklingen för tre demografiska processer; fruktsamhet, dödlighet och migration, gör antagandena om olika framtida utvecklingsförlopp. För varje process görs ett huvudantagande som anger den utveckling som anses mest trolig. Osäkerheten i antagandena anges genom att varje komponent har ett låg- och ett högalternativ.

**PÅ SENARE ÅR** har man börjat beräkna stokastiska projektioner av befolkningen. Stokastiska händelseförlopp kan beskrivas med sannolikhetsteori och beräknas med stokastiska modeller. Beräkningarna grundar sig mestadels på studier av historiska

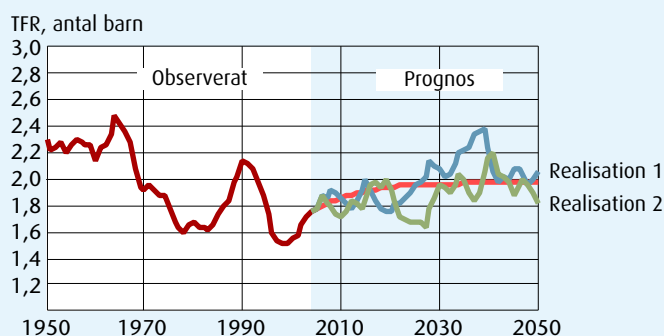
## Fruksamhetens framtida variation

### Deterministisk prognos



Den summerade fruktsamheten (TFR), det vill säga en kvinnas förväntade antal barn, är ett typiskt exempel på en stokastisk process. Det vänstra diagrammet visar observerat TFR för 1950–2005 och SCB:s prognos för 2006–2050. Den observerade TFR-kurvan har en typisk indeterministisk karaktär. Även om efterföljande års värde ofta ligger nära årets värde blir med tiden sambandet mellan årets och framtida års TFR-värden allt svagare. För prognosperioden 2010–2050 är TFR en deterministisk kurva som anger TFR:s prognostiserade medelvärde. Det är således stor principiell skillnad mellan observerat och prognostiserat TFR.

### Stokastisk prognos



Stokastiska prognoser över TFR tar i stället hänsyn till den historiska variationen i fruktsamheten. Detta görs med hjälp av tidsseriemodeller. Som ett exempel på fruktsamhetens variation visar det högra diagrammet observerat TFR för 1950–2005 och en för denna period skattad framtida medelvärdestrend som går mot nivån  $TFR = 2,0$ . Realisation 1 och Realisation 2 anger två olika stokastiska bilder av TFR under prognosperioden. Dessa är modellerade utfall av TFR-processen under perioden 1950–2005. De visar att vi under prognosperioden med stor sannolikhet kan vänta att fruktsamheten kommer att variera mellan cirka 1,6 och 2,4 barn per kvinna.

# skattar osäkerheten

demografiska tidsserier. Metoden gör det möjligt att använda variationerna i de historiska tidsserierna på de projektioner, dvs. framskrivningar, man gör. På så sätt kan man uppskatta den statistiska osäkerheten i projektionen.

Varje stokastisk projektion rör sig på ett unikt sätt in i framtiden. Projektionerna är alla olika men kommer ur samma stokastiska mekanism. När ett lämpligt antal simuleringar har gjorts kan man beräkna ett sannolikhetsintervall för den framtida befolkningen. Intervallets storlek bestäms av variationen hos den tidsserie som modellen har anpassats efter.

Även om vi ofta föreställer oss att vi utifrån kunskap om sociala, ekonomiska, politiska och andra förhållanden kan förutse dödlighet, fruktsamhet och

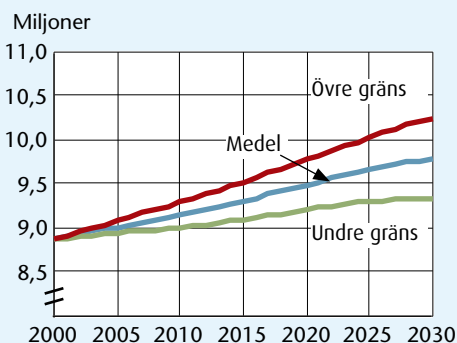
migration de närmaste åren, så finns det alltid inslag av osäkerhet i dessa antaganden. Eftersom även andra faktorer, som vi har ingen eller ganska begränsad kunskap om, också påverkar den demografiska utvecklingen så är befolkningsprognoser behäftade med en osäkerhet som växer med tiden. Ett av huvudändamålen med stokastiska befolkningsprognoser är att skatta denna osäkerhet. ●



Foto: Marie Almers Atterhall, SCB.

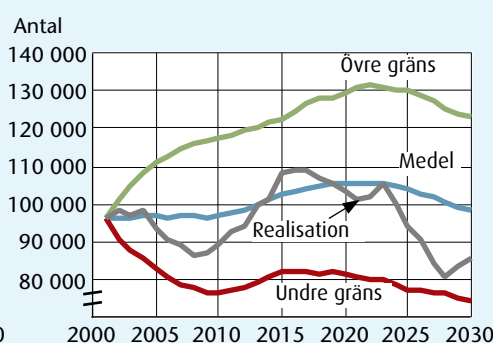
## Sannolikhetsintervall visar osäkerheten

### Prognos över befolkningen i Sverige



Diagrammet visar medelfolkmängden samt övre och undre gränserna för den svenska befolkningen under prognosperioden 2000–2030. Med 95 procents sannolikhet kommer den framtida befolkningen att röra sig inom detta intervall. Skattningarna anger en befolkning år 2030 mellan 9,3 och 10,2 miljoner, med ett medelvärde på 9,8 miljoner.

### Prognos över antal födda barn



Detta diagram visar antalet födda barn per år med ett 95-procentigt prediktionsintervall. Medelkurvan för det årliga antalet födselar ligger omkring 100 000. Diagrammet visar även ett modellerat utfall av antalet födselar under prognosperioden (realisation). Prognosen grunder sig på antagandet att den variation som har observerats under de senaste 50 åren kommer att fortsätta fram till år 2030.

## Författare

**Michael Hartmann**  
arbetar med stokastiska prognoser på Prognosinstitutet vid SCB  
tfn: 08-506 943 74  
e-post: [michael.hartmann@scb.se](mailto:michael.hartmann@scb.se)



**Gustaf Strandell**  
arbetar med statistiska metodfrågor kring befolkningsprognoser vid SCB  
tfn: 019-17 68 43  
e-post: [gustaf.strandell@scb.se](mailto:gustaf.strandell@scb.se)



## Lästips

Gottman, J. M., *Time-series Analysis*, Cambridge, Cambridge University Press 1981.