



NICO KEILMAN
Professor, Universitetet i Oslo

Befolkning: Statistisk sentralbyrå bør endre praksis og publisere sannsynlighetsprognoser¹

Med jevne mellomrom beregner Statistisk sentralbyrå (SSB) prognoser for befolkningen fram til 2100: alderssammensetningen for hvert år, årlige antall fødsler og dødsfall, størrelsen på inn- og utvandringsstrømmene, antall innvandrere med ulik landbakgrunn, osv. Publisering av en ny prognose er planlagt for den 21. juni i år. I denne kommentaren bruker jeg resultater fra SSBs forrige befolkningsprognose fra 2014 for å gjøre rede for et viktig poeng: den framtidige befolkningsutviklingen er (som alle prognoser) usikker, men noen utviklinger er mer sannsynlige enn andre. Derfor bør SSB publisere *sannsynlighetsprognoser*, og ikke bare befolkningstall for framtidige år uten at brukeren vet hvor stor sjanse SSB knytter til den ene eller den andre befolkningsutviklingen.

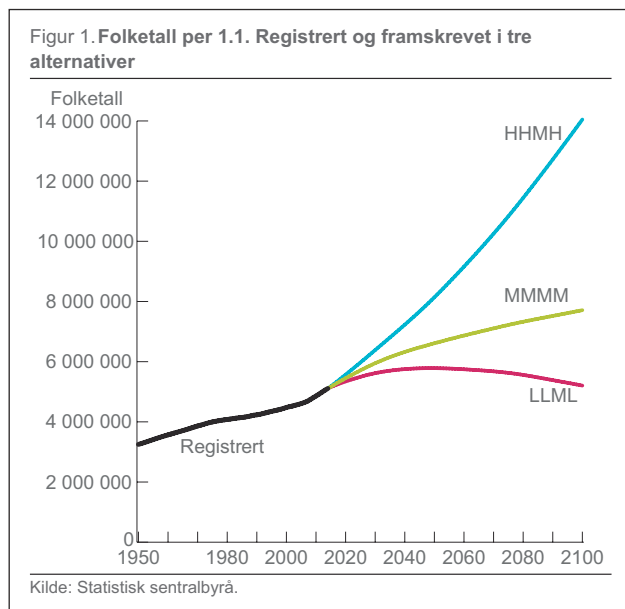
1. DAGENS PRAKSIS: ALTERNATIVER

Figur 1 viser noen resultater fra SSBs befolkningsprognose som ble beregnet i 2014. Folketallet forventes til å øke fra 5,1 millioner pr. 1. januar 2014 til 6,9 millioner i 2060 og videre til 7,7 millioner i 2100. Men en kan ikke utelukke en sterkere, svakere eller til og med negativ befolkningsvekst. Derfor viser figuren to andre baner: en bane med svakere befolkningsvekst (markert som LLML²) resulterer i 5,2 millioner i 2100 og en bane med sterk vekst (HHMH)

viser 14,0 millioner det året. Disse tre alternative baner for befolkningsutviklingen er basert på ulike forutsetninger for fruktbarhet, dødelighet og innvandring i perioden 2014-2100. I hovedalternativet (MMMM) antar SSB at kvinner får i snitt 1,8 barn i 2060, at levealderen kommer til å øke til 87,7 år i 2060, og at nettoinnvandring er 17 100 personer det året. For 2015 er disse tallene henholdsvis lik 1,73 barn pr. kvinne, 82,3 år og 29 800 personer. De andre to alternativene tar utgangspunkt i høyere eller lavere fruktbarhet, en sterkere eller svakere økning i levealderen, og en sterkere eller svakere nedgang i den årlige nettoinnvandringsstrømmen. Se Tønnessen mfl. (2014) for en nærmere forklaring.

¹ Takk til Rolf Aaberge for meget nyttige kommentarer.

² Hvert alternativ beskrives ved fire bokstaver i følgende rekkefølge: fruktbarhet, levealder, innenlandsk flytting og innvandring. M = midtels, L = lav, H = høy. I denne artikkelen begrenser jeg meg til utviklingen for landet under ett, og ser derfor bort fra alternative baner for innenlandsk flytting.



2. TO TOLKNINGER: ALTERNATIVE FRAMTIDER OG USIKKERHETSALTERNATIVER

Å beregne ulike baner for framtidig folketall (og for andre demografiske variabler) basert på forskjellige forutsetninger om dødelighet, fruktbarhet og netto innvandring er vanlig praksis blant statistiske byråer som har ansvar for demografiske prognoser. Denne praksisen går i hvert fall tilbake til 1947, da Whelpton og kolleger publiserte sin prognose for befolkningsutviklingen i USA. På denne måten ønsker man å ta høyde for det faktum at den framtidige befolkningsutviklingen er usikker. Tønnessen mfl. (2014) skriver «MMMMalternativet ... er befolkningsframskrivningenes hovedalternativ og det vi regner som mest sannsynlig. Det er lite trolig at både fruktbarheten, levealderen og innvandringen vil holde seg høy (eller lav) gjennom hele perioden, men HHMH og LLMLalternativene forteller likevel noe om usikkerheten i framskrivningene ...». Her er det flere problematiske utsagn som dessverre ikke kan gis en overbevisende faglig begrunnelse. For det første er det meningsløst å påstå at et bestemt alternativ vil inntreffe med høyere sannsynlighet enn et alternativ basert på andre forutsetninger. For det andre kan ikke spredningen i prognosene basert på forskjellige alternativer brukes som mål på usikkerheten forbundet med SSBs prognoser. Derimot viser spredningen *hvor følsomme* prognoserresultatene er for en endring i en eller flere av de demografiske variablene (samlet fruktbarhetstall for fruktbarhet, levealder ved fødsel for dødelighet, størrelsen på strømmen for nettoinnvandring) som ligger til grunn for prognosene. Dermed kan de tolkes som alternative framtider for Norges befolkningsutvikling. «Scenarier» er en hyppig brukt betegnelse i denne

sammenheng. «Hvis samlet fruktbarhetstall går mot 2 barn pr. kvinne i 2060, hvis levealderen øker mot 90,2 år og hvis nettoinnvandringen går mot 50 000 personer i 2060, så blir resultatet en befolkningsutvikling som vist i alternativet med høy befolkningsvekst (HHMH i Figur 1).» Om en slik utvikling er meget eller lite sannsynlig kan imidlertid ikke SSB si noe om, fordi usikkerheten ikke er kvantifisert.

Hadde derimot SSB ønsket at alternativene for høy og lav befolkningsvekst i Figur 1 «... kunne fortelle noe om usikkerheten ...», da måtte prognoserresultatene for et bestemt år (folketallet, størrelsen til en bestemt aldersgruppe, andel eldre i befolkningen osv.) blitt beregnet på grunn av sannsynlighetsfordelinger, ikke som tre tall.

Fordelen med en slik sannsynlighetsprognose er at brukeren blir fortalt hva SSB forventer når det gjelder mulige avvik mellom prognose og virkelighet, og hvor mye slike avvik varierer på tvers av aldersgrupper, eller mellom menn og kvinner. En sannsynlighetsprognose gir *ikke* et mer treffsikkert anslag for framtiden, men gir *rikere informasjon*. En bruker som er i stand til å forholde seg til prognoseusikkerhet er i stand til å fatte bedre beslutninger om framtidige helsekostnader, pensjonsutgifter, behov for skoleplasser e.l. enn en bruker som bare får servert tre (eller et begrenset antall) alternative beregninger uten sannsynligheter – med andre ord, deterministiske prognoser basert på forskjellige forutsetninger. En bruker som er seg bevisst kostnadene knyttet til prognoseavvik av en bestemt størrelse, kan velge en optimal strategi når en sannsynlighetsprognose foreligger. For eksempel: SSBs hovedalternativ (MMMM) viser et forventet antall personer i alder 67 eller mer i 2020 lik 820 000 (noe avrundet). Alternativene med lav vekst (LLML) og høy vekst (HHMH) gir henholdsvis 800 000 og 840 000. Anta nå at sannsynlighetsfordelingen til antall personer i alder 67+ i 2020 viser at det er mye mer sannsynlig at det virkelige tallet kommer til å ligge mellom 820 000 og 840 000, enn mellom 800 000 og 820 000³. I så fall vil pensjonsmyndighetene ikke satse på SSBs hovedalternativ for denne aldersgruppen (MMMM), men heller på et høyere tall. *Hvor mye høyere* kan en først finne ut når en sannsynlighetsprognose foreligger.

³ Analyser av treffsikkerheten i gamle prognoser viser at antall eldre i framtiden oftere er blitt underestimert enn overestimert.

3. HVORFOR ER EN DETERMINISTISK BEFOLKNINGSPROGNOSE PROBLEMATISK?

I tillegg til at usikkerheten ikke er kvantifisert i SSBs deterministiske befolkningsprognose er det to andre ulemper som jeg vil påpeke.

3.1 Et begrenset antall alternativer gir brukerne mulighet til subjektive eller politisk bestemte valg

En sannsynlighetsprognose tvinger brukeren til å vurdere en hel rekke av mulige prognoseresultater, med tilhørende sannsynligheter. Figur 2 viser en slik fordeling for predikert folketall i 2050. Kurven er basert på en sannsynlighetsprognose som jeg gjør rede for i avsnitt 4.

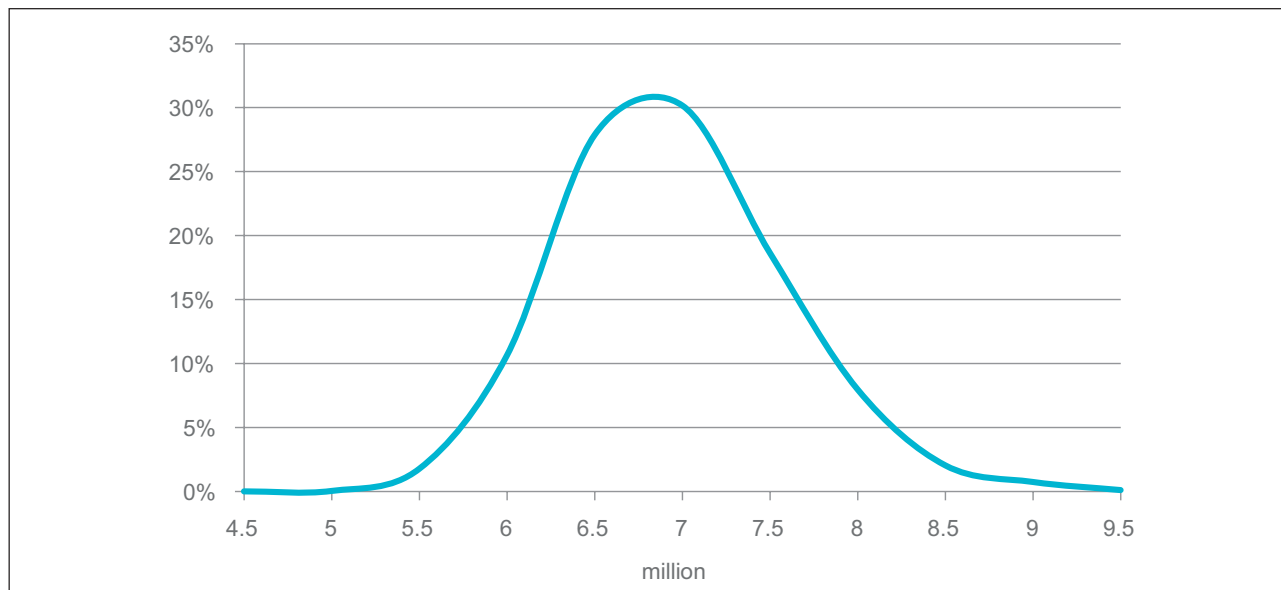
Sjansen for et bestemt tall (for eksempel 6 611 174 personer, slik som SSB har beregnet i hovedalternativet) er null⁴. Når brukeren ikke har tilgang til en slik sannsynlighetsprognose, men kun til et begrenset antall verdier (for eksempel 5 784 120, 6 611 174 og 8 138 945 personer i 2050 i SSBs alternativer med henholdsvis lav, middels og høy befolkningsvekst), vil en ofte velge en av disse verdiene basert på subjektive eller politiske grunner. Et firma som bygger en ny barneskole er mer interessert i et høyt anslag for det framtidige elevtallet enn kommunen, som må betale for nybygget.

⁴ Her betrakter jeg framtidig folkemengde som en kontinuerlig stokastisk variabel, slik at Figur 2 viser tettheten til denne variabelen.

3.2 Alternativer er inkonsistente når de tolkes som usikkerhetsalternativer

Tabell 1 viser noen resultater fra SSBs befolkningsprognose fra 2014. Beregnet antall personer som er mellom 20 og 66 år gamle i 2040, varierer fra 3,42 til 4,05 millioner. Intervallet mellom disse to verdier er 17,2 prosent, beregnet som andel personer i denne aldersgruppen i hovedalternativet. For aldersgruppen 67+ er tallene 1,10 til 1,41 millioner. Intervallet er litt bredere (24,3 prosent). Eldrekvoten derimot, definert som antall eldre delt på antall personer i alder 20-66 har en relativ intervallbredde på kun 7,5 prosent, mye smalere enn man intuitivt ville forvente. Grunnen til denne overraskende effekten er at den høye levealderen og den høye fruktbarheten i alternativet med høy befolkningsvekst *forsterker* hverandre i tallene på første og andre rad, og tilsvarende for alternativet med lav befolkningsvekst (men nå i motsatt retning). Derimot, når en beregner eldrekvoten *kompenserer* de hverandre. Tallene illustrerer et mer generelt poeng: to alternativer som er ekstreme for en bestemt variabel (for eksempel antall eldre) er ikke nødvendigvis ekstreme for en annen variabel (for eksempel eldrekvoten). Forklaringen er at man i en deterministisk prognose med tre (eller flere) ulike alternativer antar at hvert år når fruktbarheten er høy, er også levealderen høy, og motsatt for lav fruktbarhet og lav levealder. En slik perfekt korrelasjon mellom fruktbarhet og levealder er urealistisk; like urealistisk som perfekt korrelasjon mellom fruktbarhet og nettoinnvandring, eller mellom levealder og nettoinnvandring.

Figur 2. Sannsynlighetsprognose for folketall i 2050 i millioner.



Tabell 1. Antall personer i alder 20-66 og 67+ samt eldrekvoten, 2040, beregnet i tre alternativer

	Lav bef. vekst (LLML)	Middels bef. vekst (MMMM)	Høy bef. vekst (HHMH)	(H-L)/M %
20-66 (mln)	3,42	3,63	4,05	17,2
67+ (mln)	1,10	1,26	1,41	24,3
Eldrekvote (67+/20-66)	0,32	0,35	0,35	7,5

Kilde: SSB.

I tillegg til ovennevnte effekt er det en annen form av inkonsistens knyttet til bruk av deterministiske alternativer. Man antar implisitt at hvert år når fruktbarheten er høy (lav), er det 100 prosent sikkert at den også er høy (lav) neste år. Den samme forutsetningen gjelder levealder og nettoinnvandring. Slike antakelser om perfekt autokorrelasjon er urealistiske. Konsekvensen er at det er alt for stor avstand mellom banene for høy og lav befolkningsvekst på lang sikt, mens det på kort sikt kan være for kort avstand. For eksempel, en analyse av treffsikkerheten for gamle verdensprognoser viser at intervallet mellom høy og lav befolkningsvekst dekket det virkelige tallet for verdensbefolkningen kun i 50 prosent av tilfellene for prognoser 5 eller 10 år framover i tid, men i mer enn 90 prosent av tilfellene for prognoseperioder på 15 år eller mer (NRC 2000, 192).

4. SANNSYNLIGHETSPROGNOSER

Når framtiden er usikker, og når noen befolkningsutviklinger er mer sannsynlige enn andre, sier det seg selv at et anslag på framtidens befolkning bør beregnes som en sannsynlighetsprognose, og ikke som en deterministisk prognose. Dette er etablert praksis blant meteorologer (jfr. <http://www.yr.no/artikkel/slik-beregner-vi-langtidsvarslelene-1.6498765>) og samfunnsøkonomer (jfr. Norges Banks anslag på framtidig konsumprisindeks og styringsrente i Pengepolitisk rapport 1/16, samt SSBs anslag på arbeidsledighet til 2017 i Økonomisk analyse 2016/1). I mange tiår har også demografer etterlyst en slik form for prognose, jfr. Keyfitz (1981): "Demographers can no more be held responsible for inaccuracy in forecasting population 20 years ahead, than geologists, meteorologists, or economists when they fail to announce earthquakes, cold winters, or depressions 20 years ahead. What we can be held responsible for is warning one another and our public what the error of our estimates is likely to be." Til og med representanter fra statistiske byråer med ansvar for befolkningsprognoser konkluderte i 2007 at "Demographers should

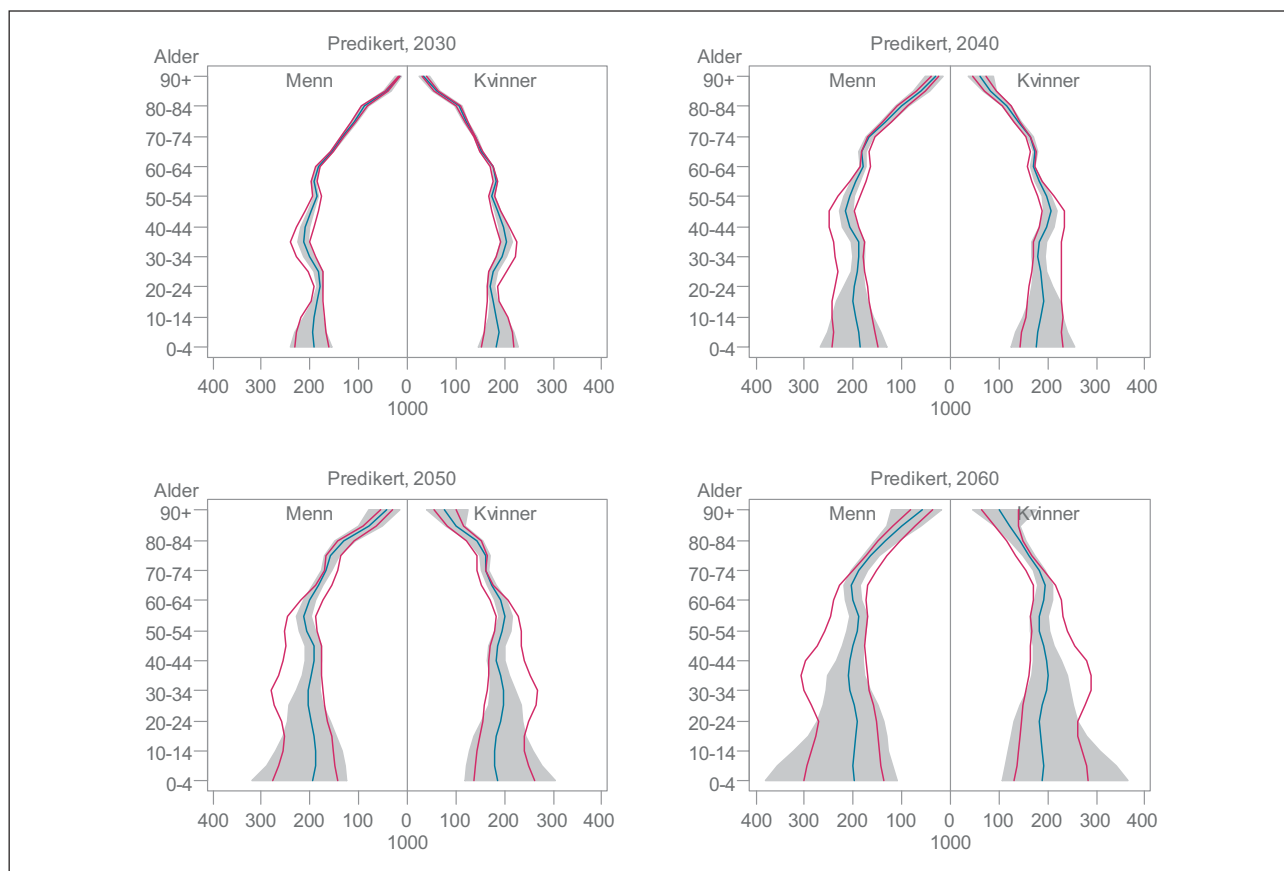
produce projections with a degree of uncertainty, but they have also the responsibility to explain to the users how to understand the projections. At the beginning, this effort may take some time to produce results, but it should be rewarding in the long-term" (UNECE/EUROSTAT 2007). I senere tid har det blitt utviklet både metoder og programvarer for å beregne slike sannsynlighetsprognoser med anslag for den framtidige befolkningen fordelt etter kjønn og alder (Bijak mfl. 2015).

4.1 Eksempler

Pr. i dag er det kun de statistiske byråene i Nederland og New Zealand som publiserer sine offisielle befolkningsprognoser i form av sannsynlighetsprognoser⁵. For byråene i en del andre land finnes det mer eksperimentelle beregninger, for eksempel Italia og Sverige. Norge er også med i denne gruppen: Aslaug Foss publiserte i 2012 en artikkel i Økonomisk analyse der hun beskrev en sannsynlighetsprognose for Norges befolkning fordelt etter kjønn og alder fram til 2060; se Foss (2012a). Utgangspunkt for denne beregningen var den registrerte befolkningen pr. 1. januar 2011. Tallene for Figur 2 kommer fra denne prognosen. Metoden for å beregne en sannsynlighetsprognose er dokumentert i artikkelen til Foss og i litteraturen som hun viser til. Her begrenser jeg meg til noen resultater.

Figur 3 viser befolkningspyramider for 2030, 2040, 2050 og 2060. Usikkerheten er kvantifisert i form av 80 prosent prediksjonsintervaller. Tolkningen er som følger: sannsynligheten for at antall menn eller kvinner i en bestemt aldersgruppe kommer til å ligge innenfor intervallet et bestemt år anslås å være 80 prosent. Ønsker man større sikkerhet, for eksempel 90 prosent, blir intervallene tilsvarende bredere. Vi ser at usikkerheten øker jo lenger fram i tid vi går. Dette er som forventet. Videre er det klart at den nederste delen av pyramiden har store intervaller, særlig på lang sikt: det er vanskelig å gi presise anslag for antall barn og ungdommer på lang sikt, fordi både *antall* kvinner i reproduktiv alder (stort sett mellom 15 og 50 år) OG deres *tilbøyelighet* til å få barn er usikre. På lang sikt er også antall eldre usikkert. Når en sammenligner de røde kurvene (alternativer for høy og lav vekst fra SSBs befolkningsprognose) med

⁵ Jfr. <http://statline.cbs.nl/Statweb/publication/?DM=SLLEN&PA=82683eng&D1=0,7,10-12,14-15,21&D2=0-4&D3=0-1,6,11,16,21,26,31,36,41,I&LA=EN&VW=T> og http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/NationalPopulationProjections_HOTP2011.aspx. Her er det verdt å nevne at FNs Population Division, som har ansvar for å beregne befolkningsprognoser for alle land i verden, også publiserer sannsynlighetsprognoser; se Raftery mfl. (2012) og <http://esa.un.org/unpd/wpp/Graphs/Probabilistic/POP/TOT/>.



Figur 3. Befolkningspyramider med 80 prosent prediksjonsintervall i grått og røde linjer for lav- og høyalternativet fra SSBs deterministiske befolkningsprognose fra 2011 (alternativer for høy og lav befolkningsvekst). Predikert 2030, 2040, 2050 og 2060

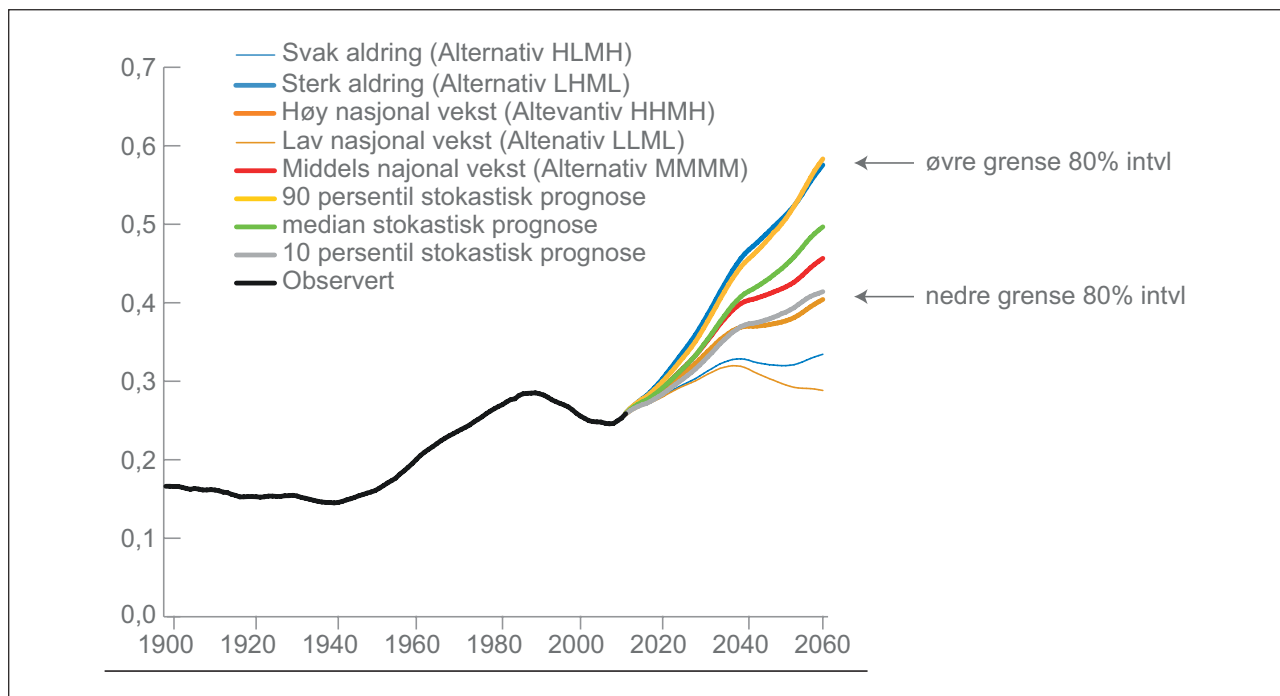
Kilde: Foss (2012a)

prediksjonsintervallene kommer det klart fram at SSBs prognose underestimerer usikkerhet for barn og eldre og overestimerer den for befolkningen i arbeidsdyktig alder. Dette er enda en illustrasjon av poenget nevnt i avsnitt 3.2.

Figur 4 viser 80 prosent prediksjonsintervallet for eldrekvoten (her definert ved aldergrensen 64/65 år) for årene 2012-2060. For et bestemt år, f. eks. 2035, anslås sjansen for at intervallet kommer til å dekke den virkelig eldrekvoten lik 80 prosent. Det er *ikke* 80 prosent sikkert at den framtidige eldrekvoten forblir *hele tiden* innenfor intervallet, men kun for hvert enkelt år. Derfor kalles prediksjonsintervallene i Figur 4 for *årwise* (eller mer generelt *punktwise*) prediksjonsintervaller. Slike intervaller er nyttige for en bruker som ønsker prognoseresultater for ett enkelt år. Når man er interessert i en sannsynlighetsprognose for en periode på to eller flere år, trenger man resultater i form av et prediksjons**band**. Tolkningen er som beskrevet ovenfor: det er 80 prosent sikkert at eldrekvoten forblir innenfor

prediksjonsbandet hele tiden. Et prediksjonsband for to eller flere år er bredere enn prediksjonsintervallet for enkelt år i denne perioden, ofte (avhengig av antall år i perioden og autokorrelasjonen i den aktuelle variabelen) *mye* bredere.

Usikkerheten på lang sikt er betraktelig, jfr. 80 prosent intervallet i 2060. En mulig konsekvens er at myndighetene som overvåker dagens folketrygdsystem kan forvente store utfordringer i andre halvdel av dette århundret: det er vanskelig å predikere forholdet mellom antall eldre og antall personer i yrkesaktiv alder på lang sikt. Dermed er det også betraktelig usikkerhet knyttet til bærekraften i dagens system.



Figur 4. Eldrekvoten: forhold mellom antall eldre (65+) og antall i yrkesaktiv alder (20-64). Registrert 2011, sannsynlighetsprognose for årene 2012-2060.
Kilde: Foss (2012)

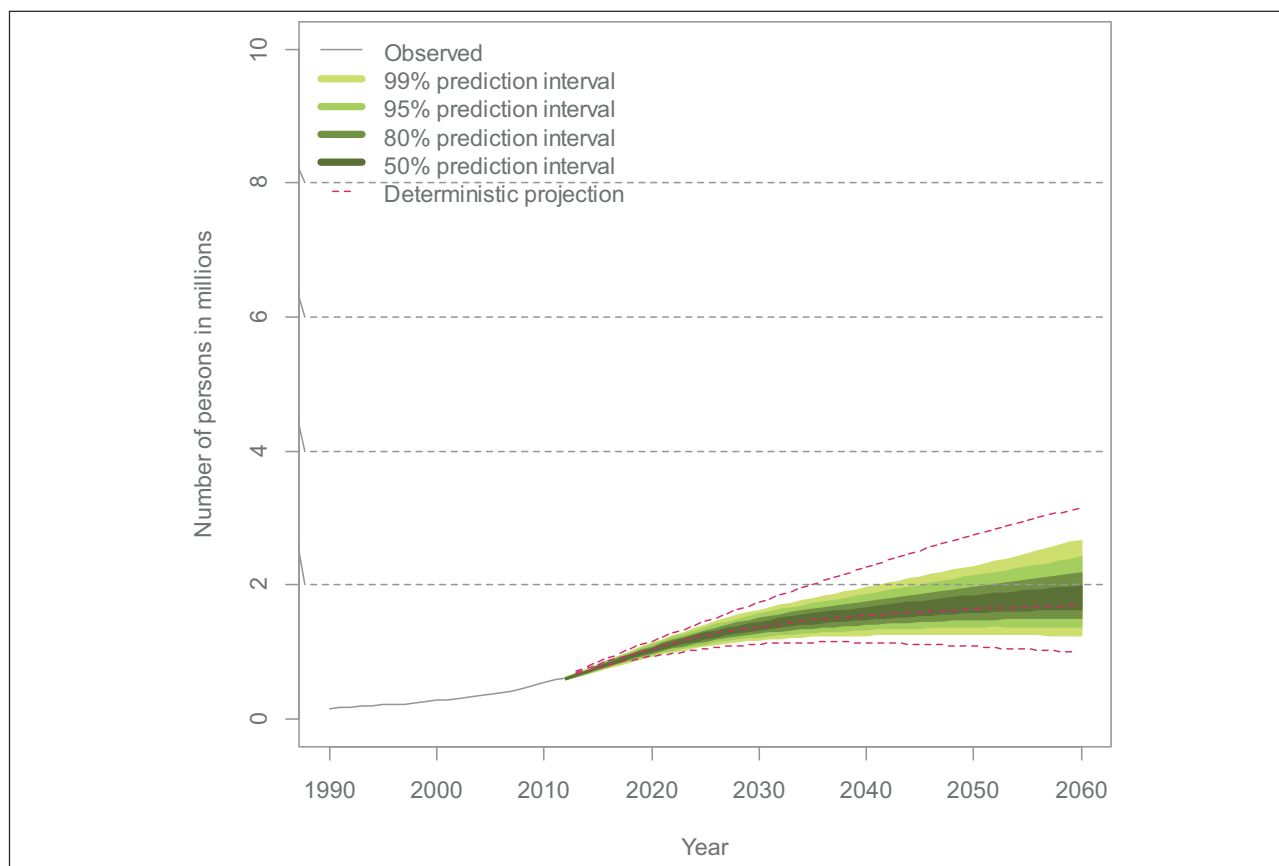
4.2 Utfordringer

Moderne statistikk og sannsynlighetsteori gir oss muligheten til å formulere prognosemodeller slik at man kan beregne prediksjonsintervaller. I tillegg finnes det programvare for å gjøre dette på en enkel måte. Allikevel er det ikke vanlig praksis blant statistiske byråer å publisere sannsynlighetsprognoser for den demografiske utviklingen. Hva kan dette skyldes?

1. En mulig grunn er at statistikkbyråene mener å ta hensyn til at mange brukere ikke er vant til å bruke befolkningstall i form av sannsynlighetsfordelinger. En spørreundersøkelse under 27 nasjonale statistiske byråer som har ansvar for landets befolkningsprognose viser at en betydelig del av disse byråene mener at deres brukere foretrekker kun et beregningsalternativ, eller at brukerne mangler innsikt i måten en kan håndtere usikkerheten på (UNECE Task Force on Population Projections, 2016). Dette poenget er en av flere grunner som Fischhoff (2012) nevner når han analyserer spørsmålet hvorfor forskere ofte er motvillige mot å være eksplisitte angående usikkerheten i deres funn. Andre grunner han nevner er: å være eksplisitt på denne måten kan av noen blitt tolket slik at forskerne er unødvendig eller overdreven forsiktige; kan føre til mindre prestisje for disse forskerne; og (kanskje viktigst) det er ikke

alltid klart hvordan man best skal presenter usikre resultater. Er det optimalt å presentere prognoseresultatene i form av prediksjonsintervaller med en fast prediksjonsgrad (f. eks. 80 prosent, som i Figurene 3 og 4), eller heller som en «vifte» med ulike prediksjonsgrader (som prognosen for innvandrerbefolkningen i Figur 5)? En annen mulighet er å beskrive usikkerheten ved hjelp av ord. En studie som Goodwin(2014) viser til, fant at usikkerheten i prognoser som Klimapanelet (IPCC) har beregnet formidles best til allmenheten når en kombinerer verbale uttrykk med numeriske sannsynligheter. For eksempel, et bestemt utfall refereres til som «sannsynlig, det vil si med en sannsynlighet mellom 67 og 90 prosent», eller som «meget usannsynlig, det vil si en sjanse på 2 til 10 prosent.»

Selv om de fleste brukerne er uvant med å bruke prognosetall i form av sannsynligheter, er mange av dem allikevel interessert i slike tall. Ovennevnte spørreundersøkelse finner at 69 prosent av en gruppe av 148 brukere vurderte det som «viktig» (40 %) eller «meget viktig» (29 %) at usikkerheten skulle foreligge i kvantifisert form. Rapporten understreker at denne gruppen på 148 ikke kan betraktes som et tilfeldig utvalg, siden de var representantene for de 27 statistiske byråene som var med i arbeidsgruppen som valgte dem ut. Samtidig gir svarene en pekepinn.



Figur 5. Sannsynlighetsprognose for innvandrerbefolkningen
Kilde: Foss (2012b)

2. En annen utfordring er at mange brukere har behov for tall som er mye mer detaljerte enn det som Figur 3 viser: den framtidige befolkningen er ikke bare fordelt etter kjønn og alder for landet under ett, men også på lavt regionalt nivå (kommune, bydel), eller fordelt etter innvandrerstatus, etter husholdningstype, osv. Kommuneplanleggere har ofte behov for tall angående utviklingen i deres kommune, av og til supplert med opplysninger om nabokommunen. Mens teknikker for sannsynlighetsprognoser som bare gir detalj for alder og kjønn er velkjente, pågår det arbeid med mer kompliserte modeller. Et viktig problem her er å estimere korrelasjonen mellom mange ulike delbefolkninger, for eksempel antall menn og kvinner i ett-års aldersgrupper fordelt etter husholdningsstørrelse (1, 2, ... 6 eller flere personer). En metode som er basert på stokastiske andeler - «random shares»; se Alho og Keilman (2010) for husholdninger, eller Foss (2012b) for innvandrerbakgrunn – gir lovende resultater. Figur 5 er et eksempel for befolkningen med innvandrerbakgrunn (både utenlandskfødte med to utenlandskfødte foreldre, og deres barn født i Norge).

5. AVSLUTNING OG ANBEFALINGER

I denne artikkelen har jeg gjort rede for at SSB bør beregne sin befolkningsprognose som sannsynlighetsprognose. Dagens praksis med et begrenset antall deterministiske alternativer har begrenset verdi, all den tid en ikke vet hvor sannsynlig det er at den virkelige utviklingen kommer til å ligge innenfor en bestemt bane definert av to slike alternativer. Noen få brukere ønsker kanskje å anvende dem for å analysere hvor robuste deres konklusjoner er for endringer i befolkningsstallene. Problemet her er å avgjøre hvor langt ut en bør og kan gå sammenlignet med hovedalternativet. Det er vanskelig å bedømme hvor sannsynlig det ene eller det andre alternativet er så lenge SSB ikke oppgir noen sannsynligheter.

Framtiden er usikker, men den ene befolkningsutviklingen er mer sannsynlig enn den andre. Derfor bør SSB gjøre som meteorologer og samfunnsøkonomer, som også lager sine anslag i form av sannsynlighetsfordelinger. En slik prognose er standard for god forskning, er viktig for å ta gode beslutninger, og mange brukere ønsker å ha innsikt i

prognoseusikkerheten i kvantifisert form, i motsetning til det man ofte tror (UNECE Task Force, 2016).

Metoder for å lage sannsynlighetsprognoser for befolkningen fordelt etter alder og kjønn er velkjente, og programvarer foreligger. Dessuten har SSB betydelig kunnskap om og verdifull erfaring med å beregne slike prognoser. Brukere som ikke ønsker å få utlevert en prognose basert på en sannsynlighetsfordeling, kan få tall for forventet verdi eller medianen i fordelingen, eventuelt supplert med tall for den nedre og øvre grensen til et 80 prosent prediksjonsintervall. En slik leveranse erstatter gammel praksis med å levere tall for hovedalternativet, supplert med tall for to alternativer for høy og lav befolkningsvekst.

Mer konkret innebærer mitt forslag at SSB ved neste korsvei (i 2018?) beregner og publiserer en sannsynlighetsprognose for befolkningen fordelt etter alder og kjønn for landet under ett. Om en slik prognose også inkluderer resultater for antall innvandrere bør vurderes. Samtidig kan SSB gjennomføre et forskningsprosjekt der formålet er å beregne en sannsynlighetsprognose for befolkningen i hver kommune, igjen fordelt etter alder og kjønn. En trinnvis tilnærming kan være fornuftig, der landets befolkning (gitt alder, kjønn og prognoseår) først fordeles over et begrenset antall landsdeler ved hjelp av stokastiske andeler, for deretter å fordele befolkningen i hver landsdel over prognoseregioner eller over fylker, og til slutt over kommuner. Fordelingen skjer tilfeldig. Hver andel antas å ha en statistisk fordeling med forventet verdi fra den deterministiske prognosen, og en varians som kan estimeres fra tidsserier med observerte andelsverdier.

REFERANSER

Alho, J. og N. Keilman (2010) On future household structure. *Journal of the Royal Statistical Society Series A* 173 (1), 117-143.

Bijak, J., I. Alberts, J. Alho, J. Bryant, Th. Buettner, J. Falkingham, J. Forster, P. Gerland, Th. King, L. Onorante, N. Keilman, A. O'Hagan, D. Owens, A. Raftery, H. Ševčíková og P. Smith (2015). Letter to the editor: Probabilistic population forecasts for informed decision making. *Journal of Official Statistics* 31 (4), 537-544.

Fischhoff, B. (2012). Communicating uncertainty: Fulfilling the duty to inform. *Issues in Science and Technology* 28 (4). <http://issues.org/28-4/fischhoff/>.

Foss, A. (2012a). Stokastiske befolkningsprognoser for Norge 2012-2060. *Økonomiske analyser* 2/2012, 30-34.

Foss, A. (2012b). Stochastic forecast of immigrant population. Paper Nordisk Demografisk Symposium Tønsberg, 13.-15. september 2012.

Goodwin, P. (2014). Getting real about uncertainty. *Foresight* 33 (Spring 2014), 4-7.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review* 8 (44), 579-593.

NRC - National Research Council (2000). *Beyond six billion: Forecasting the world's population*. Panel on Population Projections. John Bongaarts and Rudolfo Bulatao red. Washington DC: National Academy Press.

Raftery, A., N. Li, H. Ševčíková, P. Gerland, G. Heilig (2012). Bayesian probabilistic population projections for all countries. *PNAS* 109 (35), 13915-13921.

Tønnessen, M., A. Syse og K. Norgård Aase (2014). Befolkningsframskrivinger 2014-2100: Hovedresultater. *Økonomiske analyser* 4/2014, 30-36.

UNECE Task Force on Population Projections (2016). Key recommendations and good practices in the communication of population projections. Paper joint Eurostat-UNECE Work Session on Demographic Projections, Geneva, 18.-20. april 2016. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.11/2016/WP_28_TF_Report_v5_WithAppendix_.pdf.

UNECE/EUROSTAT (2007). Report Joint Eurostat-UNECE Work Session on Demographic Projections Bucharest, 10.-12. oktober 2007 <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.11/2007/1.e.pdf>.