Matematisk Modellering: Et–barnspolitik og befolkningstal

Kasper Bjering Søby Jensen, ph.d-studerende i matematikkens didaktik ved Roskilde Universitet

I LMFK–bladet 4/2012 udfoldede jeg et eksempel på "fuldbyrdet matematisk modellering" (se LMFK–bladet 2/2012 for uddybning af begrebet). Min grundidé er, at et væsentligt element i oparbejdelse af matematisk kompetence i almendannende matematikundervisning, er at kunne gå til åbne ikke–matematiske problemstillinger med et matematisk undersøgelsesapparat. Dette vil jeg gerne give gode eksempler på, som i sit tekniske niveau passer ind i den gymnasiale undervisning. Som afrunding i 4/2012 om *"hvornår Venus står op"*, udstak jeg en ny sådan problemstilling, som man har kunnet tygge på siden. Her er mine egne overvejelser over den. Problemstillingen lød:

> *»Hvordan udvikler befolkningstallet sig i et land med et–barnspolitik?«*

Problemstillingen har oplagt en berøring med matematik, idet der indgår størrelsen "befolkningstal". Ordet "udvikling" vil få personer med erfaring i typeopgaver til at lede efter datapunkter, der kan laves regression på – men de er ikke her. Den eneste anden oplysning er betingelsen "et–barnspolitik". Det kan lede tankerne hen på landet Kina, og man kan gå på jagt efter empiri i form af datapunkter gældende for dette land.

Her vil jeg gribe det analytisk an. Kan vi ud fra nogle antagelser sige noget principielt om spørgsmålet. Typisk kan man ved at beskrive en situation i dens mest ideale form sige noget om rammen for systemets opførsel. I termer af "modelleringscirklen" (se 2/2012) handler det altså om at opstille et system og oversætte det til et matematisk system.

Som udgangspunkt kan vi lade *befolkningstallet* være repræsenteret ved en tidsafhængig variabel $B(t)$. Udviklingen i befolkningstallet kan repræsenteres ved en vækstrate, der i idealsituationen antages konstant – vi kalder den r. Vækstraten er givet som differensen mellem tilkomne og frafaldne. Hvis vi ser bort fra migration, er det altså differensen mellem fødselsrate (f) og dødsrate (d) – dvs. $r = f - d$.

Fødselsraten defineres her som antallet af nyfødte børn (F) delt med befolkningstallet – dvs. $f = \frac{F}{B}$.

Da F oplagt må afhænge af B, kan man passende overveje denne sammenhæng. Hvis vi antager, at et–barnspolitikken betyder, at hver borger kun får ét barn, så vil hvert barn skabes af et par, hvis to medlemmer ikke får flere børn. Antallet af sådanne par (P) er en andel af befolkningen. Der gælder altså at $P = k \cdot B$ og der må gælde at $0 \leq k \leq 1/2$.

Hvor mange børn får det samlede antal par så? Hvis vi – meget idealiseret – antager at kvinderne er fordelt ligeligt på årgange og at gennemsnitslevetiden er T, så må det forventes, at $\frac{1}{T}$ af kvinderne føder et barn i løbet af et år. Heraf følger altså nu:

$$F = \frac{P}{T} = \frac{kB}{T} \Rightarrow f = \frac{\frac{kB}{T}}{B} = \frac{k}{T}$$

Og der kan endvidere af dette udledes:

$$0 \leq f \leq \frac{1}{2T}$$

Da det ud fra idealiseringen også relativt nemt må indses, at $d = \frac{1}{T}$ fås altså fra $r = f - d$:

$$-\frac{1}{T} \leq r \leq -\frac{1}{2T}$$

Med lidt almindelig viden om vækst med konstant rate – eller ved løsning af differentialligningen $\frac{dB}{dt} = rB$ – fås nu: $B(t) = B_0 \cdot e^{rt}$. Det ses oplagt, at i alle tilfælde vil befolkningstallet aftage, fordi r er negativ. Derudover kan undersøges hvordan udviklingen afhænger af k og T. Situationen $k = \frac{1}{2}$ giver så at sige den "optimale" situation:

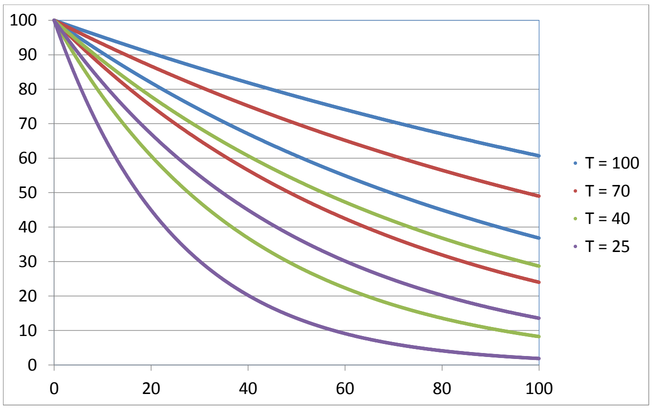
$B(t) = B_0 \cdot e^{-\frac{t}{2T}}$, mens $k = 0$ giver den anden ekstrem.

Udviklingsrummet for en befolkning med en given gennemsnitsalder kan herefter afbilledes ved graferne for de to ekstremer. På figur 1 er dette gjort for 100 år, 70 år, 40 år og 25 år. En fortolkning kan altså være, at med et–barnspolitik vil et samfund med en gennemsnitslevealder på 70 år blive reduceret med mellem 50% og 75% i løbet af 100 år.

Det der selv overraskede mig første gang jeg modellerede denne problemstilling var, at svaret først og fremmest afhang af gennemsnitslevealderen. Retrospektivt er det meget oplagt, men uden en analytisk behandling med en matematisk model, ville jeg næppe være kommet til resultatet.

Elever kan her stilles reflekterende spørgsmål. Fx hvorfor det ikke betyder noget, hvor tidligt kvinderne får deres børn. Hvis de får dem i 15 års alderen, går der jo kun 15 år før deres egne børn også får børn, osv. Får de dem i 40 års alderen skulle man synes det var anderledes. Hvorfor er det ikke det? Svaret har med idealiseringerne at gøre. Det er ligegyldigt, hvornår en kvinde får sit barn, hovedsagen er, at hun får ét og kun ét i løbet af sin levetid på T år.

Dertil kan man også reflektere over følgende afledte problemstilling: *Hvem deltager i kinesiske børns fødselsdagsfest?* Et-barnspolitik betyder jo oplagt at man ikke har nogen søskende. Men man har heller ikke nogen fætre/kusiner. Faktisk er et kinesisk barn eneste person i sit slægtsled. Konsekvensen er at "familien" opløses som institution (måske Kinas virkelige intention med politikken?). Studiet af "stamtræs–strukturer" er efter min mening en ganske hæderlig matematisk aktivitet.



Figur 1
Udviklingsrummet for en befolkning med et–barnspolitik, når gennemsnitsalderen er hhv. 100, 70, 40 og 25 år. På 1. aksen er givet antal år, mens der på 2. aksen er givet befolkningens udvikling i % i forhold til startåret.

Demografisk model

Den faglige side af mig, som er fysiker, synes der er noget æstetisk smukt over at opstille en idealiseret model med en række faktorer, som alle går ud til fordel for afhængighed af én størrelse. Derfor viste jeg begejstret ovenstående til en fysikerkollega, som dog uimponeret belærte mig om, at det virkeligt svære i demografi er, at antagelsen om ligefordeling på årgange ikke holder.

Øv! På den anden side er det jo bare en udfordring om at gå i gang med at bygge en model, som tager afsæt i en ikke–ligefordeling af befolkningen på årgange. Arbejdet med sådanne modeller kan være oplagte i samspil med samfundsfag, hvor demografiske fremskrivninger ofte bruges til at forudsige eksempelvis den fremtidige forsørgerbyrde.

Hvis afsættet er faktiske data fra Danmark, må man bygge modellen ud fra tilgængelige data. Hos statistikbanken.dk kan nemt hentes data for bl.a. befolkningens fordeling på aldersgrupper hvert år i perioden 1980–2012. Og tilsvarende kan man finde data for antal døde på de forskellige alderstrin, i de forskellige år. Afsættet kan altså være data for året 1980.

Lad $B_{i,t}$ betegne antal med alderen i for året t og lad på tilsvarende vis $D_{i,t}$ betegne antal døde. En antagelse kan nu være, at dødeligheden kun afhænger af alder. Den er derfor konstant for et givet alderstrin. Lad således d_i være dødeligheden for alderstrinnet i, som konkret kan findes som: $d_i = \dfrac{D_{i,1980}}{B_{i,1980}}$.

Der skal endvidere tages højde for fødsler. Dette kan antages kun at afhænge af antal i "forældre–alderen", som vi kan sætte til "20–39 år". Deraf kunne en konkret fødselsrate f bestemmes ved:

$$f = \frac{B_{0,1980}}{\sum_{i=21}^{40} B_{i,1980}}$$

Der kan nu formuleres et konkret bud på en matematisk model for befolkningsudvikling efter 1980:

$$B_{0,t+1} = f \cdot \sum_{i=20}^{39} B_{i,t}$$

$$B_{1,t+1} = B_{0,t}\left(1 - d_0\right)$$

$$B_{2,t+1} = B_{1,t}\left(1 - d_1\right)$$

$$\vdots$$

$$B_{i+1,t+1} = B_{i,t}\left(1 - d_i\right)$$

$$\vdots$$

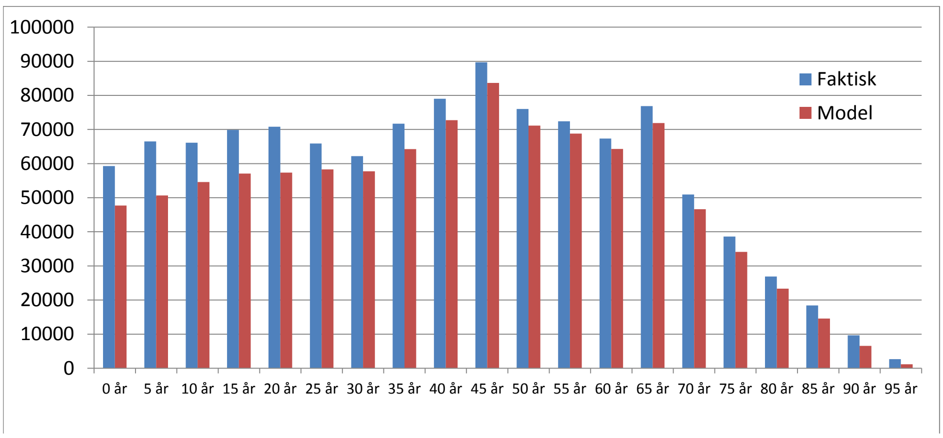
$$B_{99,t+1} = B_{98,t}\left(1 - d_{98}\right)$$

$$B_{100,t+1} = 0$$

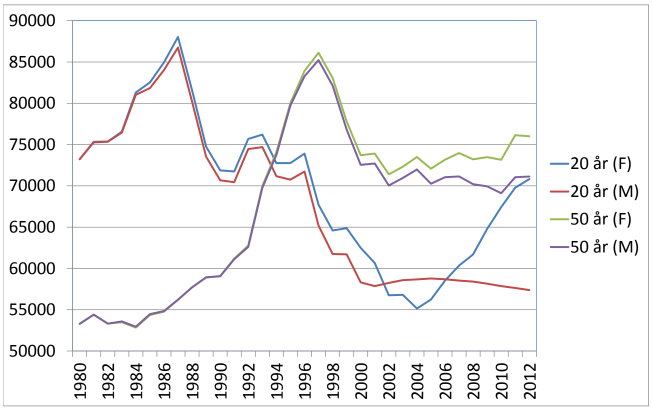
Med data fra et givet år, kan modellen implementeres i Excel. I det følgende har jeg således lavet en 1980–model baseret på talsættene $B_{1980} = (B_0, B_1, \ldots, B_{99})$ og $D_{1980} = (D_0, D_1, \ldots, D_{99})$. Herudfra er beregnet $d = (d_0, d_1, \ldots, d_{99})$ og f. Samtidigt eksisterer der et faktisk datasæt for perioden 1980–2012, som modellen kan holdes op imod.

På figur 2 ses for året 2012 både faktiske og modelberegnede tal for, hvor mange personer der er på de forskellige alderstrin (kun hvert femte er medtaget). Det ses at modellen generelt skyder under de faktiske tal. Man kan også ane at den skyder mere galt for de små aldre end for de større.

På figur 3 er afbildet den tidslige udvikling af 20–årige og 50–årige, både faktisk og modelberegnet. For 50–årige er der enorm samstemmighed de første 20 år og derefter pæn sam-



Figur 2
Faktiske tal for antal med given alder i 2012 (hvert femte alderstrin medtaget) og antal forudset af 1980–model.



Figur 3
Antal 20–årige hhv. 50–årige til forskellige årstal (1980–2012) – Faktiske tal (F) og 1980–modelberegnede (M).

stemmighed resten af perioden. For de 20–årige er der stor samstemmighed til ca. 1990 og pæn frem til ca. 2005. Herefter stiger det faktiske tal mærkbart, men modellen forudsiger slet ikke denne stigning.

På figur 4 er det derfor undersøgt, hvordan modellen opfører sig på lang sigt. Her er 1980–modellen fremskrevet til 2200 for fem alderstrin – 15, 30, 45, 60 og 75 år. Ikke overraskende har kurverne ens ”form” med forskydning. Det ses, at de alle ender ud i samme eksponentielt aftagende kurve. Matematisk relevant vil det være at undersøge denne kurve og dens sammenhæng med *f* og *d*. Dermed vil der også kunne gives nogle bud på, hvornår en befolkning vokser hhv. mindskes.

Lidt overfladisk analyse af kurverne viser, at det begynder at ”gå galt” når alderstrinnet overtages af årgange født i modelperioden. Når modellen evalueres kritisk, vil det altså være oplagt at undersøge, om antagelserne i forhold til fødsler holder, samt søge at tilpasse mekanikken i modellen efter det. På tilsvarende vis kan det overvejes, om dødsraterne bare kan holdes konstante.

På figur 5 er der ud fra faktiske data beregnet fødselsraten samt dødsrater for 40 og 70–åriges (i forhold til 1980 som er sat til 100). Her er det afgørende at huske, at mens data er faktiske, er størrelserne fødsels– og dødsrate modelkonstruktioner. På figuren ses, at fødselsraten stiger 20%, mens dødsraterne synes at aftage med over 40%. Antagelsen om konstans er nok for misvisende. Men hvordan kan vi så fremskrive befolkningen efter 2012? Det kræver antagelser. Som det kan ses af 1980–modellen, kan sådanne antagelser være ret misvisende i længden. Så det er oplagt at undersøge, hvordan sådanne antagelser kan gøres – eller rent faktisk gøres.

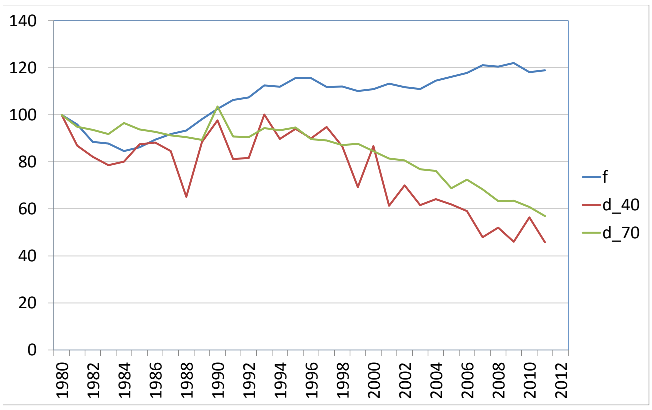
Endeligt kan man også diskutere modellens afgrænsninger. Hvilke konsekvenser har det fx, at modellen alene ser på fødsel og død, men ikke på migration? Også her kan man statistikbanken.dk levere data, som kan bruges til at undersøge dette.

Didaktiske pointer

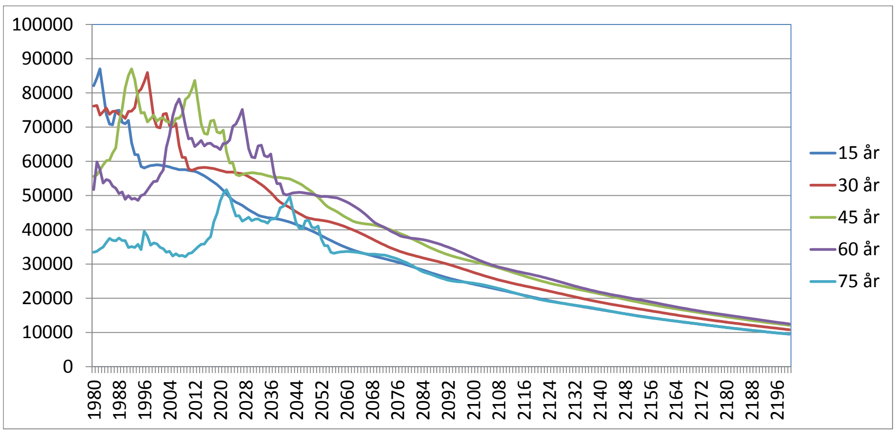
Formålet med denne artikel og undervisningsaktiviteter, baseret på den, er todelt. For det første handler det om at vise, at matematik er et værktøj, der kan bruges til noget, og herunder at vise, at bare fordi man bruger det, betyder det ikke, at man opnår indiskutable svar.

For det andet handler det om at vise, at anvendelse af matematik faktisk har matematisk karakter. Der er mange spørgsmål, som man kan stille og søge svar på, der kræver undersøgelser af matematiske systemer. Anvendelse af matematik behøver således langt fra at antage ”black box”–karakter, hvis man i øvrigt griber fat i modelleringsprocessen selv.

Endeligt rummer det selv at gribe fat i en modelleringsproces nogle ganske vide muligheder for at arbejde kritisk med modeller og modellering. Altså at forstå, at forudsigelser, der stammer fra modeller, bygger på antagelser om ting, som faktisk kan være svære at forudsige.



Figur 5
Udvikling i fødselsraten samt dødsraten for hhv. 40 og 70–årige 1980–2011.



Figur 4
Antal personer på fem forskellige alderstrin, beregnet med 1980–modellen for perioden 1980–2200.