Dommedag nu?

T. Døssing, A. D. Jackson og B. Lautrup
Niels Bohr Institutet

23. oktober 1998

Der har altid været fanatikere, som har ment, at dommedag var nær, og for en del år siden kom nogle naturvidenskabelige forskere til nøjagtig samme konklusion [1, 2, 3, 4]. Ikke ved at påvise muligheden af en altomfattende naturkatastrofe eller en frygtelig sygdom, men blot ud fra en vurdering af det totale antal mennesker, der indtil nu har levet her på Jorden.

Med statistiske argumenter i hånden hævdede disse forskere, at det mest sandsynlige ville være, at menneskehedens fremtid ville blive lige så lang som dens fortid, målt i antal individer. Alt i alt ville det mest sandsynlige totale antal individer af menneskearten altså blive dobbelt så stort, når regnskabet endeligt skulle gøres op. Med den nuværende befolkningstilvækst vil et antal af denne størrelsesorden nås inden for en overskuelig tid, og denne forudsigelse burde derfor tages meget alvorligt af os alle. Dommedag er nu, næsten!

Det samme argument anvendes også på levetider, hvad enten det drejer sig om den tid et teaterskuespil vil være på plakaten, eller hvor længe socialdemokratiet vil fortsætte med at beholde regeringsmagten. Hvis der ikke er nogen særlig sammenhæng mellem et fænomen og det tidspunkt, hvor det observeres, altså hvis alle observationstidspunkter er lige sandsynlige, så påstås det, at fænomenet med 95% sandsynlighed vil fortsætte sit liv mellem 1/39 og 39 gange den tid, det allerede har levet. Brugt på socialdemokratiet, som nu har regeret i 5 år, er forudsigelsen, at det med 95% sandsynlighed fortsætter mindst 45 dage endnu, men ikke ud over år 2193!

Disse påstande og den deraf følgende debat [5, 6] er for nylig igen kommet frem [7] og har vakt opmærksomhed i dagspressen [8, 9]. Men gentagelse gør ikke et forkert argument rigtigere. Statistik er ofte blevet brugt til at underbygge påstande, som enhver kan se strider mod sund fornuft, og dette er også tilfældet her. Argumentet for en snarlig dommedag er simpelthen baseret på fejlagtig anvendelse af statistik.

Kort fortalt består fejlen i, at en enkelt observation af et fænomen ikke kan danne grundlag for statistisk forudsigelse. Enhver statistisk forudsigelse baseret på en enkelt observation vil næsten kun fortælle noget om de *a priori* forestillinger, der gøres om fænomenet. Det er klart, at hvis man observerer, at socialdemokratiet har regeret i 5 år, så kan man uden videre konkludere, at det samlet vil være ved magten i mere end 5 år. Det er der ingen usikkerhed om. Men enhver forudsigelse af, hvor længe det vil bevare magten ud over de 5 år, afhænger af, hvorledes man forestiller sig, at det kan miste magten, og hvilken sandsynlighed man *på forhånd* tillægger de forskellige scenarier, hvorved det kan ske.

Det er naturligvis interessant at diskutere de forskellige scenarier, men så længe man ikke har yderligere information, om hvilke der vil blive realiseret, er det ikke muligt at sige mere. Påstandene er subjektive vurderinger af, hvad fremtiden kan bringe. Der findes ikke noget generelt statistisk princip, som kan fortælle os, hvor længe menneskeheden eller socialdemokratiet vil fortsætte, baseret alene på kendskab til det antal mennesker, der har eksisteret, eller den tid, socialdemokraterne foreløbig har været ved magten.

I det følgende vil vi ved hjælp af mere formelle argumenter præcisere de statistiske betragtninger for derved klart at afsløre de fejl, der er begået af dets stærkeste fortaler, astrofysikeren J. Richard Gott III fra Princeton universitetet [4, 7].

A priori sandsynlighed

Det er nemmest at analysere levetider, fordi det tillader os naturligt at benytte kontinuerte variable og integraler. Argumenterne kan uden videre oversættes til det diskrete problem om menneskehedens størrelse, fordi det totale antal individer er så kolossalt.

Intet varer evigt. Ethvert system eller fænomen har en endelig levetid, som vi betegner med T, som må være et positivt reelt tal[1]. Lad os antage, at vi til at begynde med intet ved om, hvor længe systemer af en given type faktisk vil leve, men at vi er i stand til at observere en masse sådanne systemer blive skabt, leve deres liv, for derefter at dø. Da er det muligt ved simpel optælling at bestemme en tilnærmelsesvis kontinuert fordeling af levetider, som vi skal kalde $p(T)$. Ved hjælp af denne funktion kan vi angive

[1]Hvis vi diskuterer menneskehedens størrelse angiver T det totale antal individer, der har levet, lever nu eller kommer til at leve. Alle argumenter forløber ellers på samme måde i de to tilfælde.

hyppigheden for levetider, $p(T)\,dT$, altså hvilken brøkdel af systemer, der har levetid i et lille interval mellem T og $T + dT$.

I den idealiserede grænse, hvor der er uendeligt mange systemer, kaldes hyppighedsfordelingen en sandsynlighedsfordeling. Sådanne fordelinger er velkendte i mange sammenhænge. For eksempel anvender forsikringsselskaber menneskers levetidsfordeling til at beregne præmier på livsforsikringer. Når vi antyder, at fordelingen $p(T)$ kun afhænger af selve levetiden T, men ikke af det præcise tidspunkt for systemets fødsel, udtrykker det, at verden er den samme til alle tider, i hvert fald hvad angår disse systemer. Dette er tilnærmelsesvis tilfældet for menneskers levetid indenfor en generation eller to, men bestemt ikke over mange generationer. Uden en sådan antagelse, kunne man slet ikke have livsforsikringsselskaber.

Hyppighedsfordelingen og dermed sandsynlighedsfordelingen er normaliseret, hvilket udtrykkes ved ligningen

$$\int_0^\infty p(T)\,dT = 1. \tag{1}$$

Denne ligning betyder såmænd blot, at ethvert system må have en eller anden levetid. Summen af alle levetidshyppigheder skal da være 100%.

For de fleste systemer har vi en vis viden om levetidsfordelingen. Selv om det måske ikke er muligt direkte at observere systemer gennem hele deres levetid fra fødsel til død, så er det somme tider muligt at opnå indirekte viden om levetidsfordelingen. Stjernernes liv er for eksempel kortlagt i så stor detalje gennem observation af stjerner i alle aldre, at der findes temmelig præcise forudsigelser af Solens levetid.

Men der findes unikke systemer, for eksempel livet på Jorden, menneskeheden, den teknologiske civilisation, eller universet, hvor vi kun kender et enkelt eksempel. Så er vi bogstaveligt talt på Herrens mark, hvad angår levetidsfordelingen og er tvunget til at gætte os frem til $p(T)$. Dette er, hvad der menes med *a priori* sandsynligheden for, at et system har en given levetid. Det er den sandsynlighed, vi *på forhånd* vil tillægge levetiden. Lidt mere religiøst kan man kalde det den fordeling, som Vorherre bruger, når han skaber menneskeheder eller universer. For unikke systemer er vi tvunget til at gætte på Hans fordeling. Vi kan gætte mere eller mindre intelligent og med mere eller mindre overbevisende argumenter, men for sådanne systemer er og bliver $p(T)$ et gæt. Hele sandsynlighedsbegrebet kommer faktisk i vanskeligheder, når det ikke baseres på hyppigheder. Sandsynlighed kan da omfortolkes som graden af tiltro, på engelsk "degree of belief".

A posteriori sandsynlighed

Når et systems levetid T er givet, kan man spørge om sandsynligheden for at observere systemet med alderen t, eller rettere i aldersintervallet fra t til $t + dt$. Betegnes denne betingede fordeling med $P(t|T)$ er selve sandsynligheden altså $P(t|T)\,dt$. Den må også være normaliseret

$$\int_0^T P(t|T)\,dt = 1, \tag{2}$$

fordi et system må have en eller anden alder, når det faktisk eksisterer. Læg mærke til, at integralets grænser udtrykker, at observationstidspunktet må ligge inden for levetiden. Sandsynligheden for at $t > T$ er simpelthen nul.

Ved hjælp af prior fordelingen $p(T)$ kan vi nu beregne sandsynligheden for at et tilfældigt valgt system har alderen t. Den betegner vi med

$$q(t) = \int_t^\infty P(t|T)\,p(T)\,dT. \tag{3}$$

Denne ligning udtrykker blot, at et system med alderen t må have en eller anden levetid, der er større end t. Aldersfordelingen er automatisk normaliseret, fordi de andre fordelinger er normaliserede.

Aldersfordelingen $q(t)$ er interessant, fordi den kan bestemmes uden at observere systemerne gennem hele deres liv, men kun forudsætter, at deres nuværende alder kan bestemmes. Hvis vi her og nu optæller hyppigheden af systemer med alderen t, så vil denne hyppighedsfordeling være en tilnærmelse til aldersfordelingen. Kender man aldersfordelingen $q(t)$ og den betingede sandsynlighed $P(t|T)$, så er det muligt ved hjælp af ovenstående formel at regne sig tilbage til *a priori* fordelingen $p(T)$.

Nu kommer vi endelig til det, der er interessant for dommedagsdiskussionen. Lad os spørge om sandsynligheden for at et vilkårligt system har levetiden T, når det vides, at det har alderen t. Dette er også en betinget sandsynlighedsfordeling, som betegnes med $Q(T|t)$. Den kaldes *a posteriori* fordelingen og skal også være normaliseret

$$\int_t^\infty Q(T|t)\,dT = 1. \tag{4}$$

Læg igen mærke til, at alle levetider større end alderen t er mulige. Et system, der har alderen t, må selvfølgelig have levetid større end t.

Hvis vi blot kendte denne *a posteriori* fordeling, kunne vi begynde at komme med forudsigelser om levetiden. Nu er det sådan, at der findes en vigtig sætning i sandsynlighedsregningen, kaldet *Bayes's sætning*, som bringer sammenhæng mellem de betingede sandsynligheder. Her kan den skrives

$$P(t|T)p(T) = Q(T|t)q(t). \tag{5}$$

Sætningen udtrykker simpelthen, at den ubetingede sandsynlighed for, at et system har både alderen t og levetiden T, er den samme, uanset om man beregner den med levetiden eller alderen som betingelse. Med denne sætning kan vi nu beregne *a posteriori* sandsynligheden

$$Q(T|t) = \frac{P(t|T)p(T)}{q(t)} = \frac{P(t|T)p(T)}{\int_t^\infty P(t|T')\,p(T')\,dT'}\ , \tag{6}$$

hvor vi også eksplicit har indsat formlen for aldersfordelingen. Dette er vores afsætspunkt for en rationel diskussion af, hvad der kan siges om fremtiden.

Den flade fordeling

Ofte kan man fastslå den betingede sandsynlighed $P(t|T)$ ud fra rimelige antagelser om, hvorledes alderen t er blevet bestemt. Lad os for eksempel antage, at vi observerer et system som vides at have levetiden T og konstaterer, at det har alderen t. Hvis observationen af systemet ikke på nogen måde er koblet til systemets skæbne eller eksplicit til tiden, så må det være naturligt at anse ethvert observationstidspunkt for at være lige godt. Det betyder, at fordelingen må være flad, givet ved

$$P(t|T) = \frac{1}{T} \qquad (0 \leq t \leq T). \tag{7}$$

Denne fordeling er korrekt normaliseret i det angivne interval. Udenfor intervallet sættes sandsynligheden til nul.

Hvis T betegner socialdemokratiets regeringsperiode, så betyder denne fordeling, at vi kunne have stillet dette spørgsmål på et hvilket som helst tidspunkt, mens det var ved magten. Når T betegner menneskehedens størrelse, betyder den flade fordeling, at vi og vores samtid er tilfældige individer blandt alle de mennesker, der har levet eller kommer til at leve. Med en mere malerisk sprogbrug siges den flade fordeling at udtrykke *det Kopernikanske princip:* Vi og vores tid indtager ikke in særstilling i civilisationens historie, i lighed med at vi siden Kopernikus er gået bort fra at betragte Jorden som universet centrum.

Med den flade fordeling, kan vi straks beregne posterior fordelingen

$$Q(T|t) = \frac{1}{q(t)}\frac{p(T)}{T} \qquad (0 \leq t \leq T), \tag{8}$$

hvor

$$q(t) = \int_t^\infty \frac{p(T)}{T}\, dT \tag{9}$$

blot sikrer normaliseringen. Det følger straks, at integralet i $q(t)$ må være konvergent, fordi $p(T)$ er normaliseret.

Af samme grund kan vi altid beregne middellevetiden

$$\langle T \rangle_t \equiv \int_t^\infty T Q(T|t)\, dT = \frac{1}{q(t)} \int_t^\infty p(T)\, dT \tag{10}$$

under betingelse af, at systemet har opnået alderen t. Denne størrelse er altid konvergent, fordi prior fordelingen skal være normaliseret. Derimod er det på ingen måde givet, at det næste moment af fordelingen eksisterer

$$\langle T^2 \rangle_t \equiv \int_t^\infty T^2 Q(T|t)\, dT = \frac{1}{q(t)} \int_t^\infty T p(T)\, dT. \tag{11}$$

Hvis denne størrelse divergerer, så divergerer spredningen

$$\sigma_t^2 = \langle T^2 \rangle_t - \langle T \rangle_t^2 \tag{12}$$

også. Størrelsen σ_t er også det, der menes med usikkerheden på levetiden T givet t. Selv om der altid findes en middellevetid for en mængde systemer udtrukket af fordelingen $p(T)$, så kan spredningen omkring denne middelværdi altså godt være uendelig stor.

Uanset om spredningen divergerer, så kan man altid beregne medianlevetiden som den levetid $T_{\frac{1}{2}}$, hvor den integrerede posterior sandsynlighed er præcis 50%. Den må derfor opfylde ligningen

$$\int_t^{T_{\frac{1}{2}}} Q(T|t)\, dT = \frac{1}{q(t)} \int_t^{T_{\frac{1}{2}}} \frac{p(T)}{T}\, dT = \frac{1}{2}\,. \tag{13}$$

Denne ligning siger, at systemet med 50% sandsynlighed vil opnå en samlet levetid, der er mindre end $T_{\frac{1}{2}}$. Heraf følger selvfølgelig, at systemet også med 50% sandsynlighed vil opnå en levetid, der er større end $T_{\frac{1}{2}}$.

Gott's fejltagelse

Problemet med dommedagsprofeterne er, at de tror, at de kan kigge Vorherre i kortene og vide noget om *a priori* fordelingen af mulige levetider for unikke systemer. Vi skal nu explicit vise, at Gott gør en antagelse, der låser ham fast til en bestemt prior.

Gotts fejltagelse [4, 7] bunder i den vage formulering af det Kopernikanske princip (den flade fordeling), som siger, at 95% af alle systemer med levetiden T vil have en alder t, der opfylder uligheden

$$\frac{1}{40}T < t < \frac{39}{40}T \ . \tag{14}$$

Det er helt trivielt, fordi fordelingen er flad og der mangler $2\frac{1}{2}\%$ af det fulde aldersinterval ($0 < t < T$) i hver ende og derfor i alt 5% af levetiden. Denne ulighed, som egentlig består af to separate uligheder, kan også skrives på formen

$$\frac{40}{39}t < T < 40t \ . \tag{15}$$

Hvis levetiden T er givet, vil 95% af alle systemer naturligvis også have en alder t, der opfylder denne ulighed. Trækkes levetiden t fra denne ulighed får vi

$$\frac{1}{39}t < T - t < 39t \ , \tag{16}$$

som angår den ekstra levetid, systemet får ud over den opnåede alder. Som før, givet levetiden, vil 95% af systemernes alder også opfylde denne ulighed.

Som sagt gælder ulighederne med 95% sandsynlighed, når levetiden T er kendt og alderen t ukendt. Det er meget fristende at tro, at ulighederne også gælder med 95% sandsynlighed i den omvendte situation, hvor alderen t er kendt, medens levetiden T er ukendt. Det er en fælde, man skal vare sig for, men Gott falder lige i den [4, 7].

Vi skal nu mere matematisk se, hvordan denne forveksling fastlåser prior fordelingen $p(T)$ til at være af formen

$$p(T) = \frac{\text{konst}}{T} \ . \tag{17}$$

Hermed knæsættes en ganske bestemt opfattelse af, hvordan Vorherre vælger levetider for systemerne. En opfattelse, der kan være rigtig eller forkert, men for hvilken vi ikke har nogen evidens.

Intervallet svarende til 95% sandsynlighed har naturligvis ingen særstatus. Helt almindeligt gælder det for den flade fordeling, at givet T er sandsynligheden $b - a = \int_a^b P(t|T)\,dt$ for at

$$aT < t < bT \ , \tag{18}$$

hvor $0 \leq a < b \leq 1$ er et vilkårligt interval. De 95% opnås ved at sætte $a = 1/40$ og $b = 1 - a = 39/40$. Heraf følger det ligesom før, at givet T vil sandsynligheden også være $b - a$ for at t ligger i intervallet

$$\frac{1}{b} t < T < \frac{1}{a} t \ . \tag{19}$$

Begge disse uligheder er som sagt betinget af, at vi kender levetiden.

Hvis vi i stedet ønsker at beregne sandsynligheden for at levetiden opfylder ulighederne, når alderen er givet, så skal vi benytte posterior sandsynligheden $Q(T|t)$ givet ved (9) og den er alt andet end flad. Gott's fejltagelse kan nu i almindelighed formuleres som en påstand om, at posterior sandsynligheden for, at levetiden ligger i intervallet (19) også er $b - a$, eller med andre ord

$$\int_{t/b}^{t/a} Q(T|t) \, dT = b - a \tag{20}$$

for ethvert interval $0 \leq a < b \leq 1$. Konsekvensen af denne antagelse findes ved at differentiere begge sider af denne ligning efter, for eksempel b, hvorved vi får

$$Q\left(\frac{t}{b}\middle| t\right) \frac{t}{b^2} = 1 \ , \tag{21}$$

som er gyldig for alle b. Sættes $b = t/T$ bliver dette til

$$Q(T|t) = \frac{t}{T^2} \ . \tag{22}$$

Indsættes dette i ligning (8) udleder vi at $p(T) \sim 1/T$, hvilket beviser det, vi hævdede. Det følger desuden af Bayes's sætning, at vi også har $q(t) \sim 1/t$.

Ved at hævde en i almindelighed fejlagtig påstand, bliver Gott tvunget til at foretage et unikt valg af prior, $p(T) \sim 1/T$, for hvilken hans påstand faktisk er gyldig. For at korrigere en logisk fejl i sit argument er Gott nødt til at give afkald på den frihed, der ligger i valget af prior!

Scenarier for dommedag

De foregående betragtninger viser tydeligt, at Gott's argumentation er noget vrøvl, og at forudsigelserne altid er dybt afhængige af den prior fordeling for levetider, der antages. Vi skal nu gennemgå nogle forskellige scenarier for at underbygge dette i detalje.

Den 'vage' prior: Da Gott i 1993 bliver presset i debatten [5], begynder han for første gang at bruge argumenter af bayesiansk karakter. Han hævder imidlertid for at redde sit skind, at når man ikke har nogen data om fremtiden, bør man vælge en “tilstrækkelig vag” prior fordeling, nemlig (17). Som vi har set, vil netop denne fordeling underbygge hans påstand, men det er også den eneste der gør det. Den vage fordeling har tilsyneladende ingen indbygget tidsskala, og det følger straks, at posterior fordelingen er givet ved (22). Læg mærke at dette resultat ikke afhænger af den ubekendte konstant, der indgår i denne prior.

Selv om posterior fordelingen (22) svarende til den vage prior er korrekt normaliseret, så divergerer middellevetiden og spredningen på levetiden. Den korrekte konklusion er her, at sådanne uendelige resultater viser vores totale mangel på evne til at forudsige middellevetiden på basis af de data, vi kender. I stedet beregner Gott medianlevetiden og finder

$$T_{\frac{1}{2}} = 2t \ . \tag{23}$$

Der er altså 50% chance for at levetiden vil være mindre (eller større) end den dobbelte alder for dette valg af prior. Sagt på en anden måde kan levetiden lige så vel være mindre som større end medianværdien.

Hvis man tror på Gott (og for et unikt system som menneskeheden kan der kun være tale om tro), så vil menneskeheden med 50% sandsynlighed forgå inden den bliver dobbelt så stor. Dommedag er nær. Det følger også af det vage scenarie, at der med 95% sandsynlighed vil blive mellem 180 millioner og 273 milliarder flere individer end der er nu.

Som sagt divergerer både middellevetiden og spredningen i dette tilfælde. Det strider mod det generelle resultat, som vi har udledt, at middelværdien altid vil være endelig. Problemet med dette scenarie er, at prior fordelingen ikke er normaliserbar, som den burde være. For at normalisere den er det nødvendigt at skære den af både for små og store levetider. Man kunne for eksempel erstatte den med

$$p(T) = \frac{1}{\log \frac{B}{A}} \, \frac{1}{T}, \qquad A \leq T \leq B \ . \tag{24}$$

Denne fordeling er korrekt normaliseret og identisk med den vage i hele intervallet $A \leq T \leq B$. Den ubestemte konstant er desuden blevet bestemt.

Fordelingen er til gengæld ikke nær så ‘vag’ længere. Den afhænger nu eksplicit af intervallets endepunkter, som afstikker de grænser, der på forhånd er sat for levetider. Nu kan man både beregne middellevetid og spredning, som kommer til at afhænge af intervallet. I sidste ende kommer der altså præcis det ud af dette scenarie, som bliver puttet ind. Man kan ikke få noget for ingenting!

Meteornedslag: Et andet scenarie bygger på en forhåndsantagelse om, at menneskeheden kommer af dage ved en katastrofe, for eksempel et meteornedslag som det der skete for 65 millioner år siden, og som uden tvivl var medvirkende til dommedag for dinosaurerne. Hvis vi forestiller os, at der altid er den samme sandsynlighed λ per tidsenhed for, at et meteor slår ned, så vil fordelingen af levetider for "menneskeheder", der lige er blevet skabt, være af samme form som for radioaktive atomer, der henfalder. Vi har derfor

$$Q(T|0) = \lambda e^{-\lambda T} \ , \tag{25}$$

hvor henfaldskonstanten λ er sandsynligheden per tidsenhed for et nedslag (for eksempel 1 per 25 millioner år). Ved hjælp af ligning (8) finder vi så prior fordelingen

$$p(T) = \lambda^2 T e^{-\lambda T} \tag{26}$$

Faktoren T skyldes, at når vi udvælger et atom eller en menneskehed tilfældigt blandt alle de, som eksisterer på et givet tidspunkt, vil de længstlevende have størst chance for at blive udvalgt. Igen hænger alle forudsigelser på antagelsen om, at netop dette scenarie afgør vores skæbne, og at vi har et bud på størrelsen af henfaldskonstanten.

Genetisk sammenbrud: Et scenarie, der indrømmet er noget extremt, går ud på, at en art som menneskeheden uddør, fordi den opsamler så mange genetiske fejl, at individerne ikke længere er levedygtige. Efter et vist antal, for eksempel T_0, reproduktioner er det bare slut, så at vi har

$$p(T) = \delta(T - T_0) \ . \tag{27}$$

Her kan alting også beregnes, men igen afhænger det af, at vi har et godt gæt for T_0.

Indsamling af data

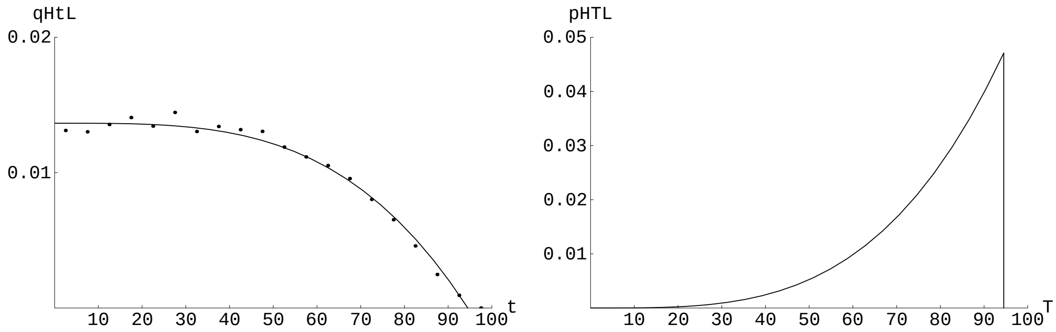
Uden viden om *a priori* fordelingen kommer vi ikke videre. Vi kan ikke afgøre om det ene eller det andet scenario er det rigtige og har derfor ingen forudsigelser om fremtiden.

Hvad skal der til for at opnå data om $p(T)$? Svaret er, at vi må gå ud og kontakte andre verdener (a la Startrek) og spørge dem, om hvor gamle, de er, eller om hvor mange individer, de er lige nu. Sandsynligheden er $q(t)$ for at svaret er t. Med tilstrækkelig meget data kan vi danne et histogram over $q(t)$ og dermed indirekte opnå viden om prior fordelingen $p(T)$.

Hvis den betingede fordeling $P(t|T)$ er flad, kan vi beregne prior fordelingen ved simpel differentiation af posterior fordelingen (9) og finder

$$\left.\frac{dq(t)}{dt}\right|_{t=T} = -\frac{p(T)}{T} \ . \tag{28}$$

Bemærk, at højre side altid er negativ, så at $q(t)$ altid er aftagende. Det er en direkte konsekvens af antagelsen om en flad betinget foredeling og kan kun forventes opfyldt for denne fordeling.



Figur 1: *Aldersfordelingen (venstre) og levetidsfordelingen (højre) for Danmarks befolkning. Data er taget fra Danmarks statistik på* `http://www.dst.dk` *og korrigeret som beskrevet i teksten.*

I Fig. 1 er denne metode benyttet til at beregne prior fordelingen af levealdre ud fra befolkningens nuværende aldersfordeling. Antagelse om en flad betinget fordeling betyder i dette tilfælde, at befolkningen er i en stationær tilstand, hvilket det varierende fødselstal viser, ikke kan være rigtigt. For at korrigere for dette, har vi renormaliseret befolkningens aldersfordeling med det faktiske fødselstal. Korrektionen kan sagtens undgås på bekostning af formalismens enkelthed.

Den derved fremkomne fordeling er vist i venstre delfigur. Som den fuldt optrukne kurve viser, får man et forbavsende fint fit med funktionen

$$q(t) = \frac{\gamma+1}{\gamma T_0}\left(1 - \left(\frac{t}{T_0}\right)^{\gamma}\right) \qquad (0 < t < T_0) \ , \tag{29}$$

hvor $T_0 = 94.5$ år og $\gamma \approx 3.45$. Den deraf følgende prior fordeling er vist i højre side af figuren. Den er

$$p(T) = \frac{\gamma+1}{T_0}\left(\frac{T}{T_0}\right)^{\gamma} \qquad (0 < T < T_0) \ . \tag{30}$$

og vokser monotont op til $T_0 \approx 95$ år for derefter at falde dramatisk mod nul. Middellevetiden bliver

$$\langle T \rangle = \int_0^{T_0} T p(T)\, dT = \frac{\gamma + 1}{\gamma + 2} T_0 = 77.2 \text{ år}, \tag{31}$$

medens den forventede levetid for et menneske, der har opnået alderen t bliver

$$\langle T \rangle_t = \int_t^{T_0} T Q(T|t)\, dT = \frac{\gamma}{\gamma + 1} \frac{T_0^{\gamma+1} - t^{\gamma+1}}{T_0^{\gamma} - t^{\gamma}} \tag{32}$$

For et menneske der lige er født bliver den forventede levetid 73.3 år, medens et menneske, der har opnået denne alder, kan forvente at blive 85 år gammel.

Dette eksempel viser, at det er relativt enkelt at udlede levetidsfordelingen ud fra aldersfordelingen, forudsat der er tilstrækkeligt med data. De data, der er brugt i Fig. 1, består af mere end 5 millioner danskeres nuværende alder. Det er indlysende, at kendskab til en enkelt tilfældigt udvalgt danskers nuværende alder ikke kan føre til nogen som helst brugbar konklusion om levetidsfordelingen $p(T)$.

Konklusion

Vi har gennemgået nogle generelle teknikker fra sandsynlighedsregningen, som er brugbare til at vurdere dommedagsprofetiernes betydning. De har vist os, at enkeltstående begivenheder afslører meget lidt om de underliggende sandsynlighedsfordelinger. Påstande om entydige forudsigelser er altid baseret på skjulte antagelser vedrørende *a priori* fordelinger. Vi har påpeget, at den eneste rationelle metode til at løse denne og lignende problemstillinger er at indsamle mere data. Når sandsynlighedsberegning leder til konflikt med den sunde fornuft, er det en god ide at konstruere et formelt argument, som kan afsløre mulige skjulte antagelser og logiske fejlslutninger. De teknikker, vi har diskuteret her, danner en passende ramme for denne analyse.

Specielt har vi vist, at Gott's dommedagsprofeti er resultatet af et præcist, men ganske uberettiget, valg af prior. Et valg, han er tvunget til på grund af en logisk fejl i sit argument. Selv om man kan *vælge* at tro, at dommedag er nær, så er der ingenting, der tvinger en til det. Dommedag er ikke en uundgåelig konsekvens af sandsynlighedsregningen.

Litteratur

[1] B. Carter
Phil. Trans. Roy. Soc. Lond. **A310**,347–363 (1983)

[2] J. Leslie
Bull. Canad. Nucl. Soc. **10**, 10–15 (1989)

[3] H. B. Nielsen
Acta Physica Polonica **B20**, 427–468 (1989)

[4] J. Richard Gott III
Implications of the Copernican principle for our future prospects
Nature **363**, 315–319 (1993)

[5] Steven N. Goodman, Alan L. Mackay, P. Buch, J. Richard Gott III
Future prospects discussed
Nature **368**, 106–108 (1994)

[6] T. Kopf, P. Krtous, and D. N. Page
Too soon for doom gloom?
University of Alberta preprint (1994)

[7] J. Richard Gott III
A grim reckoning
New Scientist **156**, 36–39 (1997)

[8] Dan H. Andersen
Dommedag er nær — sandsynligvis
Berlingske Tidende, 24. februar (1998)

[9] J. Bang
Et dyt i basunen
Berlingske Tidende, 3. marts (1998)