

# OM ANVENDT VIDENSKAB

Når man arbejder videnskabeligt med problemstillinger, som har stor praktisk betydning i samfundet, må man som regel bruge andre statistiske metoder, end hvis man arbejder med grundforskning. Derfor bør den anvendte videnskabs metoder være på skemaet i de naturvidenskabelige uddannelser.

## Forfatter



Christian Frølund Damgaard er professor, dr. scient. Han arbejder med plantepopulationsbiologi og den matematiske og statistiske modellering af planteøkologiske processer  
Institut for Bioscience,  
Aarhus Universitet  
cfd@bios.au.dk

**H**ypotesetestning er en central aktivitet i den videnskabelige metode, og videnskabeligt arbejde er ofte, efter filosofen Karl Popper, karakteriseret som at gøre sit yderste for at falsificere konkrete arbejdshypoteser om fx kausale sammenhænge i naturen. Men det at arbejde videnskabeligt med en problemstilling omfatter mere end blot at teste hypoteser.

Et eksempel på en videnskabelig hypotese kunne være, at planter konkurrerer med hinanden, når de gror tæt sammen. Hypotesen kan højst sandsynlig føres tilbage til bondestenalderen, og det har åbenlyst ikke være muligt at falsificere hypotesen generelt, selv om tusinder af planteøkologer har undersøgt plantekonkurrence siden da. Men hvad er det så egentligt man laver, når man arbejder videnskabeligt med plantekonkurrence?

Fænomenet plantekonkurrence har stor praktisk betydning for mennesket, fx indenfor landbrug, hvor det er vigtigt at kende den optimale så-tæthed for at få det bedste udbytte. Det har betydet, at man i stedet for at teste grundlæggende videnskabelige hypoteser har bedrevet anvendt videnskab og lagt vægt

på dels at kvantificere effekten af plantekonkurrence samt undersøge miljøets betydning, dels at forstå og kvantificere usikkerheder i vores nuværende viden.

## Anvendt forskning

De problemstillinger, som belyses i den anvendte forskning, er typisk meget relevante for nutidens mennesker: Hvad er fx den mest hensigtsmæssige fordelingspolitik i følge den gældende økonomiske teori? Eller hvad er sandsynligheden for, at en given meteor udsletter alt liv på jorden? Det er derfor underligt, at der ikke er mere fokus og forståelse for den anvendte forsknings metoder. Måske skyldes det, at den anvendte forskning ofte er mere matematisk sofistikeret end den hypotesetestende grundforskning og således er sværere at kommunikere til ikke-fagpersoner.

I grundforskning genereres uspecifik generel viden eller mere præcist videnskabelige arbejdshypoteser, som man ikke hidtil har kunnet falsificere. I anvendt videnskab målretter man denne generelle viden mod en konkret problemstilling ved hjælp af indsamlede data, som man mener på en afgørende måde kan belyse denne problemstilling. I dette arbejde med at målrette og

anvendeligøre vores generelle viden har såkaldt Bayesiansk statistik vist sig at spille en nøglerolle.

## Bayesiansk statistik

Ved hjælp af Bayesiansk statistik kan man opdatere uspecifik *prior* viden til mere specifik og anvendelig *posterior* viden ved hjælp af data. Thomas Bayes (1701 - 1761) var en engelsk præst, som først formulerede det matematiske grundlag for metoden ved hjælp af betingede sandsynligheder. Men det var den franske matematiker og astronom Pierre-Simon Laplace (1749 - 1827), som først anvendte metoden i anvendte videnskabelige problemstillinger, fx til at beregne sandsynligheden for, at solen vil stå op i morgen givet en prior viden om, hvor mange morgener solen hidtidig havde stået op. Desværre var metoden i praksis uanvendelig for de fleste konkrete problemstillinger, idet den typisk er meget beregningstung. Dette praktiske problem forsvandt dog med computeren, og det er nu en relativ simpel opgave at foretage endog meget komplicerede Bayesiansk statistiske beregninger med hjælp fra en kraftig computer.

For en Bayesiansk statistiker er en sandsynlighed ikke noget, man kan genfinde i den fysiske verden. Sand-

## Bayesiansk analyse

Hvis to plantearter konkurrerer i et givet miljø, så er der, efter matematikerne Lotka og Volterra, fire forskellige muligheder: De to arter kan koeksistere, art 1 udkonkurrerer art 2, art 2 udkonkurrerer art 1, eller afhængig af startbetingelserne vil enten art 1 eller art 2 udkonkurrere den anden art. For at undersøge, hvilken af de fire delhypoteser som bedst forklarer, hvad der sker, når de to arter konkurrerer i naturen, kan man udføre et konkurrenceeksperiment. I en traditionel frekventist analyse af et sådant eksperiment vil man enten acceptere eller forkaste hver af de fire delhypoteser, og det er ikke

muligt at udnytte informationen om, at de fire delhypoteser er komplementære. Hvis man derimod analyserer konkurrenceeksperimentet vha. Bayesiansk statistik, vil resultatet være det intuitivt rigtige; nemlig en tildeling af en sandsynlighed til hver af de fire delhypoteser givet data fra konkurrenceeksperimentet.

Hvis man derudover har supplerende viden om konkurrenceforholdet mellem de to planter, fx har observeret i anden sammenhæng, at den ene art som regel udkonkurrerer den anden art, da er det i en Bayesiansk analyse muligt at

inddrage denne viden i analysen af konkurrenceeksperimentet som prior viden.

Endelig er den Bayesianske tilgang simpelthen mere intuitiv end frekventist statistik i forhold til, hvordan vi uformelt syntetiserer viden ud fra kendsgerninger og observationer, og interessant så svarer den Bayesianske tankegang nøje til den tyske filosof Immanuel Kant's empiriske domme (i Kritik af den rene fornuft, 1781) hvor man ved hjælp af sansning kan transcenderer fra det almene og subjektive (prior viden) til det specifikke og objektive (posterior viden).

synligheder er alene et menneskeskabt begreb, som vi anvender til at beskrive vores usikkerhed om et fysisk fænomen. Således giver det mening for en Bayesiansk statistiker at beregne sandsynligheden for, at en bestemt hypotese kan forklare de indsamlede data givet antagelser omkring de mulige hypoteser (se boks).

Bayesiansk statistik afviger på flere andre punkter fra den, for nuværende, hyppigst anvendte form for statistik (frekventist statistik), som først blev udviklet for lidt mere end 100 år siden ved Francis Galton, Karl Pearson og Ronald Fisher og som havde den store praktiske fordel, at man kunne regne med den i hånden. Opbygning af ny viden ved hjælp af frekventist statistik sker ved, at man enten forkaster eller accepterer en relativ simpel nulhypotese baseret på en beregning af sandsynligheden for, at data kan antages at være i overensstemmelse med nulhypotesen. Hvis nulhypotesen forkastes, er det rimeligt at opstille en mere detaljeret og interessant hypotese. Derudover anser en frekventist statistiker anvendelsen af prior viden som en uacceptabel, subjektiv måde at manipulere med resultatet af et eksperiment.



Den franske matematiker og astronom Pierre-Simon Laplace er af den anvendte videnskabs helte. Han var den første til at bruge Bayesiansk statistik på anvendte videnskabelige problemstillinger. Maleri af Sophie Feytaud, 1842

### Bayesiansk statistik bør på skemaet

De danske naturvidenskabelige fakulteter fremhæver ofte den frie hypotesetestende grundforskning som værende essentiel for vores samfunds udvikling og undervisningen af de studerende. Dette er uden tvivl rigtigt, og dog vil jeg hævde, at vi er nødt til at ændre undervisningen på de naturvidenskabelige fakulteter, så den i højere

grad introducerer de studerende til kvantitativ, anvendt videnskab og dennes metoder. Bl.a. mener jeg, det er nødvendigt at udvide statistikundervisningen på de klassiske naturvidenskabelige fag, så de studerende også introduceres til Bayesiansk statistik. Ellers er jeg bange for at kandidaterne fra de naturvidenskabelige fakulteter i fremtiden ikke vil være relevante inden for rådgivning og produktion, men kun kan få arbejde som undervisere.

Derudover er det vigtigt, at vi, som producerer anvendt viden, konsekvent begynder at estimere og kommunikere kvantitative bud på usikkerhederne af vores resultater til den undrende offentlighed, inklusiv i vores rådgivning til styrelser og det politiske system.

Ellers er jeg bange for, at vores troværdighed på lidt længere sigt vil lide betydelig skade. Kommunikationen af fx vejrudsigter har vist, at det sagtens kan lade sig gøre på en måde, så alle kan være med. Fx kan man forestille sig, at rådgivningen om naturplejen på et givet areal i fremtiden også vil indeholde en beregnet sandsynlighed for, at den anviste naturpleje vil lede til det ønskede resultat inden for en given årrække. ■

**Kilder og videre læsning**  
McGrayne, S. H. (2011) The theory that would not die, Yale University Press.

Hahn, R. (2005) Pierre Simon Laplace, 1749-1827: A Determined Scientist, Harvard University Press.

Janes, E. T. (2003) Probability theory. The logic of science, Cambridge University Press.

Damgaard C. (1998). Plant competition experiments: Testing hypotheses and estimating the probability of coexistence. Ecology 79: 1760-1767.