

forsøgte at finde et nyt ståsted. At videnskaben er en social aktivitet var so-leklart, men konsekvenserne af dette var mindre håndgribeligt.

Man kan udlægge disse forskellige tendenser inden for videnskabs-filosofien som krisetegn, positiv introspektion, nødvendig selvransagelse m.v. Men det, der står tilbage, er, at man begyndte at interessere sig for videnskaben som andet end abstrakte strukturer og nu så den som aktivitet, som praksis, som knyttet til instrumenter, laboratorier, sociale organisatio-ner, repræsentationer, eksperimenter. Det gav et frisk pust ind i de ellers abstrakte filosofiske spekulationer. Og man blev opmærksom på den tætte sammenhæng mellem videnskab og teknologi, noget der ellers havde været næsten fraværende i Kuhns og Poppers værker. Videnskabsfilosofien blev konfronteret med virkeligheden. Samtidig forekom det, at de mange so-cialkonstruktivistiske casestudies ikke nødvendigvis behøvede at blive tolket som ren debunking af videnskaben, men faktisk kunne ses som god empirisk historie- eller samfundsforskning. Man kunne måske endda lære noget af dem om hvordan man organiserede, ledede, styrede, evaluerede og kom-munikerede videnskab. Det var jo den måde, man anvendte casestudies på inden for andre grene af samfundsvidenskaben, og den amerikansk-engelske videnskabssociolog Steve Fuller (f. 1959) forsøgte således fra slutningen af 1980-erne at skabe et forskningsprogram – såkaldt social epistemologi – der skulle være forskningspolitisk involveret og give resultater, der kunne bru-ges inden for ledelse og organisation af forskning.

Statistisk signifikans – et lærestykke

Her til sidst vil vi se på et konkret eksempel på samspil mellem både metodologiske, videnskabsfilosofiske og sociale problematikker. Det drejer sig om statistisk signifikans.

Utrolig megen forskning består i at indsamle data og bagefter underka-ste disse en statistisk analyse, for derved at nå frem til bekræftelse eller for-kastelse af en hypotese. Statistik er på den måde et nøgleområde inden for videnskaben, og samtidig et område med en meget interessant historie. Ge-netikeren og informationsteoretikeren Ronald Fisher (1890-1962) vender tilbage i næste kapitel. Men Fisher var også, måske endda først og fremmest, statistiker. I Mellemløstiden kodificerede han den måde, man skulle lave statistisk analyse af empiriske data. Statistikken havde udviklet sig igennem

slutningen af 1800-tallet, bl.a. gennem udviklingen af sandsynlighedsregningen. Væsentlige navne var Francis Galton (1822-1911) og Karl Pearson (1857-1936), der begge også var “eugenikere”, som vi skal se. De og mange andre arbejdede med at indsamle og systematisere store mængder empiriske data, og de var især interesserede i, hvad man kunne uddrage af disse.

Fisher arbejdede med, hvordan man på forhånd skulle tilrettelægge videnskabelige forsøg, så de gav brugbare data, og med hvordan man ud fra disse kunne drage holdbare konklusioner. Fisher arbejdede især med situationer, hvor man skulle afgøre, om ens data var af en sådan karakter, at det var sandsynligt, at en given hypotese ikke kunne bekræftes af dem, idet de kunne være fremkommet som resultat af rene tilfældigheder. Når man har en hypotese, kan man lave to typer fejl. Enten kan man ud fra sine data slutte, at det forholder sig i overensstemmelse med hypotesen, og senere finde ud af, at det gjorde det ikke. Eller man kan hævde, at det ikke forholder sig i overensstemmelse med hypotesen, og så senere finde ud af, at det gjorde det faktisk. Fisher udarbejdede en række tests, der skulle sikre, at man ikke tillagde en hypotese forkert status. Han mente, at hvis man kunne slutte, at der var en sandsynlighed mindre end 1 til 20 for, at data var et resultat af tilfældigheder, og derfor ikke bekræftede hypotesen, kunne man hævde, at hypotesen var sand. Deraf kommer udtrykket, at det er “statistisk bevist, at ...” Fisher skabte et begreb om statistisk signifikans og om det rette niveau for denne signifikans – typisk fem procent.

Utallige “dobbeltblinde kontrollerede” forsøg, og andre forsøg blot gennemført med kontrol-grupper inden for medicin, landbrugsforskning (der var Fishers eget gebet), psykologi, økonomi osv. har anvendt hans metoder, som næsten fik status som identisk med “den videnskabelige metode”, desuagtet at man inden for f.eks. fysik og astronomi meget sjældent ser denne metode anvendt. Her ræsonnerer man oftere ud fra modeller og årsagssammenhænge, ud fra “nomologiske maskiner”, som videnskabsteoretikeren Nancy Cartwright (f. 1943) har karakteriseret det. Men i ganske mange tidsskrifter inden for psykologi, økonomi og andre samfundsvidenskaber samt i lægevidenskaben er langt hovedparten af artiklerne baseret på vidensudsagn, der hævdes at være viden, fordi de er statistisk signifikante.

Fisher hævdede, at der fandtes en universel metode til at afgøre kvaliteten af de videnspåstande, som man kunne komme med på basis af eksperimentelle data. Det var for så vidt en overbevisning, der passede med

samtidens bestræbelser på at finde grundlaget for en enheds-videnskab og for at se al videnskab som empirisk funderet – altså en art logisk positivisme. Men allerede i forbindelse med fremkomsten af Fishers arbejder i 1920'erne og 30'erne kom der indvendinger imod centrale dele af hans teser. Statistiske metoder kunne ikke være universelle og uafhængige af den substans, det felt, det drejede sig om. Fisher havde som en konsekvens af sine synspunkter hævdet, at der var symmetri mellem de to ovennævnte fejltyper. Men det er klart, at hvis man f.eks. bruger statistiske metoder til at styre kvaliteten i en fødevarerproduktion, er det ikke et brugbart synspunkt. Hvis en kyllinge-producent vil indføre et system med prøver for salmonella og derfor beder om vejledning fra statistikeren, er der selvkært forskel på forebyggelse af de to typer fejl. Det er katastrofalt at lave en test, der siger, at kyllingerne ikke har salmonella, hvis de faktisk har det, mens det alene koster penge for producenten, at lave en test, der siger, at de har salmonella, hvis de rent faktisk ikke har det. Der er altså ikke symmetri, og det påpegede andre, såkaldt "teoretiske", statistikere hurtigt. Statistisk signifikans er ikke et mål for videnspåstandes pålidelighed eller brugbart som handlevejledning. Der må viden om selve substansen og situationen til.

Ikke desto mindre blev Fishers metoder kanoniserede og opnåede videnskabelig status. Fisher selv affærdigede kritikerne med, at de var rene teoretikere, der ikke havde erfaring fra konkret empirisk forsøgsarbejde. Da Kuhn og socialkonstruktivistene begyndte deres angreb på den videnskabelige objektivitet og de retoriske midler, hvormed den videnskabelige institution opretholdt disse forestillinger, kom også begrebet "statistisk signifikans" og de tilhørende tests under analyse og angreb. De blev set som netop retoriske trick, der skulle overbevise, men var uden reelt grundlag. Ordet "signifikans" blev ifølge disse retoriske analyser af forskerne ofte tolket som "betydningsfuld", og dermed sluttede de, at hvis en statistisk test viste signifikans, så var resultatet også dermed betydningsfuldt. Der var dog reelt tale om forvanskning af Fisher, der alene havde sagt, at hans test med en rimelig sandsynlighed kunne sikre, at bestemte data ikke skyldtes tilfældighed. Man kan altså f.eks. ved en stor undersøgelse finde ud af, at et sæt kostvaner giver en middellevelevetid på 74,6 år og et andet sæt en middellevelevetid på 74,7 år. Hvis testen så viser statistisk signifikans, så kan man med stor sandsynlighed hævde, at denne forskel faktisk skyldes forskellen i kostvaner. Men man kan selvfølgelig ikke slutte, at denne forskel er betydningsfuld i

den forstand, at den f.eks. vil få store dele af befolkningen til at skifte kostvaner. Hvis forskellen havde været 74,6 og 79,7 år, ville man sikkert kunne se en anden form for konsekvens. Det var i øvrigt den slags diskussioner, der måtte føres omkring undersøgelserne af sammenhængen mellem rygning og lungekræft i 1950'erne og 60'erne.

Statistisk signifikans er altså ingen garanti for "signifikans". Den erkendelse har ikke desto mindre manglet i et utal af tidsskrifter, hvor denne type tests er blevet brugt uden en diskussion af, hvad de egentlig siger, og hvad man kan slutte ud fra dem. Der har ikke været nogen epistemologisk disciplin, kunne man sige. Fishers tidlige kritikere, f.eks. Egon S. Pearson (1895-1980) og Jerzy Neyman (1894-1981), havde også en videnskabelig kontrovers med ham. Han "vandt" i den forstand, at hans anbefalinger og metoder fik en fortolkning, der blev utrolig udbredt (faktisk findes der ganske mange software-pakker, der giver adgang til at lave statistiske signifikans-tests à la Fisher – uden at de helt gør klart, hvad det er, der sker, og hvad betydningen af det er). Og kigger man på videnskabelig aktivitet generelt, vil man se, at brugen af statistiske signifikans-tests er endog meget stor.

"Det er statistisk bevist, at ..." Argumentet er et redskab i den sociale konstruktion af videnskabelige kendsgerninger, ingen tvivl om det. Men en væsentlig del af deres overtalelseskraft skyldes en åbenlys uklarhed omkring, hvad der ligger i begrebet signifikans. En lang række tidsskrifter og videnskabelige discipliner er i de senere år begyndt at kræve andre former for statistisk behandling af empiriske data, nogle nægter sågar at optage artikler baseret på fisherske statistiske analyser for signifikans – det gælder bl.a. inden for medicin, især epidemiologi, og dele af psykologi, økologi og pædagogik.