

Hvorfor bruke bayesiansk tankegang?

Selv et (tilsynelatende) tørt fag som statistikk kan ha store disputer som engasjerer fagfolk dypt og fører til danning av markerte skoleretninger. Diskusjonen mellom tilhengere av bayesiansk tankegang og statistiske metoder og mer tradisjonelle statistikere bør være av interesse også for leger. Aursnes og medarbeidere (1) gir i dette nummer av Tidsskriftet et bilde av denne diskusjonen sett fra et bayesiansk ståsted. I artikkelen settes statistiske problemstillinger på dagsordenen, og det er ingen tvil om at leger bør ha kunnskap om bayesianske metoder.

Aursnes og medarbeidere kritiserer med rette den overdrevne bruken av p-verdier i medisinsk statistikk. Men de kommer bare i liten grad inn på konfidensintervaller, som er langt mer informative og som derfor bør foretrekkes (2). De legger stor vekt på Bayes' formel, men eksemplet som benyttes (1), inneholder ingen spesielle bayesianske elementer. Lesere som er interessert i bayesiansk statistikk, anbefales en artikkel av David Spiegelhalter, som er en fremtredende bayesianer, for å få en god oversikt over disse metodenes plass i medisinen (3).

Hva dreier denne diskusjonen seg egentlig om? Motsetningene bunner dels i forståelsen av begrepet «sannsynlighet». Man kan gi minst to forskjellige definisjoner. I den ene definisjonen kan man for eksempel tenke på terningkast, og si at sannsynligheten for å få to øyne er lik 1/6, siden frekvensen av toere vil gå mot 1/6 når antall kast øker. Dette er den *frekventistiske* definisjonen. Den bayesianske ser derimot sannsynlighet som *subjektiv*, uttrykk for ens tro på et forhold. En spesialist kan f.eks. anslå en 60 % sannsynlighet for at en bestemt ny medisin vil virke godt før han har resultater fra kliniske utprøvinger. I praksis vil bruken av sannsynlighet ofte være en blanding av det frekventistiske (f.eks. data fra epidemiologiske studier) og det subjektive.

Så langt er det hele ukontroversielt. Motsetningene oppstår når data fra vitenskapelige forsøk skal analyseres statistisk. Bayesianere vil her insistere på at vekten skal ligge på subjektiv sannsynlighet. Man kan for eksempel tenke på en klinisk utprøving av et nytt medikament. Forut for dette foreligger det en viss tro på medikamentets nytte sammenliknet med en tradisjonell behandling. I bayesiansk tankegang opererer man med begrepet apriorisannsynlighet for medikamentets effekt, basert på oppfatninger blant fagfolk. Denne apriorisannsynligheten kombineres så med data fra det kliniske for-

søket og resulterer i en aposteriorisannsynlighet. Dette gir beregninger som erstatter de klassiske p-verdier og konfidensintervaller. I praksis ser man likevel ofte at konklusjonene av en bayesiansk og en klassisk (frekventistisk) analyse er ganske sammenfallende.

I klassisk statistikk, altså den retning som dominerer i medisinsk-vitenskapelige tidsskrifter og lærebøker og som baserer seg på et frekventistisk sannsynlighetsbegrep, vil man ønske å skille apriorioppfatninger fra analysene av dataene. Dette er parallelt med skillet mellom resultat og diskusjon i en vitenskapelig artikkel. Først gir man resultatene fra studien, og så diskuterer man relasjonen til faglige oppfatninger, biologiske teorier, tidligere kjente resultater etc. Resultatdelen må derfor nødvendigvis inneholde temmelig standard frekventistiske analyser. Dette er utvilsomt noe av årsaken til at man knapt finner subjektive bayesianske analyser i medisinsk forskningslitteratur.

Det er også innvendinger av mer prinsipiell art. Blant annet kan kvaliteten av faglige oppfatninger variere sterkt. Det er eksempler på at feilaktige oppfatninger har stått meget sterkt også blant spesialister. Et eksempel er oppfatningen i 1980-årene om at spedbarn burde legges på magen, noe som senere viste seg å øke risikoen for krybbedød.

Sett utenfra kan man sikkert undre seg over at uenigheten mellom bayesianere og klassiske statistikere kan føre til intense debatter. Det dreier seg i virkeligheten om en forskjellig filosofisk vinkling på statistiske analyser. Sterke personligheters engasjement har ført til kontroverser over mange tiår.

Tradisjonelt har subjektiv sannsynlighet stått sterkt i det man har tenkt på som bayesianske metoder. I de senere år har imidlertid en objektiv bayesianisme vokst frem og er i dag dominerende. Denne benytter de samme matematiske modeller som bayesianerne tradisjonelt har brukt. Apriorisannsynlighetene er i liten grad uttrykk for subjektive oppfatninger, de er snarere en del av et teknisk apparat der man gjerne opererer med såkalte ikke-informative apriorifordelinger. Denne utviklingen har skutt fart med nye beregningsmetodikker som er tatt i bruk i de senere år, «Markov chain Monte Carlo»-metodene. Disse teknikkene er stort sett helt uproblematisk også for de klassiske statistikere, og flere av dem har tatt dem i bruk.

Utviklingen i de senere år har derfor vært at gapet mellom klassiske statistikere og bayesianere er blitt mindre. De store emo-

sjonelle debatter føres i langt mindre grad enn før. De fleste statistikere er pragmatisk innstilt, og man bruker det som til enhver tid er et nyttig hjelpemiddel til å løse problemene. Bayesianske metoder var hovedtema i en doktoravhandling ved Universitetet i Oslo i 2001 (4, 5).

Bayesianske teknikker kan være nyttige innen mange fag. I medisinen er det en forholdsvis beskjeden bruk av bayesianske metoder. I Medline fra 1996 til april 2002 finner man 1 359 referanser, og søker man etter «Markov chain Monte Carlo», finner man 196 referanser. Likevel – lærebøker i medisinsk statistikk og medisinsk-vitenskapelige tidsskrifter er utvilsomt dominert av de klassiske frekventistiske metoder. En viktig grunn til dette er at disse metodene er lett tilgjengelige i statistiske programpakker. Man kan uten særlige problemer kjøre de mest kompliserte analyser, mens bayesianerne i liten grad har nedfelt sine metoder på denne måten. Det finnes likevel noen slike programpakker, bl.a. BUGS, som er utviklet i et medisinsk-statistisk miljø i Cambridge (6).

I årene fremover vil man sannsynligvis se økt bruk av de objektive bayesianske metoder. Også de subjektive metoder vil bli brukt i sammenhenger der dette er naturlig. Lærebøker og kurs i medisinsk statistikk vil sikkert også bli preget av dette.

Odd O. Aalen

Seksjon for medisinsk statistikk
Universitetet i Oslo
Postboks 1122 Blindern
0317 Oslo

Odd O. Aalen (f. 1947) er professor i medisinsk statistikk ved Universitetet i Oslo

Litteratur

1. Aursnes I, Natvig B, Tvette IF. Bayesiansk tenkemåte på vei inn i medisinsk statistikk? Tidsskr Nor Lægeforen 2002; 122: 1369–72.
2. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet 2002; 359: 57–61.
3. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. Health Technology Assessment 2000; 4: 1–130.
4. Osnes K. Different choices of priors in Bayesian disease mapping: a case study. J Epidemiol Biostat 1998; 3: 277–87.
5. Osnes K, Aalen OO. Spatial smoothing of cancer survival: a Bayesian approach. Statistics in Medicine 1999; 18: 2087–99.
6. BUGS. www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml.