



TILFELDIG GANG

Nr. 3, årgang 6

September 1989

Utgitt av Norsk Statistisk Forening

Innhold

	Side
Hva slags stoff bør TG inneholde?	2
Fra formannen i Norsk Statistisk Forening	4
American Statistical Association 150 år. <i>Jostein Lillestøl</i>	5
På tide med Bayes-analyse i Norge nå? <i>Bent Natvig</i>	7
Vi har lest S.M. Stigler: The history of statistics. <i>Henrik Dahl</i>	14
Klipp fra debatten utenfor statistikernes rekker	15
Nytt fra universiteter og høyskoler	16
Kalender for internasjonale møter	17
Call for Papers – Probabilistic Safety Assessment and Management (PSAM), Beverly Hills 1991	18



Hva slags stoff bør TG inneholde?

Mange faglige foreninger støtter utgivelsen av vitenskapelige tidsskrifter. Spesielt i de senere år har det imidlertid også vært en oppblomstring av mer uformelle publikasjoner som bringer generelle nyheter ut til medlemmene. På denne måten står f.eks. American Statistical Association bak Amstat News, og The Biometric Society gir ut Biometric Bulletin. Slike tidsskrifter synes å dekke et stort praktisk behov hos de vanlige foreningsmedlemmene, både som rene nyhetsformidlere og i sammenheng med mer uformelle diskusjoner medlemmene imellom.

Det er likevel ikke almen enighet om nytten av denne typen publikasjoner. I mainnummeret i år av Biometric Bulletin finner vi f.eks. et innlegg av den kjente britiske statistikeren David J. Finney, som i sterkt kritiske vendinger tar opp utgivelsen av dette bladet:

"I know that you and your predecessors have put an immense amount of hard work into the Biometric Bulletin. I was always doubtful about this venture, and I am still unconvinced that the publication has a value commensurate with the trouble and cost. At least let us be careful to keep it as an ephemeral and discourage inclusion of articles of lasting value that ought to find their place in a normal journal."

Nå er det et åpent spørsmål om TG kan sammenligne seg med Biometric Bulletin. Norsk Statistisk Forening har ikke de samme ressurser som Biometric Society, og det er en åpenbar forskjell i ytre standard mellom de to publikasjonene. Det kan likevel være interessant å forfølge Finneys tankegang overført til våre lokale forhold. Slutter man opp om det synspunktet Finney forfekter, bør redaktøren av TG sky alle artikler som på noen måte antyder en overgang til arbeider publisert i "seriøse" tidsskrifter.

Allerede i det første nummeret under den nye redaksjonen i TG blir imidlertid dette prinsippet åpenlyst brutt. Et av temaene som nylig ble tatt opp på det 5. Norske Statistikermøte på Inderøy, var Bayesiansk statistikk. Redaktøren har bedt en av hovedaktørene i denne forbindelsen om å utarbeide et skriftlig innlegg, tilpasset rammen i TG. Resultatet er Bent Natvigs artikkel i dette nummeret. En vurdering av den etterfølgende debatten på Inderøy gir ikke grunnlag for å tro at akkurat dette emnet skal utløse de helt store kontroverser i dag. Likevel representerer Natvigs artikkel i sin form noe nytt for TG.

Redaktøren har den største respekt for David J. Finney (og hans arbeider – enten de måtte dreie seg om probitanalyse eller populasjoner av kattehaler!). Likevel er det den nye redaksjonens hensikt å prøve å bryte med tankegangen skissert i sitatet ovenfor. En publikasjon som TG vil nettopp være det riktige sted for en uformell meningsbrytning, der en ikke skal behøve å bekymre seg om noe er av "lasting value" eller ikke. Det ligger ingen negativ vurdering av mer "vitenskapelige" tidsskrifter i dette synet, men bare et ønske om å tilby rom for

debatt tilpasset forholdene for den relativt lille gruppen av norske statistikere. I tiden fremover vil derfor stoffet i TG bli forsøkt supplert med innlegg som i sterkere grad drar i retning av diskusjon omkring faglige spørsmål. Alle innlegg som også fører diskusjonen videre mottas selvsagt med takk.

I det første nummeret av TG, som utkom i april 1984, ble det bl. a. understreket at bladet ville formidle "sladder av almen interesse og betydning for statistikere." Slike temaer vil naturligvis fortsatt innta en sentral plass i TG's spalter. Det kan diskuteres om den debatten som fra tid til annen utspiller seg i dagspressen om statistiske spørsmål, hører med i denne kategorien. Likevel kunne ikke redaktøren dy seg for å sakse fra en meningsutveksling som nylig foregikk i Bergens Tidendes spalter. Det hele blir presentert her under overskriften "Klipp fra debatten utenfor statistikernes rekke", selv om en av debattantene utgikk fra våre egne tropper. Den tidligere formannen i NSF må tydeligvis ha reagert akkurat på samme måte som TG's redaktør i forbindelse med de mange lettvinte statistiske undersøkelsene presentert som "nyheter" i Bergenspressen. For ordens skyld har vi naturligvis også tatt med et motinnlegg. Hvis det dukker opp flere sladretemaer av almen interesse for statistikere, kan det bli aktuelt å ta med tilsvarende klipp også i fremtidige numre av TG.

FRA FORMANNEN I NORSK STATISTISK FORENING

Hva er det mange elsker å hate? Mange lesere vil svare Melodi Grand Prix. Andre vil ifølge utvalgsundersøkelser utført av vår nyvalgte sekretær, Ørnulf Borgan, svare Avdeling C. Med formannen også valgt fra samme sted burde utgangspunktet kanskje være litt frynsete.

La oss hale frem trumfess med en gang - Erik Mohn. Han var professor II ved Avdeling C fra 1983 frem til sin altfor tidlige død på en skitur i Nordmarka i mars i fjor. Riktignok hadde han sin hovedstilling på Norsk Regnesentral, men for oss ved Avdeling C var han vår mann. Under hans formannstid fra 1984-85 ble foreningen løftet opp til det nivå den ligger på i dag. Fra å være en Oslo-forening ble den en nasjonal forening med lokale underbruk. Tilfeldig Gang ble etablert for å binde foreningen sammen. Den faglige virksomheten fikk en ny dimensjon ved innføringen av etterutdanningsseminarer i 1985. De sporene Erik Mohn trakk opp har de etterfølgende styrever under Jostein Lillestøl og Bo Lindqvist fulgt etter i på en glimrende måte.

Det nåværende styre utgjør et sterkt lag og har ingen unnskyldning for en eventuelt dårlig innsats. Det har vi da heller ikke tenkt å gjøre. Etter et velbegrunnet forslag fra Eivind Damsleth, Norsk Hydro, tar en sikte på et etterutdanningsseminar i "Dataanalyse og grafiske metoder" i Oslo i februar 1990 med Dr. William S. Cleveland, Bell Lab. som foreleser. Nærmere detaljer om seminaret vil bli utsendt så snart de foreligger.

Styret med den nye redaktør av TG, Ivar Heuch, i spissen ønsker også å videreutvikle meldingsbladet, kanskje særlig når det gjelder det faglige stoffet samt som et forum for meningsutvekslinger. La det være helt klart at "redaktørplakaten" også gjelder for TG, slik at styret ikke vil blande seg bort i den redaksjonelle linje Ivar Heuch vil måtte ønske å følge. Det er viktig at den enkelte er villig til å bidra med stoff enten uoppfordret eller etter press fra redaktøren.

Statistikkfaget står meget sterkt i Norge akkurat nå. Ved flere av våre akademiske institusjoner og også utenfor disse fins det faglige miljøer av god internasjonal klasse. Dette ble bl.a. klart synliggjort i 1988 både ved den Nordiske Konferansen i Matematisk Statistikk i Åbo og ved det Europeiske Statistikermøte i Øst-Berlin. Med økte ressurser har også de akademiske institusjoner kunnet ta opp nye anvendelsesområder og ha et langt tettere samarbeid med eksterne miljøer. Endelig kommer at en er i ferd med å samordne statistikkfagets utvikling med utviklingen i informasjonsteknologi. Fra å være et påheng til de matematiske instituttene, står også nå plutselig statistikkfaget helt sentralt når fremtidens utvikling av matematikkfaget drøftes.

Dette leder til mitt siste poeng. Innslaget av statistikk og kanskje spesielt sannsynlighetsregning må styrkes i den norske skole. Dette må ikke gå på bekostning av den tradisjonelle, grunnleggende matematikkundervisningen, men komme i tillegg. Avgjørende er det at innslaget fra vårt fagområde illustrerer spennvidden i anvendelsene og ikke gror fast i rød og blå kuler og bridgehender. Til våren skal det gis et etterutdanningskurs over en uke for matematikklærere i den videregående skole. Avdeling C har tatt i mot utfordringen å dekke 40% av undervisningen. Den sjansen skal vi utnytte maksimalt. Vi sier ikke mer.

Oslo 25. august 1989

Bent Natvig

American Statistical Association – 150 år

American Statistical Association (ASA) ble stiftet i Boston i 1839, og har nylig feiret sitt 150-årsjubileum. Initiativet til foreningen ble tatt av fem menn med interesse for å samle inn og presentere fakta om det amerikanske samfunn: en prest, en lege, en redaktør, en kongressmann og en lokalhistoriker. Det var betydelig interesse for foreningen, og den fikk straks over 50 medlemmer. Disse tilhørte den første generasjon amerikanere med aritmetikk i skolen, og det var i samfunnet en økende tro på nytten av denne.

Den første tid var "arkivarer" og leger de dominerende medlemsgrupper, og bare noen få hadde videre matematisk skolering eller akademisk tilknytning. Mens "arkivarene" var tilbøyelig til å ville registrere alt som kunne tenkes å interessere senere historikere, var andre mer opptatt av å avdekke "laws of human life", bl. a. ved analyser basert på "censuses", som i Boston var utført hvert tiår siden 1800. I en analyse av disse uttrykker Shattuck, trolig den best statistisk skolerte av stifterne, sin målsetting slik:

"These are not the speculations of a visionary theorist, but the legitimate deductions from sober facts. We are not a theorist – an experimentalist. We have no sympathy with the opinions of some modern reformers, who seem to be governed by theories founded on uncertain and partial data, or vague conjecture. We are a statist* – a dealer in facts. We wish to ascertain the laws of human life, developed by the natural constitution of our bodies, as they actually exist under the influence that surround them, and to learn how far they may be modified and improved. This can only be done by an accurate knowledge of the facts that are daily occurring around us."

Foreningens aktiviteter falt raskt inn i de mer sofistikerte statistiske strømninger som var under utvikling i England, Tyskland og Frankrike. I 1888 startet utgivelsen av det tidsskrift vi i dag kjenner som Journal of the American Statistical Association. Selv om foreningen og dens medlemmer hadde betydelig påvirkningsevne i statistiske spørsmål hos de styrende myndigheter, var foreningen ikke enerådende, og neppe å regne for en nasjonal forening før etter århundreskiftet.

I 1910 var ASA omgjort fra et lærd selskap av amatører til en profesjonell organisasjon av statsansatte statistikere og akademikere som arbeidet med statistikk, hovedsakelig i samfunnsfagene, men også i næringslivet. Foreningen hadde i stigende grad konkurranse fra profesjonelle organisasjoner i de ulike anvendelsesområder, i første rekke innen økonomi og sosiologi. I tiden rundt 75-årsjubileet i 1914 strevde foreningen med å finne sin identitet.

I begynnelsen av 1920-årene ble agendaen klarere og utvidet. ASA fanget opp de begrepsmessige og teoretiske gjennombrudd hos den engelske biometriske skole,

* Uttrykket "statistician" ble ikke brukt før på 1900-tallet.

som la grunnlaget for faget statistikk slik vi kjenner det i dag. Samtidig fortsatte man de tidligere oppgavene knyttet til innsamling, lagring og rapportering av statistiske data, og la vekt på at ASA skulle være en "økumenisk" organisasjon med bredest mulig medlemsbase. I 1930 var hovedtyngden av medlemmer fortsatt økonomer i forvaltning og næringsliv, deretter andre samfunnsvitere av ulikt slag, men landbruk, naturvitenskap og ingeniørfag var begynt å gjøre seg gjeldende.

De siste 50 år kan ASA vise til en stadig stigende aktivitet. Arrangering av faglige møter knyttet til forskning og utdanning, samt omfattende publiseringsevne, er de mest synlige uttrykk, men rådgivende virksomhet overfor næringsliv og forvaltning (bl. a. setting av standarder), er også sentrale arbeidsoppgaver. Ved 150-årsjubileet har ASA over 15 000 medlemmer organisert i 11 seksjoner etter faglige interesser og i 75 regionale avdelinger.

Jubileet ble feiret på ASA's årlige hovedmøte 6.-10. august, denne gang lagt til Washington, DC. Undertegnede deltok som representant for Norsk Statistisk Forening. Noen inntrykk: Møtet hadde et enormt omfang med opptil 20 sesjoner samtidig, med en rekke temautstillinger og produktpresentasjoner. Variasjonen i tema var enorm, fra hard matematikk (i samarbeid med IMS) til korttricks i elementærundervisningen, det var sannelig ikke lett å velge. Jeg hadde stort utbytte av møtet (herunder en t-skjorte med påskrift "I am statistically significant"), og takker NSF for reisestøtten. Inntrykk knyttet til mitt verv i NSF's undervisningsutvalg vil jeg forsøke å formidle senere i TG.

Jostein Lillestøl

P.S. Hva har Florence Nightingale, Alexander Graham Bell og Jostein Lillestøl til felles med IBM og Bank of Greece? Medlemsskap i ASA!

PÅ TIDE MED BAYES-ANALYSE I NORGE NÅ?

Bent Natvig, Universitetet i Oslo

Foredrag Norsk Statistikermøte, Inderøy, 12-14 juni 1989
(Presentasjonen her er noe forkortet).

Sammendrag.

Hensikten med foredraget er å presentere hovedidéer i Bayesiansk statistisk metodelære og prøve å avklare forskjeller fra den statistiske metodelære som i dag i hovedsak presenteres ved norske universiteter og høyskoler. Spesielt vil det legges vekt på motsetningene mellom sannsynlighetsberegninger gjort før og etter at data er samlet inn, ved bl.a. å sammenligne det klassiske konfidensintervall med det Bayesianske troverdighetsintervall.

1. Innledning.

Klassisk hypotesetestingsteori er først og fremst egnet til å teste vitenskapelige hypoteser. Den er ikke konstruert for, og dermed heller ikke særlig egnet for, å danne grunnlag for beslutninger under usikkerhet. Ved slike beslutninger i forbindelse med reelle anvendelser vil det som oftest være direkte dumt å ikke ta hensyn til godt skjønn i tillegg til mer konkrete data. Et jurymedlem i en mordsak er f.eks. nødt til å skjønnsmessig vurdere troverdigheten av nøkkelvitner i tillegg til å basere seg på mer håndfaste bevis. En lege som vurderer om en pasient skal gjennomgå en hjerteoperasjon, kan ikke bare basere seg på generelle estimater for at slike operasjoner er vellykket. Legen må også, til dels skjønnsmessig, ta hensyn til pasientens fysiske og psykiske helsetilstand. Et forsikrings-selskap som tilbys å forsikre et kjernekraftverk, må gjøre dette på grunnlag av en risikoanalyse av kjernekraftverket. En slik analyse har ingen muligheter til å bli rimelig bra hvis en ikke tar hensyn til ingeniørenes skjønn og erfaringer med de teknologiske komponenter og psykologenes og sosiologenes erfaringer med de menneskelige komponenter. Den Bayesianske angrepsmåte er konstruert for å ta hensyn til slikt skjønn på en måte som er konsistent med sannsynlighetsregningens regler, og er derfor særlig velegnet i forbindelse med beslutninger under usikkerhet.

2. Á priori og á posteriori sannsynlighetstettheter.

Vi skal se på et eksempel fra akselerert levetidstesting der $n=100$ metallstrenger utsettes for en høy statisk belastning. La

$$X_i = \text{Levetid for } i\text{-te metallstreng, } i = 1, \dots, n$$

Gitt feilintensiteten λ , antas nå $X_1, \dots, X_n$ å være uavhengige, identisk eksponensielt fordelte med feilintensitet λ . Betegn dataene fra forsøket med $D_1 = (x_1, \dots, x_n)$. Rime-lighetsfunksjonen for λ gitt D_1 er da:

$$L(\lambda|D_1) = \lambda^n \exp(-\lambda \sum_{i=1}^n x_i)$$

λ betraktes som *fast*, men ukjent for statistikeren. Vi kaller den en *tilfeldig størrelse*. Den opprinnelige usikkerhet statistikeren har om hvor λ måtte befinne seg, uttrykkes ved en *á priori sannsynlighetstetthet* $\pi_0(\lambda)$. Vi har da en *subjektivistisk Bayes metode*. Hvis λ er en tilfeldig variabel som beskriver "fysisk variasjon" fra metallstreng til metallstreng, har vi en *empirisk Bayes metode*. Det er noe ganske annet.

Etter at dataene D_1 er observert, oppdaterer statistikeren sin usikkerhet om hvor λ måtte befinne seg. Denne uttrykkes ved en *á posteriori sannsynlighetstetthet* $\pi(\lambda|D_1)$. Hemmeligheten er nå at Bayes' teorem er skreddersydd for denne oppdateringen og gir umiddelbart

$$\pi(\lambda|D_1) = \frac{L(\lambda|D_1)\pi_0(\lambda)}{\int_0^\infty L(\lambda|D_1)\pi_0(\lambda)d\lambda} = kL(\lambda|D_1)\pi_0(\lambda), \quad (2.1)$$

der k bare er en normeringskonstant.

Anta nå at vi i en viss forstand er uten informasjon om λ før forsøket settes igang. Vi velger derfor *á priori sannsynlighetstettheten* $\pi_0(\lambda)$ utfra en uniform modell:

$$\pi_0(\lambda) = \begin{cases} M^{-1} & 0 \leq \lambda \leq M \approx 10^{12} \\ 0 & \text{ellers} \end{cases} \quad (2.2)$$

Det er da ikke så vanskelig å vise at

$$\pi(\lambda|D_1) \approx \left(\sum_{i=1}^n x_i \right)^{n+1} \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) / \Gamma(n+1) \quad 0 \leq \lambda < \infty \quad (2.3)$$

Høyre side her er en gamma tetthet med *formparameter* $a = n + 1$ og *skalaparameter* $b = \sum_{i=1}^n x_i$. Denne siste størrelsen er forsøkets totale testtid.

Anta nå at vi utfører et nytt forsøk uavhengig av det første og ender opp med dataene $D_2 = (y_1, \dots, y_m)$. Vi benytter da høyre side av (2.3) som vår *á priori sannsynlighetstetthet* med parametrene a og b som gitt over. Fra (2.1) blir nå *á posteriori tettheten* en gamma tetthet med formparameter $a + m = n + m + 1$ og skalaparameter $\sum_{i=1}^n x_i + \sum_{i=1}^m y_i$. Skalaparameteren kan fortsatt fortolkes som forsøkenes totale testtid. En antagelse om en *á priori gamma tetthet* i vår modell vil altså lede til at også *á posteriori tettheten* er gamma. Gamma fordelingen sies derfor å være en *naturlig konjugert á priori fordeling* for feilintensiteten i den eksponensielle modell. Hvis vi for formparameteren a i en slik fordeling har at $a > 1$, kan $a - 1$ fortolkes som antall observasjoner i et tidligere forsøk og skalaparameteren b som den tilhørende totale testtid. Dette tidligere forsøket kan være reelt eller fiktivt. Denne fortolkning, samt matematiske bekvemlighetshensyn, er grunnen til at en ofte velger en *naturlig konjugert á priori fordeling* i den eksponensielle modell.

3. Litt Bayesiansk inferensteori.

Det er viktig å legge merke til at utgangspunktet for en Bayesiansk inferensteori for en parameter θ nettopp er å posteriori tettheten for parameteren. La oss først se på punkttestimering av θ på grunnlag av å posteriori tettheten $\pi(\theta|D)$. Anta vi benytter estimatet a , mens den korrekte verdien altså er θ . Tapet ved å ikke anslå θ korrekt måles ved tapsfunksjonen $L(a, \theta)$. Bayesestimatet, $\hat{\theta}$, er den verdi av a som minimerer forventet å posteriori tap, dvs. oppfyller:

$$E[L(\hat{\theta}, \theta)|D] = \min_a \int_{\Theta} L(a, \theta) \pi(\theta|D) d\theta$$

Hvis vi har en kvadratisk tapsfunksjon, dvs.

$$L(a, \theta) = (a - \theta)^2,$$

er det lett å vise at

$$\hat{\theta} = E(\theta|D)$$

Bayesestimatet er da følgelig forventningen i å posteriori fordelingen.

Anta vi har en mengde $C \subset \Theta$ slik at

$$\int_C \pi(\theta|D) d\theta = 1 - \alpha.$$

Da kalles mengden en Bayes $100(1 - \alpha)\%$ *troverdighetsmengde* for θ . Hvis C er et intervall, kaller vi det et *troverdighetsintervall*. Et $100(1 - \alpha)\%$ troverdighetsintervall for θ er følgelig et intervall der vi med sannsynlighet $1 - \alpha$ tror at θ befinner seg *etter at dataene D er samlet inn*.

Anta for den eksponensielle modellen at å priori sannsynlighetstettheten for feilintensiteten λ er en gamma tetthet med formparameter a og skalaparameter b . Forsøket resulterer i at m enheter feiler etter en total testtid $\sum_{i=1}^m y_i$. Å posteriori tettheten er da gitt ved en gamma tetthet med formparameter $a + m$ og skalaparameter $b + \sum_{i=1}^m y_i$. Hvis nå $\tilde{\Lambda}$ har denne sannsynlighetstettheten, følger det at

$$2(b + \sum_{i=1}^m y_i) \tilde{\Lambda} \sim \chi^2_{2[a+m]}$$

Følgelig er:

$$\left[\frac{\chi^2_{\alpha/2, 2[a+m]}}{2(b + \sum_{i=1}^m y_i)}, \frac{\chi^2_{1-\alpha/2, 2[a+m]}}{2(b + \sum_{i=1}^m y_i)} \right] \quad (3.1)$$

et $100(1 - \alpha)\%$ troverdighetsintervall for λ .

La oss sammenligne med et klassisk konfidensintervall idet vi lar:

$$Y_i = \text{Levetid for } i\text{-te metallstreng, } i = 1, \dots, m$$

Siden nå:

$$2\lambda \sum_{i=1}^m Y_i \sim \chi_{2m}^2,$$

følger det at

$$\left[\frac{\chi_{\alpha/2, 2m}^2}{2 \sum_{i=1}^m Y_i}, \frac{\chi_{1-\alpha/2, 2m}^2}{2 \sum_{i=1}^m Y_i} \right]$$

er en $100(1 - \alpha)\%$ intervallestimator for λ . Legg merke til at fortolkningen nå er denne. Det er en sannsynlighet på $1 - \alpha$, før forsøket utføres, for at intervallestimatorens (stokastiske) grenser omslutter den faste parameteren λ .

Anta nå at forsøket utføres og vi registrerer at $\sum_{i=1}^m Y_i = \sum_{i=1}^m y_i$. Vi får da følgende $100(1 - \alpha)\%$ intervallestimat for λ :

$$\left[\frac{\chi_{\alpha/2, 2m}^2}{2 \sum_{i=1}^m y_i}, \frac{\chi_{1-\alpha/2, 2m}^2}{2 \sum_{i=1}^m y_i} \right] \quad (3.2)$$

Hva sier dette intervallet om hvor parameteren λ måtte befinne seg? Svaret er: ingenting! Enten ligger parameteren i intervallet eller så gjør den det ikke. *Klassisk inferens beskriver ikke vår usikkerhet om parameteren etter at dataene D er samlet inn.* Populært sagt - vi kjenner sannsynligheten for å treffe blinken før vi skyter, men vi vet ingenting om hva som skjedde etter at skuddet er avfyrt.

Anta nå at en likevel ved anvendelser feilaktig fortolker (3.2) som et $100(1 - \alpha)\%$ troverdighetsintervall for λ . Ved å sammenligne (3.1) og (3.2) svarer dette numerisk til en Bayes metode med á priori "gammatetthet" for λ med parametre $a = b = 0$, dvs.

$$\pi_0(\lambda) = 1/\lambda, \lambda \geq 0$$

Dette er en uekte á priori tetthet siden

$$\int_0^{\infty} \lambda^{-1} d\lambda = \infty.$$

Om vi ser bort fra dette siste, kan vi likevel slå fast at for de fleste anvendelser vil $\pi_0(\lambda)$ gitt over legge helt urealistiske store vekter på små verdier av λ . Hvis vi har mye data,

dvs. at en i (3.1) har $m \gg a$, $\sum_{i=1}^m y_i \gg b$, vil forskjellen på (3.1) og (3.2) være liten, og betydningen av á priori fordelingen være minimal.

Resultatet over gjelder faktisk ganske generelt. Har en mye data, vil en Bayes metode numerisk ofte lede til omtrent samme resultat som en velkjent klassisk metode. Har en lite data, vil forskjellen kunne være stor, og den klassiske metode vil implisitt ofte svare til en á priori informasjon som kan være helt urimelig.

La oss avslutte dette avsnittet med et eksempel som illustrerer Bayes metodikken ytterligere. Anta vi på bakgrunn av en á posteriori gamma tetthet med formparameter $a+m$ og skalaparameter $b + \sum_{i=1}^m y_i$ ønsker å beregne sannsynligheten for at en ny metallstreng overlever 1000 timers høy statisk belastning. La

$X =$ Levetid for ny metallstreng

Vi veier da overlevelsessannsynligheten $\exp(-\lambda 1000)$ med á posteriori usikkerheten i λ og får

$$\begin{aligned} P(X > 1000) &= \int_0^{\infty} e^{-\lambda 1000} (b + \sum_{i=1}^m y_i)^{a+m} \lambda^{a+m-1} \exp(-\lambda(b + \sum_{i=1}^m y_i)) / \Gamma(a+m) d\lambda \\ &= \frac{(b + \sum_{i=1}^m y_i)^{a+m}}{(1000 + b + \sum_{i=1}^m y_i)^{a+m}} = \left(1 + \frac{1000}{b + \sum_{i=1}^m y_i}\right)^{-(a+m)} \end{aligned}$$

Hvis vi nå har mye data, dvs. $m \gg a$, $\sum_{i=1}^m y_i \gg b$, får vi følgende tilnærmede uttrykt ved sannsynlighetsmaksimeringsestimatet $\lambda^* = m / \sum_{i=1}^m y_i$ for λ :

$$P(X > 1000) \approx (1 + \lambda^* 1000 / m)^{-m} \approx \exp(-\lambda^* 1000)$$

Dette siste er nettopp det klassiske sannsynlighetsmaksimeringsestimatet for den søkte sannsynlighet.

4. Et eksempel på en troverdighetsformel.

Anta vi er interessert i parameteren $\theta = 1/\lambda$ i vår eksponensielle modell. Den naturlig konjugerte á priori fordeling for θ er nå den inverse gammatettheten gitt ved:

$$\pi_0(\theta) = b^a \theta^{-(a+1)} \exp(-b/\theta) / \Gamma(a), 0 \leq \theta < \infty, a, b > 0$$

Anta videre at vi benytter en av følgende forsøksplaner:

- a) Fullstendig observasjon inntil et bestemt antall forsøksenheter har feilet
- b) Fullstendig observasjon inntil et fast tidspunkt
- c) Invers binomisk forsøksplan. Dette vil si at en enhet blir testet til den feiler eller til den når en bestemt alder. Deretter blir den erstattet med en ny enhet. Dette gjentas inntil et forhåndsspesifisert antall enheter har feilet.

La for alle planene

k =antall enheter som feiler

T =total testtid

Anta at (k, T) er observert. Uansett plan er (k, T) en *Bayes sufficient* funksjon av dataene, dvs. at for enhver á priori tetthet for θ vil á posteriori tettheten bare avhenge av dataene via (k, T) .

Ved bruk av den naturlige konjugerte á priori fordeling for θ blir á posteriori fordelingen også invers gamma der formparameteren a oppdateres til $a + k$ og skalaparameteren b til $b + T$. Bayesestimatet, $\hat{\theta}$, for θ når en benytter en kvadratisk tapsfunksjon blir nå

$$\begin{aligned}\hat{\theta} = E(\theta|(k, T)) &= \frac{b + T}{a + k - 1} = \frac{b}{a - 1} \frac{a - 1}{a + k - 1} + \frac{T}{k} \frac{k}{a + k - 1} \\ &= \hat{\theta}_0 \omega + \theta^*(1 - \omega),\end{aligned}\tag{4.1}$$

der $\omega = (a - 1)/(a + k - 1)$, $\hat{\theta}_0$ er forventningen i á priori fordelingen, mens θ^* er sannsynlighetsmaksimeringsestimatet. (4.1) kalles en troverdighetsformel (eng. *credibility formula*). Den viser at Bayesestimatet for θ kan skrives som et veiet gjennomsnitt av det tilsvarende estimatet á priori og sannsynlighetsmaksimeringsestimatet. Legg merke til at ettersom k , antall enheter som svikter, øker, blir mer vekt lagt på sannsynlighetsmaksimeringsestimatet og mindre på á priori estimatet. Formelen viser hvordan vårt endelige estimat for θ er balansert mellom ingeniørens skjønn og de rene data fra forsøket, og inviterer *formelig* til samarbeid mellom ingeniøren og statistikeren. Slike formler er brukt i forsikringsmatematikken siden før 1920.

Legg også merke til at så lenge $k = 0$, så vil $\theta^* = \infty$ uansett total testtid T . Dette er helt urimelig. Bayesestimatet for $k = 0$ er derimot

$$\hat{\theta} = \frac{b + T}{a - 1}$$

Dette er endelig for $a > 1$ og øker lineært i den totale testtiden T . Dette er naturlig. Dess lengre tid det går før noen enheter svikter, dess lengre vil vi anslå enhetenes forventede levetid å være. Bayesianere, uttrykker dette slik: "Bayesians learn while waiting (for the first failure)".

5. To begrensninger ved bruk av Bayesianske metoder.

- i) En naturlig konjugert á priori fordeling er som oftest tilstrekkelig brukbar til å beskrive vår opprinnelige usikkerhet om parameteren θ . Innvendingen mot å bruke

en slik fordeling er at dette tvinger á posteriori fordelingen til á være medlem av samme klasse av fordelinger uansett hva slags nye data vi måtte ha fått.

- ii) Hvis vi har en parametervektor $\underline{\theta} = (\theta_1, \dots, \theta_s)$, og det er urealistisk á anta at $\theta_1, \dots, \theta_s$ er uavhengige á priori, vil vi ofte ha store problemer med á modellere vår á priori innsikt om disse avhengighetsforholdene. Har vi lite data, vil á posteriori fordelingen kunne bli svært lite robust overfor endringer i á priori antagelsene om avhengighet.

Innvendingen i i) imøtegås ved á benytte numeriske integrasjonsmetoder ved oppdateringen i stedet for á velge en naturlig konjugert á priori fordeling. Innvendingen i ii) er langt alvorligere og kan virke dempende på bruken av Bayesianske metoder i modeller med mange parametre.

6. Anvendelser.

De miljøene som er kommet lengst i Norge når det gjelder anvendelser av Bayesianske metoder synes á være Norsk Regnesentral samt Senter for Industriforskning ved Arne Bang Huseby. En interessant anvendelse ble gitt av Abrahamsen et al. (1989) ved dette møte. I tillegg er foredragsholderen involvert i det fellesnordiske forskningsprogrammet TERO-TEKNOLOGI med prosjektet "Feildatainformasjon ved Bayesiansk statistikk". Dette programmet har til hensikt á oppná bedre pålitelighet og reduserte vedlikeholdskostnader for industrielt utstyr. Endelig er det en klar kobling mellom Kalman filtrering og Bayes metodikk, se f.eks. Meinhold og Singpurwalla (1983), som ytterligere understreker anvendelsesmulighetene for Bayesianske metoder.

Referanser

- Abrahamsen, P., Halvorsen, K.B. og Omre, H. (1989) Hvor mye olje er det her? - En anvendelse av Bayesiansk romlig statistikk. Foredrag Norsk Statistikermøte, Inderøy.
- Barlow, R.E. & Proschan, F. (1985) Inference for the exponential life distribution. In: Theory of Reliability, 143-164. Soc. Italiana di Fisica, Bologna, Italy.
- Meinhold, R.J. & Singpurwalla, N.D. (1983) Understanding the Kalman filter. The American Statistician, 37, 123-127.
- Natvig, B. (1989) Bayesianske metoder i pålitelighetsanalyse. Kapittel 4 i: Pålitelighetsanalyse med teknologiske anvendelser. Matematisk institutt, Universitetet i Oslo.

Vi har lest

Stephen M. Stigler: The History of Statistics. The Measurement of Uncertainty before 1900.
Harvard University Press 1986.

Selv om statistikkfaget er forholdsvis nytt, har det en interessant historie. Ved å kjenne denne historien får statistikere perspektiv på den aktiviteten de bedriver til daglig. Lærere i statistikk kan kanskje lettere forstå de begrepsmessige problemene studenter sliter med i dette faget, når man vet at heller ikke åndshøvdingen som Euler og Gauss uten videre kunne se den mest fruktbare måten å angripe statistiske problemer på.

Boken tar sitt utgangspunkt i Legendres lansering av minste kvadraters metode i 1805. Det er verd å merke seg at Legendre hadde rikelig med praktisk erfaring med data før dette gjennombruddet. Som medlem av kommisjonen for måling av meridianen fra Barcelona til Dunquerque i 1795 fikk han stifte bekentskap med store mengder geodetiske måledata. Dette arbeidet ledet til lanseringen av meter som $1/10\,000\,000$ av en meridiankvadrant. For å vise at Legendres resultat ikke var noen liten prestasjon gir Stigler oss eksempler på tidligere mislykkede forsøk bl.a. Euler som i 1749 hadde angrepet problemet med Saturns og Jupiters bevegelser. Han trengte å løse 75 ligninger med 6 ukjente, men kom ikke fram til noen brukbar metode.

Publiseringen av Legendres arbeide i 1805 tiltrakk seg stor oppmerksomhet bl.a. fra Laplace og Gauss. Gauss var ikke snauere enn at han hevdet å ha utviklet metoden før Legendre.

Minste kvadraters metode ble raskt tatt i bruk i astronomi og landmåling, men i samfunnsfag var man langt mer tilbakeholdende. Qetelet innskrenket seg mye til ren opplisting av data og til sine hjertesaker: Gjennomsnittsmennesket og normalpopulasjonen.

Gjennombruddet kom med Francis Galton, som via sin interesse for arv kom bort i "regresjonproblemet". Dette besto for Galton i det fenomenet at det synes som om sønner av fedre som avviker fra normal intelligens får intelligens som ligger mellom farens intelligens og normen. Siden Galton ikke hadde de nødvendige matematiske kunnskaper til å angripe problemet analytisk, gjorde han utspekulerte empiriske forsøk. Et med sukkererter ble avgjørende. Han hadde oppfunnet en "maskin", som han kalte Quincunx, og som vi kjenner igjen som et Brett med spiker plassert slik at kuler sluppet fra toppen illustrerer sentralgrenseteoremet anvendt på Binomialfordelingen. Denne brukte han til å lage en slags analog til det teoretiske resultatet at en blanding av normalfordelinger er normal. Det han hadde oppdaget var egentlig den binormale fordelig og dens velkjente egenskaper. Selve utledningen av formeluttrykket for den binormale fordeling fikk han matematikeren J. Hamilton Dickson til å gjøre ut fra de egenskaper Galton mente den burde ha.

Etter hvert overtok mer matematisk skolerte folk som Edgeworth, Pearson og Yule. Det ble Yule som omkring år 1900 fullførte den multiple regresjonsmodellen, slik vi kjenner den i dag.

Henrik Dahl

Bergens Tidende, 15. juli 1989

Nyheter

Nær annenhver mann i Bergen er utro

Bergens Tidende, 4. august 1989

TOR ARNE FANGHOL

Kvinner tenker på utroskap, mens menn gjør det. 45 prosent av menn i Bergen har vært utro, mens bare knapt 30 prosent av kvinnene har vært fysisk utro. Det viser en undersøkelse foretatt av studenter ved Bedriftsøkonomisk Institutt i Bergen. Undersøkelsen viser også at 40 prosent av dem som har vært utro, har fortalt det til sin faste partner.

Undersøkelsen bygger på et utvalg av 201 personer i Bergen. Studentene har forsøkt å vise sammenhenger mellom kjønn og utroskap, livssyn og utroskap og om det er spesielle aldersgrupper som er mer utro enn andre.



Bergens Tidende, 2. august 1989



Tvilsomme undersøkelser

Herr redaktør.

Bergens Tidende har den senere tid ofte referert resultater fra ulike undersøkelser om den vanlige bergensers synspunkter og øvrige gjøren og laten. Med store oppslag og kategoriske overskrifter sias «fakta» fast, også når det ikke ligger noen seriøs statistisk undersøkelse bak.

Ikke sjelden formidles resultater fra studentoppgaver (ortest BI), som typisk tar utgangspunkt i 200 spurte bergensere. Dette gir i beste fall feilmarginer på pluss/minus 7 prosentenheter på anslag for de tilsvarende størrelser blant alle bergensere, og trekkes det konklusjoner om undergrupper (f. eks for begge kjønn og ulike aldre), blir feilmarginene lett det dobbelte! Med så store marginer kan en sjelden trekke noen som helst generelle konklusjoner.

Når en så tar med at det sjelden dreier seg om tilfeldige utvalg i disse undersøkelsene (er bergensere i friluft i området rundt Torvalmenningen det?), blir det hele meningsløst.

Det ferskeste og av de mest groteske eksemplene er i BT 15. juli, er en undersøkelse av nevnte slag om bergenseres sidesprang i ekteskapet. Med fete typer kunnjøres at «Nær annenhver bergenser er utro». Dette skjer side om side med referat fra en seriøs undersøkelse foretatt av Statistisk Sentralbyrå om et annet samlivstema. At det ikke er sammenfall mellom overskriftens «er» og artikkelens «har vært» får så være. BTs oppslag ble for øvrig ukritisk videreført i NRKs lokalsending samme dag.

BT har ved en rekke anledninger også selv benyttet meningsmålingsinstitutter for å belyse lokale spørsmål, med svært lave stikkprøvestørrelser. Hvis man ikke er villig til å betale det en seriøs undersøkelse koster, burde man avstå. Man kan ikke skyldes på at leseren får ta det som det er, med en viss underholdningsverdi, så lenge avisen gir resultatene det minste skinn av palitelighet.

BT er blitt flinkere de senere år til å presentere uryddig statistisk informasjon grafisk, men det finnes dessverre nok av «statistiske analfabeter» i staben, og ikke alle er ferievikarer på jakt etter «agurkstoff».

Bergen, JOSTEIN LILLESTØL

Medias meningsmålinger

Herr redaktør.

Jeg viser til Jostein Lillestøls leserinnlegg i BT 1. august med tittelen «Tvilsomme undersøkelser».

I utgangspunktet er jeg enig med Lillestøls argument: Pressen er generelt for lite flink til å formulere utsagn i henhold til de statistiske begrensninger som ligger i en hver utvalgsundersøkelse, og i henhold til en undersøkelses seriositet. Dette gjelder bl.a. ved referat av studentoppgaver. Her bør pressen være mer varsom.

Når Lillestøl sier at Bergens Tidende benytter «svært lave stikkprøvestørrelser» i sine egne undersøkelser, så er dette direkte feil. I våre undersøkelser for Bergen opererer vi vanligvis med utvalgsstørrelser på 600 - 800 respondenter - en svært akseptabel utvalgsstørrelse for denne type undersøkelser.

Når det gjelder representativitet, så er det meningsmålingsinstituttens ansvar å påse at utvalgene blir trukket så tilfeldig som mulig. De instituttene som gjennomfører BTs redaksjonelle målinger et stort sett Opinion A/S og Norsk Gallup Institutt. Jeg har full tiltro til disse mht. måten utvalgene blir trukket på.

En ting er de tekniske begrensningene som ligger i en hver utvalgsundersøkelse (utvalgsstørrelse, representativitet etc.), en annen ting er formuleringen av spørsmålene og kommenteringen av resultatene. Spørsmålsformuleringer vil alltid være en kilde for diskusjon. Jeg tør imidlertid hevde at BT, i samarbeid med instituttene, har arbeidet mye med å få spørsmålene mest mulig «objektive» i henhold til den problemstilling man ønsker å belyse.

Personlig synes jeg også at redaksjonen er blitt langt flinkere til å kommentere og presentere undersøkelsesresultater.

Den største feil som fremdeles gjøres er at objektive, tekniske opplysninger vedr. undersøkelser ikke blir presentert. Når undersøkelsesresultater kommenteres, skal alltid den eksakte ordlyden av spørsmålene gjengis. Likledes skal publikum få opplysninger om utvalgsstørrelse, intervjuform, utvalgsområde og de statistiske feilmarginer som gjelder for undersøkelsen.

HELGE HOLBÆK-HANSEN
Markedskonsulent
Bergens Tidende

NYTT FRA AVD.C

Matematisk institutt, UiO

Cand.scient.

Pål Kjærsgaard : *'Metoder for eksakt beregning av pålitelighet basert på minimale sti-/kuttmengder'*

Veileder : Arne Bang Huseby

Publikasjoner

Gordon E. Willmot & Bjørn Sundt : *'On Posterior Probabilities and Moments in Mixed Poisson Processes'*

Storebrand, juni 1989

Bent Natvig : *'Pålitelighetsanalyse med teknologiske anvendelser'*

Universitetet i Oslo, august 1989

NYTT FRA UNIVERSITETET I BERGEN

(Generell oppdatering siden forrige gang TG brakte noe derfra, i desember 1987!)

Uteksaminerte cand. scient.

Stein Håkon Lygre: *'Ekstremverdiestimat av stasjonære Gaussiske prosessar basert på den avleia omhyllingskurveprosessen.'* Veileder: Trygve Nilsen (juni 1988).

Bjørn Henrik Auestad: *'Grafisk analyse av ikke-lineære tidsrekkemodeller.'* Veileder: Dag Tjøstheim (juni 1988).

Geir Ottersen: *'Express-88 - et statistisk ekspertsystem.'* Veileder: Ivar Heuch (desember 1988).

Aasmund Kvamme: *'Prissetjing av konvertible obligasjoner.'* Veileder: Jostein Paulsen (august 1989).

Ansettelses

Rolv Skjærven er utnevnt til professor i medisinsk statistikk og informasjonsbehandling ved Universitetet i Bergen.

KALENDER FOR INTERNASJONALE MØTER

1990

- 3.-5. april 5th Statistical & Scientific DataBase Management (SSDBM) Meeting, Charlotte, North Carolina, USA. Informasjon: Z. Michalewicz, Dept. of Computer Science, Victoria University at Wellington, P.O. Box 600, Wellington, New Zealand.
- 4.- 8. juni International Conference on Bootstrapping and Related Techniques, at Universität Trier. Informasjon: Prof. Dr. W. Sendler, Universität Trier, FB IV - Mathematik, Postfach 3825, D-5500 Trier, Federal Republic of Germany.
- 13.-15. juni 7th Annual Quality and Productivity Research Conference at the University of Wisconsin, Madison, USA. Informasjon: T.J. Snodgrass, Dept. of Engineering Professional Development, 801 Extension Building, Univ. of Wisconsin, Madison WI 53706, USA.
- 2 6.-9. aug. American Statistical Association, Joint Statistical Meetings, Anaheim, California, USA. Informasjon: ASA, 1429 Duke Street, Alexandria, VA 22314-3402, USA.
- 2 13.-18. aug. 2nd World Congress of the Bernoulli Society and 53rd Annual Meeting of the Institute of Mathematical Statistics, Uppsala, Sverige. Informasjon: Lars Holst, Matematiska Institutionen, Chalmers Tekniska Högskola, S-41296 Göteborg, Sverige.
- 19.-24. aug. Third International Conference on Teaching Statistics (ICOTS III), Dunedin, New Zealand. Informasjon: The Secretary, ICOTS III, Dept. of Mathematics and Statistics, University of Otago, POB 56, Dunedin, New Zealand.
- 9.-15. sept. COMPSTAT 1990, Dubrovnik, Jugoslavia. Informasjon: Mrs. Vesna Mildner, University Computing Center, 41000 Zagreb, Engelsova bb, Jugoslavia.

Call for Papers

Probabilistic Safety Assessment and Management (PSAM)

*An International Conference Devoted to the Advancement
of System-based Methods for the Design and Operation
of Technological Systems and Processes*

Sponsored by the Society for Risk Analysis

February 4-7, 1991 • The Beverly Hilton • Beverly Hills, California, USA

SCOPE

The purpose of PSAM is to provide a forum for the presentation of scientific papers covering both methodology and applications of system-based approaches to the design and safe operation of technological systems and processes. These include nuclear plants, chemical and petroleum facilities, defense systems, aerospace systems, and the treatment and disposal of hazardous wastes.

The following is a list of topics within the scope of the meeting:

1. Safety management and decision making.
2. Risk-based regulation.
3. Qualitative and quantitative design objectives.
4. Evaluation of alternate technologies and processes.
5. Probabilistic and non-probabilistic models for safety assessment.
6. Uncertainty analysis.
7. Uncertainties in physical and chemical phenomenology.
8. Expert judgment in safety studies.
9. Human reliability.
10. Risk-based methods for improving operator training.
11. Computerized control systems and operator aids.
12. Artificial intelligence in support of safety management.
13. Software system safety.
14. Implications of advances in computer power.
15. Earthquakes, fires, tornadoes, and natural phenomena.
16. Multi-hazard analysis, e.g., fires following earthquakes.
17. Vulnerability and safeguards analysis.
18. Aging of systems, structures, and components.
19. Reliability-based design.
20. Accident management.
21. Communicating the results of risk assessment and management to peers, decision makers, and the public.

INSTRUCTIONS FOR SUMMARY SUBMISSION

Four copies of a summary (800 - 1,200 words; typed, single-spaced) should be submitted to the General Chairman not later than Friday, April 6, 1990. Summaries should contain a title and include all authors' names, affiliations, and telephone, Telex, and FAX numbers. Authors should indicate the primary and one alternate category with which their papers are most closely identified using the numbering scheme in the Scope section of this announcement. Full papers will be due August 15, 1990.

GENERAL CHAIRMAN

Professor George Apostolakis
Mechanical, Aerospace, and Nuclear Engineering Dept.
University of California
405 Hilgard Avenue
Los Angeles, CA 90024-1597 USA
Telephone: (213) 825-1300
FAX: (213) 825-0761
TELEX: 3716012 UCLA LSA

SENIOR ADVISORY BOARD

B.J. GARRICK, *Pickard, Lowe, and Garrick, Inc., Chairman* • C.A. CORNELL, *Stanford University*
G.W. ELVERUM, JR., *TRW, Inc.* • K. INOUE, *Kyoto University* • T.A. MEAKER, *European Space Agency*
D. OKRENT, *University of California, Los Angeles* • V. PILZ, *Bayer AG*
N.C. RASMUSSEN, *Massachusetts Institute of Technology* • A.D. SLAY, *Gen. USAF Ret., Slay Enterprises, Inc.*
B. TOLLEY, *Commission of the European Communities*

LOCAL ORGANIZING COMMITTEE

D.G. BROWNE, *University of California, Los Angeles* • A.A. DYKES, *Pickard, Lowe, and Garrick, Inc.*
T. FOGLE, *TRW, Inc.* • L.R. FROEBE, *IT Corporation* • S.B. GUARRO, *The Aerospace Corporation*
S. KAPLAN, *Pickard, Lowe, and Garrick, Inc.* • M. KAZARIANS, *A.D. Little, Inc.*
R.J. MULVIHILL, *Planning Research Corporation* • M.G. STAMATELATOS, *ASI Systems International*
J.A. SUTCLIFFE, *University of California, Los Angeles*





NESTE NUMMER

Neste nummer av TG vil komme ut i midten av desember. Stoff til bladet må være redaksjonen i hende senest 1. desember. Det kan sendes til:

TG v/Ivar Heuch
Matematisk institutt
Universitetet i Bergen
Allegt. 55
5007 Bergen

Telefon (05) 21 28 35 (direkte)
(05) 21 28 38 (instituttkontoret)

STYRETS ADRESSE

Post som skal til styret i NSF og ikke til redaksjonen i TG skal sendes til:

Norsk Statistisk Forening v/Bent Natvig
Matematisk institutt, avd. C
Universitetet i Oslo
Postboks 1053, Blindern
0316 Oslo 3

TORGERSEN, LIV
REKSTRYKKEVERKET
BRAMMENVEIEN 60
OSLO 2

88