ISSN 0803-8953

TILFELDIG GANG
Nr. 2, årgang 17
Mai 2000

Utgitt av Norsk Statistisk Forening

# INNHOLD


| | Side |
|---|---|
| Redaksjon: | 1 |
| Fra redaktøren. *Odd Aalen* | 2 |
| Fra formannen. *Geir Egil Eide* | 3 |
| **Intervju** | |
| Fra sveivemaskiner til World Wide Web. Et intervju med Liv Høyland. *Bo Lindqvist* | 4 |
| **Fra fagets historie** | |
| Erling Sverdrups radioforedrag for 45 år siden. *Knut Håkon Brox* | 10 |
| Hva kan vi slutte av statistikk? Første foredrag. *Erling Sverdrup* | 11 |
| **Statistiske reisebrev** | |
| Chicago: De levde i rus og blues. *Håkon K. Gjessing* | 20 |
| Et reisende sabbatsår. *Nils Lid Hjort* | 22 |
| **Debatt** | |
| Greit nok med statistikk, men hvilke metoder er det som etterspørres? *Lars H. Vik* | 24 |
| Referat fra debattmøte i Statistisk Forening Oslo 6.4.2000. *Inge Helland* | 26 |
| Statistikkfaget fortjener et bedre rykte. *Bjørn Vassnes* | 29 |
| Om statistikeres image i norske aviser. *Xeni K. Dimakos* | 31 |
| Er tiden moden for nytenkende og mer kreativ rekruttering til statistikkfaget? *Ulf Indahl* | 32 |
| Yrkeslivets behov: Et bidrag til diskusjonen om statistikkfaget. *Tormod Næs* | 33 |
| Datavarehus og dataminig – hukommelse og intelligens? *Knut Opdal* og *Ulf Indahl* | 36 |
| **Pusleri** | |
| Pusleri-spalten. *Jostein Lillestøl* | 39 |
| **Faglig stoff** | |
| 500 skoleklasser får besøk av matematikere og statistikere! *Ørnulf Borgan* | 40 |
| Alarm! Statistikk i ny 3-årig økonomisk-administrativ utdanning. *Jostein Lillestøl* | 42 |
| **Meldinger** | |
| Innkalling til ekstraordinært hovedmøte. *Geir Egil Eide* | 45 |
| Programkomité for det 11. norske statistikermøtet i år 2001. | 46 |

# *Redaksjon:*

Odd Aalen (redaktør)
Grete Usterud Fenstad (redaksjonssekretær)
Ingrid Glad
Nils Lid Hjort

Redaksjonens adresse:
Odd Aalen
Seksjon for medisinsk statistikk
Universitetet i Oslo
P.b. 1122 Blindern, 0317 Oslo

Tlf: 22 85 12 36
Fax: 22 85 13 13
e-mail: odd.aalen@basalmed.uio.no

Trykking og utsending av TG utføres av
Statistisk sentralbyrå

## *Bidrag til Tilfeldig Gang:*

Alle slags bidrag til bladet er velkomne, enten det gjelder nye temaer av interesse for norske statistikere, eller man har kommentarer til innlegg som allerede har stått i bladet. Stoff til neste nummer må være redaksjonen i hende innen 1.9.2000.

Det greieste er at stoffet sendes inn som Word-fil (Word 97 eller lavere versjon) til redaktørens e-mail-adresse, se ovenfor. Alternativt kan vi ta en ren tekst-fil, men vi vil helst ikke ha TEX-format med mindre dette er nødvendig på grunn av matematikken.

Vi har etablert en nettversjon av TG, se følgende internettadresse:
*www.math.uio.no/nsf/web/index.htm . De som leverer bidrag til TG og som ønsker å reservere seg mot at bidraget deres legges ut på nettet, må gi beskjed om dette ved innsendelse av stoff.*

## *Fra redaktøren:*

*New England Journal of Medicine: Anvendelse av statistikk i medisin utgjør en av de store landevinninger på det medisinske fagfelt i siste årtusen*

Årtusenskiftet har gitt opphav til mange interessante betraktninger som både skuer fremover og tilbake. The New England Journal of Medicine, som er et av verdens mest anerkjente medisinske tidsskrifter, ga i årets første nummer en oversikt over de viktigste landevinninger innen medisinen de siste tusen år. Det listes opp 11 felter, fra grunnforståelsen av den menneskelige anatomi og fysiologi, så som hjertets funksjon og blodomøpet, til utvikling av røntgen og antibiotika og moderne DNA-tekonologi. Meget interessant er det at et av de elleve feltene er "Application of Statistics to Medicine"! Dette stiller altså på linje med oppdagelsen av den grunnleggende menneskelige fysiologi, av narkose, immunsystemet, antibiotika, røntgen og moderne medikamenter. Det er i sannhet et ærverdig selskap vi er havnet i!

Navnene som nevnes under omtalen av statistikk er velkjente: Fermat, Pascal, Gauss, Bayes. Videre omtales John Graunt som laget noen av de første dødelighetstabeller. Av de nyere navn, nevnes selvsagt Fisher, Neyman og Cox. (Neyman, som virket i Berkeley, USA, hadde neppe likt at han blir presentert som russisk immigrant, han var jo polakk og meget stolt av det!)

Hvorfor er statistikk så viktig i medisinen? Først og fremst skyldes dette fremveksten av de kliniske forsøk og av epidemiologien. De randomiserte kliniske prøvninger, som utføres i titusentall hvert år, danner det presise vitenskapelige fundament for all moderne medikamentell behandling, og også for andre medisinske intervensjoner. Epidemiologien dreier seg om hvilke forhold (risikofaktorer) som fører til utvikling av sykdom. På begge felter er målet å si noe om kausalitet. Når randomisering kan iverksettes, som ved de kliniske prøvninger, er dette forholdsvis uproblematisk. Verre er det innenfor epidemiologien der man vil ha en evig usikkerhet om hvorvidt en ukjent faktor er den egentlige forklaringen på de assosiajoner man finner. Likefullt, målet er å trekke kausale slutninger, og den klassiske, og overforsiktige, holdningen innen vårt fag at dette ikke er mulig, må man legge fra seg. Et poeng her er selvsagt at mange sannheter i medisinen har en mer begrenset levetid enn de evigvarende matematiske sannheter som utgjør den tradisjonelle, og litt nevrotiske, referanse-rammen innen statistikkfaget.

Nominasjonen av statistikkfagets inntreden på den medisinske arena blant de "Top Eleven" begivenheter innen medisin fra siste årtusen, sier også noe om statistikkens almenne betydning. Det er såvisst ingen grunn til noen duknakket beskjedenhet på vegne av statistikkfaget.

Odd Aalen

## *Fra formannen:*

Kjære medlemmer.

Styret har i henhold til $ 4 i foreningens vedtekter vedtatt å avholde ekstraordinært hovedmøte for å få gjort nødvendige oppdateringer av vedtektene såvel for foreningen som for foreningens fond. Detaljene gis i egen innkalling, men består i hovedsak av justeringer som skal gi a) mulighet for etablering av nasjonale fagseksjoner i foreningens regi, b) formannen en fremtidsrettet tittel, og c) utvidet anvendelsesområde for foreningens fond. Forslagene skulle være rimelig ukontroversielle og kan forhåpentligvis vedtas på møte i Grimstad i juni. Møtet vil også gi en ekstra mulighet utenom de regulære hovedmøtene (annethvert år) for kommunikasjon mellom styret og medlemmene om aktuelle saker.

Styret har også diskutert bruken av foreningens adresseregister spesielt med tanke på bruk av epost-adressene og har gjennomført en liten spørreundersøkelse blant medlemmene med epost-adresse om bruken (se annetsteds i TG). På bakgrunn av denne vil styret behandle saken pånytt i neste styremøte.

Med tanke på det 11. norske statistikermøtet i 2001, der Statistisk forening i Bergen er ansvarlig arrangør, har foreningen oppnevnt en programkomite bestående av statistikere såvel fra teoretiske og anvendte universitetsmiljøer som fra næringslivet. Innspill til komiteen (som er annonsert annetsteds i TG) om neste års program ønskes velkommen.

Den mest hektiske aktiviteten i foreningens organer den siste tiden har undervisningsutvalget stått for. Etter flere innlegg i forrige nr. av TG har en del medlemmer aktivisert seg og kommet med innspill til utvalget vedrørende undervisningsplanene for den videregående skolen. En høringsuttalelse vil bli sendt til Kirke- , undervisnings- og forskningsdepartementet. I tillegg har undervisningsutvalget også arbeidet med statistikkfagets plass i de nye planene for den økonomisk-administrative linje ved de statlige høyskoler. Her er det usedvanlig viktig å se ting i sammenheng og undervisningsutvalgets rolle i dette arbeidet er blitt av stadig større viktighet.

Til sist vil jeg nevne at foreningens netttjenester er under stadig forbedring. TG er igjen tilgjengelig på nettet etter å ha lagt nede i lengre tid, og i tillegg er foreningens hovedside oppdatert. Spesielt vil jeg nevne at en på hovedsiden nå har muligheten til å melde seg på foreningens nye uformelle og åpne diskusjonsforum: NSF-forum. Det er å håpe at så mange medlemmer og andre (!) som mulig benytter dette til utveksling av kunnskap og ideer.

Geir Egil Eide

# Fra sveivemaskiner til World Wide Web....

## Et intervju med Liv Høyland.

**Liv Høyland kan se tilbake på over 30 års virksomhet i statistikk-undervisning og har gått foran i bruk av datamaskiner og prosjekter i undervisningen. Men for Liv var veien fram til hovedfagseksamen i statistikk både lenger og mer svingete enn for de fleste.**

Bo Lindqvist

Da Liv Høyland i 1994 fratrådte sin stilling som førsteamanuensis i statistikk ved NTH, hadde hun formidlet grunnleggende statistikk-kunnskaper til norske sivilingeniører gjennom 26 år. Men Liv har sannelig ikke ligget på latsiden som pensjonist heller. Hun har blant annet undervist flere etterutdanningskurs i statistiske metoder for kvalitetsforbedring, og for tiden er hun av alle ting ansvarlig for instituttets første 100% internettbaserte kurs. Dette er et 3 vekttalls introduksjonskurs i statistikk, primært beregnet på lærere i den videregående skolen. Dermed har Liv vært aktivt medvirkende i en eventyrlig utvikling av regneredskaper brukt i undervisningen - fra de primitive sveivemaskinene som ble brukt på 50- og 60-tallet, via den første hullkort-datamaskinen "Nusse" ved Universitetet i Oslo, de første PC-ene på 80-tallet og endelig til dagens virkelighet med "ubegrenset" regnekapasitet og den verdensomspennende veven. Faktum er at Liv gikk i bresjen for å ta i bruk PCer i statistikkundervisningen.

**MINITAB og selvvalgte prosjekter i forsøksplanlegging**

- Etter en omlegging av grunnkursene i statistikk ved NTH i 1984 fikk jeg i oppdrag å legge opp det vi kalte Statistikk 2. Selv om det var endel vanskeligheter med tilgjengeligheten av PC-er, fikk jeg innført bruk av MINITAB i dette kurset. Meningen var at det skulle være et brukerkurs, og jeg la vekt på at studentene skulle forstå hva de gjør når de kjører en regresjonsanalyse eller variansanalyse. Øvingene på MINITAB var obligatoriske, og en av dem var større enn de andre og talte 10% av eksamenskarakteren. Studentene la ned mye arbeid i denne øvingen, og jeg tror de lærte veldig mye av å måtte jobbe selvstendig med et datasett på denne måten.

- Men ikke alle var like begeistret for kurset?

- Nei, noen av de mest teoretisk interesserte blant studentene på linjen for fysikk og matematikk klaget på at faget ble for umatematisk og "banalt". Men min ide var nettopp å gjøre faget mindre matematisk for å nå fram med den statistiske tankegangen. Og veldig mange fant Statistikk 2 både morsomt og viktig. Kanskje bidro det også til en utvidet bruk av statistikk blant sivilingeniører. Et godt bevis på dette er at mange kom tilbake senere under arbeidet med diplomoppgaven eller kanskje etter at de var begynt å arbeide, og fortalte at de etter mønster fra Statistikk 2 nå skulle utføre en regresjonsanalyse. Men ofte trengte de litt assistanse for å få dette helt til.

- I faget forsøksplanlegging gikk du enda lenger med prosjekter?

- Ja, og dette ble mitt favorittfag. "Statistisk forsøksplanlegging" har utvilsomt vært det morsomste og mest givende faget jeg har undervist. For å lære faget skikkelig er det nødvendig å planlegge, utføre og analysere

et forsøk på egen hånd. Prosjektene jeg ga i dette faget, gikk ut på nettopp dette, og med en helt selvvalgt problemstilling.

- Hva slags forsøk utførte studentene?

- Studentene presenterte et utall av forskjellige forsøk, om enn ikke alltid like seriøse. For eksempel var det en som skulle gå langrenn over en viss distanse og la opp et forsøk med følgende tre faktorer: skitype, støveltype og - som siste faktor - farge på lua! Ellers kunne forsøkene dreie seg om for eksempel popping av popcorn eller varighet av blomster. Populære "laboratorier" var studentbyenes felleskjøkken. En dr.ing. student ønsket å studere virkningen av alkohol på balanseevnen. Han lagde et sinnrikt apparat for å kunne måle balanse på en objektiv måte, og la virkelig sjelen sin i dette prosjektet!

- Hvordan ble faget mottatt?

- Det virker som om alle innså betydningen av å legge opp og analysere sine egne forsøk. Faget var også populært blant doktorgradsstudenter i ulike fag, som hadde bruk for forsøksplanlegging i forskningen. Flere har henvendt seg til meg med spørsmål etterat de kom i jobb. En morsom episode skjedde da jeg en gang sto i en kø i USA. Plutselig er det en mann som vinker og roper til meg på norsk lenger bak i køen. Han forteller at han har vært min student og at han nå er i USA for å legge fram resultatet av flerfaktorforsøket han gjennomførte i faget.

Prosjektet i forsøksplanlegging som Liv initierte, består fremdeles. Spesiell reklame fikk det under OL i Nagano i 1998, der skiskytteren Halvard Hanevold fortalte pressen om det statistiske forsøket han hadde gjennomført som del av sine sivilingeniørstudier. Han suderte ulike varianter av skytestilling og skyteprosedyrer, og kanskje var prosjektet medvirkende til OL-gull?

**Egen studietid**

- Hvordan hadde det seg at du startet å studere realfag?

- Jeg hadde alltid likt matematikk og tok artium på reallinjen i 1944. Jeg husker at vi var bare 4 jenter, mot 20 gutter, i klassen. Jeg startet realfagsstudier på Blindern i 1947 og tok matematikk bifag i 1948. Deretter tok jeg bifag i kjemi og botanikk og ble cand.mag. i 1950. Selv om matematikk var favorittfaget mitt og det jeg helst ville drive med, valgte jeg å gå videre med kjemi. Det var nemlig meget uvanlig at jenter tok matematikk hovedfag den gangen, så det kom liksom ikke på tale. Dessuten, med matematikk hovedfag var skolen stort sett den eneste mulige arbeidsplassen, og jeg hadde den gang slett ikke lyst til å undervise.

- Så du begynte på hovedfag i kjemi?

- Jeg tok jobb på landbrukshøgskolen på Ås og deretter på biokjemisk institutt ved Universitetet i Oslo. Samtidig tok jeg hovedfagslabber i kjemi. Men etter at Arnljot og jeg giftet oss og fikk barn, ble jeg hjemmeværende, slik det var vanlig den gangen. Studiene ble dermed lagt på hyllen.

- Men altså ikke for alltid?

- Vi dro til Berkeley høsten 1957, der Arnljot startet sine PhD studier. Høsten 1959 kom vi så tilbake til Norge, der Arnljot overtok et lektorat på Blindern. Jeg bestemte meg da for å ta opp igjen studiene, men jeg droppet kjemien. Jeg syntes jeg hadde nok med å koke på kjøkkenet hjemme om jeg ikke i tillegg skulle stå og koke på en lab om dagen. Jeg hadde nå lyst til å starte med statistikk. Siden statistikk ikke var opprettet som fag under mine adjunkt-studier ti år tidligere, måtte jeg ta opp faget helt fra grunnen av.

- Hvordan var det å begynne som student igjen etter disse årene?

- Selv om jeg var godt over 30 år, ble dette en fin opplevelse for meg. De andre studentene på kursene var veldig hyggelige og behandlet meg som alle andre på sin egen alder. Siden Arnljot underviste noen av kursene jeg deltok i, måtte Sverdrup steppe inn og gi eksamener. Selv om dette kunne være en ulempe for de andre studentene, var det ingen misnøye av den grunn.

- Du hadde nå tre barn. Hvordan var det med barnepass?

- De to eldste gikk i en barnepark en viss del av dagen, og den yngste, som ble født i 1960, hadde jeg med meg på Blindern. Her hadde de installert et rom for babyer og hadde en student til å se etter dem. Det er klart at dette var lite tilfredsstillende, men det fantes ikke barnehager slik som nå.

- Og så ble det statistikk hovedfag?

- Vi dro igjen til Berkeley i 1962, der Arnljot fullførte sin doktorgrad. Da vi kom tilbake i 1963, bestemte jeg meg for å ta statistikk hovedfag. Siden det var så lenge siden jeg tok min lavere grads eksamen, måtte jeg søke om å få ta embetseksamen. Sverdrup var her meget velvillig og satte opp en plan som jeg fikk godtatt.

- Hva var temaet for hovedoppgaven?

- Tittelen var "Estimation of Follow-up Studies", med Sverdrup som veileder. Men jeg traff ham bare en eller to ganger under arbeidet. Jeg tror ikke det var unormalt. Ellers vil jeg fremheve Sverdrup som en meget god foreleser.

- Var det vanskelig å motivere seg for studiene når du samtidig var mor og husmor?

- For meg var hovedfagsstudiet på det tidspunkt å betrakte som en hobby, og jeg syntes statistikk var gøy å holde på med. Men det er klart at jeg også hadde andre enn faglige ting i tankene når jeg var på forelesning. En gang, på en litt rotete forelesning der de fleste allerede hadde falt av lasset, kom jeg på at jeg måtte skrive handleliste for dagens husholdningsinnkjøp. Jeg satte i gang og skrev og skrev. Da de andre så dette, trodde de at nå kom det noen virkelige gullkorn fra foreleseren, som de måtte få med seg de også. Og alle skrev som besatt. Etterpå spurte de meg hva det var jeg fant så viktig. De ble nok litt snurt da jeg fortalte at det var handlelisten jeg skrev på!

- Men med tre unger hjemme ble det vel ikke mye tid til lesing?

- Nei, studiet ble nok litt tøffere enn jeg selv var klar over. Det ble mye lesing og arbeid utover natten etter at ungene var sovnet. Dette var slitsomt, og like før endelig muntlig eksamen i 1965 fikk jeg et skikkelig sammenbrudd. Men jeg kom da gjennom eksamen, selv om det tok lang tid før jeg kom ovenpå igjen. Så jeg vil gjerne råde de unge til å kjenne sin begrensning.

**Til Trondheim og NTH**

- Og så begynte du som foreleser i Trondheim?

- Vi flyttet til Trondheim høsten 1965, etter at Arnljot hadde begynt i sin stilling ved NTH allerede 1. januar. På det tidspunkt tenkte jeg ikke å bruke min statistikkeksamen til noe annet enn muligens noen konsulentjobber. Jeg var faktisk meget bestemt på at jeg ikke skulle forelese. Men i Trondheim var det mangel på lærere, så jeg ble overtalt til å gi noen forelesninger på Norges Lærerhøgskole (senere AVH, nå del av NTNU). Den første forelesningen var på min 40 års dag, så det var en spesiell måte å feire denne dagen på. Jeg syntes det var morsomt å undervise. Studentene var interesserte, og jeg likte å prøve å få dem til å forstå. Spesielt var det hyggelig at flere jenter tok statistikkfag. Når de så at jeg hadde greidd det, så måtte jo de også klare det, sa de.

- Hvordan kom du til NTH?

- Det ble opprettet en stilling som høgskolelektor i statistikk ved Matematisk institutt på NTH, og styreren Sigmund Selberg ba meg om å søke denne. Jeg var jo litt betenkt, siden Arnljot på den tiden var dosent på det samme instituttet. Men Selvberg mente at dette ikke hadde noen betydning, da *han* var sjefen på Matematisk institutt! Jeg hadde ikke annet enn hovedoppgaven å vise til, men da jeg nettopp hadde fått brev fra Tore Dalenius i Stockholm med forepørsel om jeg var interessert i å få denne publisert i tidsskriftet til International Statistical Institute, var dette tilstrekkelig til at jeg fikk fast ansettelse fra 1. mai 1968.

Dette var altså den spede begynnelse for statistikkmiljøet i Trondheim. NTHs og senere NTNUs statistikkgruppe har siden hatt en meget kraftig vekst. Et stort antall statistikere rundt omkring i landet har fått sin utdanning i Trondheim. Dette skal Liv og Arnljot ha æren for ved at de i alle år vektla kvalitet i undervisningen og det å ha et bredt og relevant undervisningstilbud. Økt aktivitet ga behov for flere ansatte, som i neste omgang betydde økt fagtilbud osv.

- Da vi begynte, var det grunnkurs i statistikk over to semestre for alle studenter ved NTH bortsett fra arkitektstudentene. Dette utgjorde i seg selv en stor undervisningsvirksomhet, men vi var også ivrige på å kunne tilby andre statistikk-kurs. Hver enkelt av oss underviste derfor alltid et videregående kurs ved siden av et grunnkurs. I tillegg engasjerte vi timelærere fra Oslo som kom og ga spesialiserte kurs.

- Var det vanskelig å få gehør høyere opp for behovene?

- Nei, jeg har inntrykk av at de fleste innså behovet for mer statistikkundervisning.

- Hvordan var det å forelese på "mannsbastionen" NTH i denne første tiden?

- Dette var jo tiden for studentopprørene, og det var nok noen som ville prøve å sette foreleseren til veggs ved å stille intrikate spørsmål. Men det var professorene som var de egentlige skyteskivene for opprøret, og det var også en fordel å være kvinnelig foreleser.

- Miljøet på instituttet?

- Jeg vil i den sammenhengen gjerne fremheve matematikeren Sigmund Selberg. Han gjorde instituttet til en hel familie, og var den naturlige familiefaren. I tillegg til NTH var det jo også matematikere på Lærerhøgskolen, og jeg minnes mange hyggelige fester blant Trondheims matematikere.

Det er her også naturlig å legge til at Liv og Arnljot opp gjennom årene har bidratt meget sterkt til å skape et godt og hyggelig sosialt miljø rundt seg.

**Kvinnesak**

- Du nevnte at du satte pris på at så mange jenter valgte statistikk. Ved studieretningen for industriell matematikk ved NTNU har vi i alle år sett at jenteandelen har vært mye større i statistikk enn i matematikk og numerikk. Er statistikk et jentefag?

- Det er mulig at jenter velger statistikk framfor matematikk fordi statistikk virker mindre teoretisk og mer matnyttig. Jenter blir skremt av ting som de ikke finner jordnære nok.

- Men mens sønnen har gått i sine foreldres fotspor og studert statistikk og aktuarfag, har dine døtre valgt annerledes. Ligger det noe bak her?

- Jeg har ikke prøvd å påvirke dem i noen retning. Men de var selv interesserte i å ta en utdanning, og det synes jeg personlig er det største jeg har oppnådd. For 40 år siden var det ingen selvfølge. Da ble det forventet av meg at jeg skulle være hjemme og ta meg av barna.

- Men vår yngste datter bestemte seg tidlig for at hun «iallefall ikke skulle gå på NTH». Og en dag slo hun i bordet under middagen og ropte: «Hører jeg *Wilcoxon* en gang til så drar jeg min vei!» Dette var i forbindelse med at vår eldste datter tok et statistikkfag for å få ekstrapoeng til medisinstudiet, og vi tenkte nok ikke på at statistikken godt kunne få hvile under middagen.

- Hva mener du om likestilling?

- Jeg synes mye positivt har skjedd i likestillingens navn. Det er helt klart store holdningsendringer i forhold til da jeg var ung, for eksempel at mannen i dag tar et mye større ansvar for hus og barn. Men jeg er absolutt imot alt som har med kjønnskvotering og ekstrapoeng for jenter å gjøre. Jeg tror det virker mot sin hensikt og gjør at jentene får stempel for å være dårligere enn gutta. For eksempel vil det kunne føre til uttalelser som «Hun ble ansatt fordi hun er kvinne».

**Forskning og faglig virksomhet**

Etter å ha vært først og fremst mor og husmor ved de første oppholdene i Berkeley i 1957-59 og 62-63, hadde Liv sitt første virkelige sabbatsår som fagperson ved University of California i Berkeley i studieåret 1979/80.

- Jeg fulgte en del kurs i Berkeley, men syntes at de var litt for teoretiske. Jeg ville for eksempel gjerne lært mer om forsøksplanlegging, men fant kursene lite rettet mot konkrete anvendelser. Men for all del, det var et fint år, og det var veldig morsomt å oppleve for eksempel Blackwell og Hodges som forelesere. Men jeg sa allerede den gangen at hvis vi skulle ha et nytt friår, ville jeg dra til et sted hvor de var mer opptatt av anvendelser.

- Og slik ble det?

- Ja, i studieåret 1987/88 dro vi til University of Wisconsin, Madison, og dette ble et svært utbytterikt år for oss begge. Vi kom i kontakt med "Center for Quality and Productivity Improvement", med George Box som grunnlegger og naturlige forgrunnsfigur. Her ble vi satt inn i et nytt og viktig felt, og vi fikk anledning til å følge kurs med forelesere som Bisgaard, Fung og Hunter. Spesielt sistnevnte var det en opplevelse å høre på.

- Hvordan ble interessen for kvalitetsforbedring fulgt opp etter Madison-oppholdet?

- Vi ble begge veldig engasjerte da vi var i Madison og kom oppglødde hjem med et ønske om å skulle revolusjonere industrien. Vi ville få ingeniørene i industrien til å innse betydningen av statistisk forsøksplanlegging. Jeg føler at vi kanskje ikke lyktes helt her, men vi har da nådd fram til mange gjennom å gi etterutdanningskurs i feltet. Spesielt har store bedrifter som Elkem og Hydro vært flinke til å sende ingeniører på disse kursene.

- Hvordan er dine erfaringer med å drive etterutdanningskurs?

- Det er veldig inspirerende å møte folk som gjør bruk av forsøksplanlegging i sitt daglige arbeid og som dermed er motiverte for å lære og forstå mer. Vi har også fått mange hyggelige tilbakemeldinger fra tidligere kursdeltakere som skriver til oss for å få hjelp til å komme videre. Vi gjorde forresten også reelle flerfaktorforsøk i kursene. Spesielt populært var forsøket der vi målte skumhøyde ved skjenking av øl i glass. Faktorene var øltype, glasstype og glassets vinkel. Deltakerne målte, regnet og konkluderte ivrig, og jeg tror dette var både morsomt og lærerikt.

Liv og Arnljot er kjent som flittige deltakere på konferanser. Spesielt har de deltatt meget regelmessig på de nordiske og norske statistikermøtene.

- Jeg reiste sjelden mens barna var små, men setter nå stor pris på å kunne møte kolleger for utveksling av ideer og tanker. Jeg synes det er viktig at mange møter opp, ikke minst på de norske statistikermøtene. Såvidt jeg husker var det Steinar Bjerve som tok initiativet til det første møtet på Røros i 1979. Det er flott at Norsk Statistisk Forening har klart å holde disse møtene i hevd siden. Arnljot og jeg ser nå forresten fram til å delta på det nordiske møtet i Grimstad i juni.

**Statistikk som hobby og yrke**

- Har det blitt tid til hobbyer?

- Ved siden av husholdning og familieliv har statistikken vært min hobby. Mye av fritiden har jeg brukt på å lese statistikkbøker. Men både Arnljot og jeg er veldig glade i naturen. Derfor har vi satt veldig stor pris på å kunne være på hytta ved Meråker så mye som mulig. Det er ingenting som det å gå i fjellet. Vi lo litt, forresten, da vi en gang var i Wien og så at de hadde laget en kunstig foss. Tenk så mye flottere det er å se på vår egen foss ved hytta, sa vi til hverandre. Jeg er ikke noe glad i å gå i store byer og se på bygninger og museer. Naturen er det flotteste for meg!

- Et standard spørsmål til slutt: Ville du gjort de samme valgene dersom du kunne gjort dem på nytt?

- Utvilsomt. Jeg er veldig glad for at jeg tok opp igjen studiene mine i 1959, og at jeg valgte statistikken. Mange på min alder forble hjemmeværende i alle år, men det hadde ikke vært noe for meg. Jeg har isteden vært priviligert som har kunnet ha hobbyen min som yrke. Likevel må jeg innrømme at jeg hele tiden har satt familien og de hjemlige sysler høyest. I jobbsammenheng har jeg sett det som min oppgave å formidle statistikk-kunnskaper til studentene på en slik måte at de blir interesserte og motiverte til å gå videre med studier. Forskning har det derfor blitt lite av for min del, for selv om lysten har vært der, har tiden ikke strukket til. Med min store interesse for faget var det heller ikke lett å gå av med pensjon. Det har derfor vært godt å kunne holde på med kursvirksomhet og annen faglig virksomhet også som pensjonist.

*Og det håper vi at Liv fortsetter med i mange år framover!*

# ERLING SVERDRUPS RADIOFOREDRAG FOR 45 ÅR SIDEN

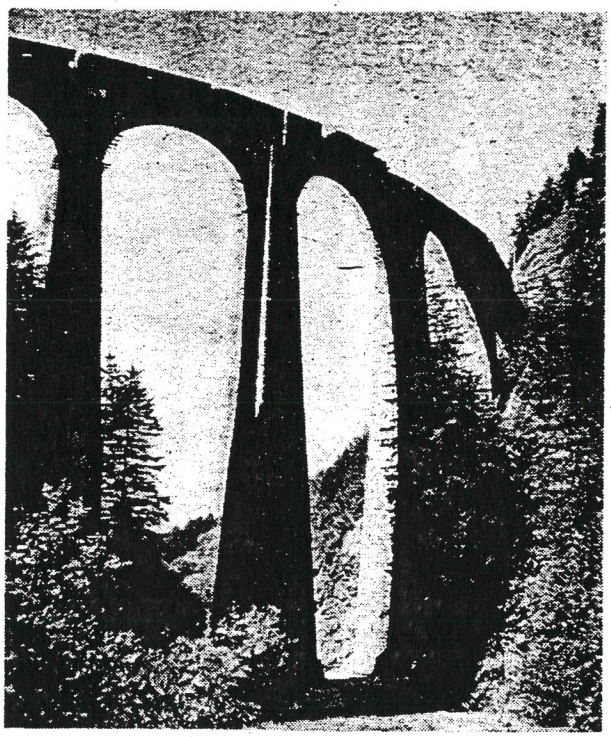
I serien «Universitetenes radioforedrag» bidro Erling Sverdrup på onsdagene 1., 8. og 15. desember i 1954 under tittelen «Hva kan vi slutte av statistikk ?». Foredragene gikk direkte på luften fra kl. 19.35 til kl. 20.00.

Erling Sverdrup (1917-94) var nylig utnevnt til professor i forsikrings-matematikk ved Forsikringsteknisk Seminar, og initiativet til hans del-tagelse i denne folkeopplysnings-virksomheten kom fra Trygve Haavelmo, leder av komiteen for «Universitetenes radioforedrag» og professor i økonomi.

Erling Sverdrup bygget senere opp utdannelsen av matematiske statistikere ved Universitetet i Oslo. Samtidig publiserte han vesentlige forskningsbidrag innenfor teoretisk statistikk og forsikringsmatematikk, og engasjerte seg bredt i anvendt forskning og forsikringsvirksomheten. En omtale av Sverdrups liv og virke finnes i festskriftet til hans 70-årsdag (*Twelve contributions to a tradition in statistical and actuarial science*, R. Norberg (ed.), Almqvist & Wiksell).

Manuskriptene til de tre foredragene ble funnet i Sverdrups etterlatte papirer i fjor høst. De gjengis i originalform her og i de to neste utgavene av dette tidsskrift. Opptak av foredragene finnes dessverre ikke.

Knut Håkon Brox



*De sveitsiske jernbanene går over pass og gjennom ville fjellandskaper. (Skolekringkastinga — klokka 12.25).*

Fra Programbladets annonsering for 8.12.54:

**13.30 Grammofonkonsert**

*Benedetto Marcello:* Konsert i c-moll for obo og strykeorkester: Allegro moderato. Adagio. Allegro. (Roger Reversy med Suisse Romande-orkestret. Dirigent Ernest Ansermet). (Decca LX 3100). *R. Strauss,* tekst *Mackay:* Heimliche Aufforderung, verk 27, nr. 3. (Hilde Konetzni, Hermann von Nordberg). (Col. LB 101). *Paul Dukas:* Villanelle. (Dennis Brain, Gerald Moore). (Col. DB 3300). *F. Schubert,* tekst *Rückert:* Du bist die Ruh', verk 59 nr. 3. (Dietrich Fischer-Dieskau, Gerald Moore). (HMV DB 21349). *M. Ravel:* Preludium, fra «Le Tombeau de Couperin», suite. (Suisse Romande-orkestret. Dirigent: Ernest Ansermet). (Decca LXT 2821).

**14.00 Lokale meldinger**

**Slutt**

*

**16.30 Minemeldinger**
**Kulingvarsler**

**16.35 Ettermiddagskonsert**

Bergensensemblet.
Leder: *Erling Knudsen.*
(Bergen).

*Alfred Evensen:* Norsk dans nr. 1.
*Jolly Kramer Johansen:* Seljefløyte.
*Pietro Mascagni:* Intermesso.
*Alphons Czibulka:* Kjærlighetsdrøm etter ballett.
*Johannes Hanssen:* Østerdalsrapsodi.
*Franz Heinz:* Humoreske.
*Louis Revel:* Cherry Ripe.
*Aksel Fredriksen:* Gavott.
*Erik Meyer-Helmund:* Sigøynerorkester.
*Peter Hindar:* Valser fra gamle Bergen.
*H. C. Lumbye:* Drømmebilder.

**17.30 Problemklubben**

**17.45 Gammel dansemusikk**

Arne Vikens kvartett.

*Anders Viken:* Lurifaks, polka.
*Adolf Paulen:* En aften på Alhus, reinlender.
*Anders Viken:* Gubben, masurka.
*Per Bolstad:* En aften på Mork, vals.
*Arne Viken:* Dullen, reinlender, og Glad som fuglen, polka.

**18.00 God norsk**

Bokmål.
Rektor *Marius Sandvei.*
(Kristiansund).

**18.15 Ungdomsrevyen**

(Tromsø).

**19.00 Værmelding**
**Dagsnytt. Aktuelt**
**Meldinger**

**19.35 Hva kan vi slutte av statistikk?**

Foredrag av professor dr. *Erling Sverdrup.*
(Universitetenes radioforedrag).

**20.00 Kirsten Flagstad synger sanger av Jean Sibelius**

Ved flyglet: *Waldemar Alme.*
Demanten på marssnön (J. J. Wecksell). Arioso. Den första kyssen, og Flickan kom ifrån sin älsklings möte (Johan Ludvig Runeberg). Var det en dröm? (J. J. Wecksell).

**20.20 Hva er hendt siden sist?**

Odda—Sulitjelma.

Radioforedrag.

# Hva kan vi slutte av statistikk ?

Erling Sverdrup.

Første foredrag.

Vi er etterhvert blitt vant til å se fenomener i natur og samfunn belyst ved tall eller statistikk. Slike tall er fremkommet ved å måle ett eller annet på bestemte steder og tidspunkter, men hensikten med tallene er nesten alltid at de skal få oss til å tro på noe almengyldig. Vi hører f.eks. om at en landbruksstasjon har utført noen statistiske forsøk med forskjellig slags kunstgjødning. Rapporten fra stasjonen forteller at en slags kunstgjødning ga større høstutbytte enn andre og det er åpenbart meningen at vi skal tro at denne kunstgjødning alment er bedre enn andre. Eller det kan være tale om statistiske forsøk for å belyse fangstutbytte ved nylongarn sammenlignet med alminnelig garn. Hensikten er naturligvis ikke bare å bringe på det rene hva slags garn har fisket best i konkrete tilfeller. Hensikten er å si noe alment om hvilket garn som fisker best. Vi hører også om fabrikker som ikke kontrollerer alle de ferdige produkter de sender ut på markedet, men f.eks. bare hvert femtiende. Produsenten mener på denne måte å få belyst kvaliteten av samtlige de produkter han sender ut på markedet. Eller det kan være målinger av virkningen av bestemte oppdragelsesmetoder og undervisningsmetoder eller målinger av folks intelligens, følelsesliv, meninger eller etterspørsle etter bestemte varer. Ved alle slike målinger er det meningen at vi skal ledes til å trekke almengyldige slutninger.

Folks innstilling overfor denslags statistiske undersøkelser synes å variere mellom to ytterpunkter. Som det ene ytterpunkt den usvikelig tro på alt som kan uttrykkes i tall eller som kan måles. Tallene er så å si det presise uttrykk for fakta. Den skotske

matematiker fra slutten av det forrige århundre, Lord Kelvin, har uttrykt det slik: "Når du kan måle det du snakker om og uttrykke det i tall, da vet du noe om det; men hvis du ikke kan måle det og hvis du ikke kan uttrykke det i tall, da er din viden tynn og utilfredsstillende." Hvis denne bemerkning skal tas bokstavelig, har lord Kelvin utvilsomt felt en altfor hård dom over mange fremragende forskeres resultater. Stikk imot Kelvins syn har vi det som gir uttrykk for ekstrem skepsis overfor alt som smaker av statistikk. Med statistikk kan man bevise alt det man ønsker å tro på, blir det sagt, og man får høre den gamle historie om at det er tre grader av løgn: løgn, grov løgn og statistikk. Det er fra tid til annen gitt mange eksempler på absurde eller motstridende resultater som man kan komme frem til ved å bruke statistikk. Slike eksempler kan være nyttige som en advarsel mot misbruk av statistikk, men hvis man av disse eksempler drar den slutning at statistiske undersøkelser fra naturens side er upålitelige, er man på villspor. Statistiske undersøkelser er naturligvis så pålitelige eller upålitelige som menneskene gjør dem.

Vi skal se på noen eksempler på hva slags vanskeligheter statistikkeren kommer opp i når han skal finne noe almengyldig av de begrensede opplysninger som et statistisk materiale gir ham.

Vi kan tenke oss at vi ved statistiske undersøkelser vil bringe på det rene hvilken av to fabrikater av vaskemasiner er best. Vi vasker like skitne plagg i begge maskiner og måler hvithetsgraden for hvert enkelt plagg etter vask. Vannets temperatur, vaskepulver og andre vilkår for vask i de to maskiner vil vi gå ut fra er de samme. Nå regner vi at gjennomsnitlig hvithetsgrad for hver av de to maskiner og sammenligner de to tall. Gjennomsnittet ved den første vaskemaskin viser seg å være større enn ved den andre. Men kan vi av dette slutte at den første type maskin er bedre enn den andre ? Det er klart at det ~~ville være store betenkeligheter ved å gjøre det~~ kan vi ikke uten videre gjøre, for, som vi

35 Bard

pleier å uttrykke det, forskjellen kan jo bero på tilfeldigheter. La oss ~~Vi kunne jo~~ tenke oss at vi hadde vasket alle plagg i en maskin og delt plaggene vilkårlig i to bunker. Det ville være høyst forbuasende om gjennomsnittlige hvithetsgrad for plaggene i de to bunker skulle vise seg å være den samme ~~like store~~. Den forskjell som vi ganske sikkert vil få ~~kan altså ikke bare på forskjellig fabrikat~~ skyldes andre omstendigheter som vil innvirke på hvitheten. Altså behøver ikke forskjellen i hvithetsgraden ved de to maskiner skyldes ~~fabrikat på tross av at vi har søkt å gjøre vilkårene for vaskingen i de to maskiner så like som mulig~~. Statistikerens problem er naturligvis ikke bare å fremlegge de to gjennomsnitt. Han må vurdere hvorvidt de gir grunnlag for å påstå at den ene maskin i det lange løp vil vaske best.

Dette var et eksempel på hva slags statistiske problemer som reiser seg når det gjelder å vurdere kvaliteten av varer.

La oss se på et eksempel fra et helt annet område. I Norge ble der i perioden 1946 - 1950 født omtrent 67 tusen barn. Av disse viste det seg at 17 pr. tusen født i ekteskap var dødfødt mens 25 pr. tusen født utenfor ekteskap var dødfødt. Det er nærliggende å se det høye antall dødfødte blant fødsler utenfor ekteskap i sammenheng med de sosiale og økonomiske kår under hvilke de ugifte mødre lever.

Men før vi innlater oss på denslags forklaringer er det kanskje best å vise litt forsiktighet. En ekstrem skeptiker kunne si: Av disse tallene kan vi da ikke slutte noen ting. Det er da bare et tilfeldig påfunn av statistikeren å inndele mødrene etter om de er gifte eller ikke. ett eller annet kjennetegn, f.eks. etter om de har rødt hår eller ikke. Det er klart at antall dødfødsler pr. 1000 ikke vil være det samme i de to grupper.

Skeptikeren har naturligvis rett i at vi ikke uten videre av den ting at antall dødfødsler pr. 1000 i ekteskap er større enn utenfor ekteskap kan slutte at det er større sjangse forat en nyfødt utenfor ekteskap vil være dødfødt. Spørsmålet er derfor om differensen mellom 25 og 17 i dette tilfelle er så stor at den bare med rimelighet kan forklares ved at sjangsen for dødfødsel i det store og hele er forskjellig i de to grupper, eller om den kan tilskrives tilfeldigheter.

Eksemplene med vaskemaskiner og dødfødsler er to eksempler på den situasjon som statistikeren til stadighet er i. Hans arbeide er ikke avsluttet når en har presentert resultater av målinger.

Det oppstår et spørsmål om tolkning av tallene. Hva kan vi si om de realiteter som ligger bak ? For å kunne besvare denslags

spørsmål må statistikeren kombinere det som på forhånd er kjent om det som er målt og talt med den nye viden som det statistiske materiale gir ham. Det som på forhånd er kjent er den faglige innsikt til teknikeren, legen, agronomen eller andre, alt etter hva statistikken gjelder. Å gjøre bruk av denne viden til å fortolke det statistiske materiale er statistikkrens oppgave.

Jeg har i det jeg har sagt tildels vært nødt til å bruke en relativt svevende uttrykksmåte. Jeg har brukt uttrykk som "differensen kan tilskrives tilfeldigheter", "sjangsen for dødfødsel", "gjennomsnittlig hvithetsgrad i det lange løp". Det stemme med en uttrykksmåte som er svært alminnelig blant folk. Man kommer neppe nærmere løsningen av det praktiske spørsmål ved å gruble over hva folk mener med disse uttrykk. Men uttryksmåten tyder utvilsomt på at ved statistiske undersøkelser er man nødt til å operere med sjangser og tilfeldigheter og da er det åpenbart best å gjøre det på en så rasjonell måte som mulig. Det er en gren av matematikken som beskjeftiger seg med sjangser og tilfeldigheter, nemlig sannsynlighetsregningen. / 15 min. Sannsynlighetsregningen er det naturlige hjelpemiddel for statistikeren. I denne fagvitenskap erstatter man de svevende uttrykk som jeg mener med presise matematiske begreper og man lager seg det tankeverktøy som statistikeren trenger. Det er bruken av dette verktøy som har gjort det mulig å utvikle rasjonell metoder til å tolke statistiske materialer. Man får derved et grunnlag for å vurdere slike spørsmål som jeg nevnte. Statistisk teori basert på sannsynlighetsregning kalles <u>matematisk statistikk</u> og det er denne gren av den statistiske teori som er emne for mine foredrag. / 10 min. Matematisk statistikk og sannsynlighetsregning brukes i dag i en rekke fagvitenskaper, i medisin, landbruksforskning, industriell forskning, biologi, meterologi, forsikring, fysikk o.s.v.

For å gi et innblikk i statistikkerens tenkemåte skal jeg først si noen ord om sannsynlighetsregningen.

Sannsynlighetsregningen ble skapt i det 17. århundre, det århundre som var så rikt på pionerer i matematikk. Mot slutten av århundre skapte Leibnitz og Newton differential- og integralregningen, det uundværlige hjelpemiddel for nesten all praktisk anvendelse av matematikken, også anvendelsene i statistikken. Ved begynnelsen av århundre skapte filosofen og matematikeren Descartes den analytiske geometri og dermed anskueliggjørelsen av tallmaterialer ved grafiske billeder. Diagrammer er jo i dag blitt allemannseie, de brukes av forretningsmenn, teknikere, reklamemenn o.s.v.

Sannsynlighetsregningen oppsto uavhengig av disse landevinninger. Den ble opprinnelig utviklet som en matematisk teori for hasardspill. I tider da man innen visse samfunnslag risikerte anseelige pengesummer på kompliserte spill med kort og terninger, oppsto det naturligvis et sterkt behov for rasjonelle metoder til å beregne fordeler og mangler såvel ved de forskjellige spilleregler som ved de forskjellige måter å spille på. Det var for å få utført en slik beregning at en lidenskapelig fransk spiller Chevalier de Mere i 1654, altså for temmelig nøyaktig 300 år siden henvendte seg til den kjente filosof og matematiker Blaise Pascal for å få løst følgende problem: Anta at to spillere spiller mot hverandre og hver av dem gjør samme innsats på 32 pistoler, pistoler var en europeisk pengeenhet på den tid. Det spilles en serie spill. Ved hvert spill har de to spillere like stor sjangse til å vinne. Den som først vinner 3 spill kan innkassere hele innsatsen på 64 pistoler. Spørsmålet var nå: Anta at spillet avbrytes når f.eks. den ene spiller har vunnet 2 spill og den anden 1 spill. Hva er en rettferdig fordeling av innsatsen ? Pascal løste dette problem. Han fant at første spiller skulle ha 48 pistoler og andre 16: Hvis den første spiller hadde vunnet 2 spill og den andre ingen skulle den første ha 56 pistoler og den andre 8 pistoler. Løsningene forutsatte at det var rettferdig å fordele gevinsten i forhold til de to spilleres sannsynlighet for å vinne på det tidspunkt da spillet ble avbrutt. Det trengtes altså en klar definisjon av begrepet sannsynlighet samt regler for hvorledes man skulle regne med dette begrep. Dette utvikler Pascal i en brevveksling med en annen betydelig matematiker fra det 17. århhndre, nemlig Pierre Fermat. Disse to regnes derfor vanligvis som grunnleggerne av sannsynlighetsregningen.

Noen vil kanskje undre seg over at betydelige matematikere som Fermat og Pascal kunne kaste seg over et slikt problem som dette problem om delingen av innsatsen med slik begeistring.

Forklaringen er neppe at det var så store økonomiske interesser knyttet til hasardspill. Grunnen er vel heller at matematikere til alle tider er blitt tiltrukket av problemer som stiller store krav til oppfinnsomhet og tankevirksomhet uten hensyn til den direkte praktiske betydning. Men det betyr ikke at matematikere bare blir tiltrukket av problemet på samme måte som en sjakkspiller som finner adspredelse i intens tankevirksomhet. Matematikerens interesse blir vakt når han øyner noe prinsipielt nytt i tankegangen. Han vil derfor alltid forsøke å generalisere og han vil rasjonalisere sin tankevirksomhet ved å søke en abstrakt formulering. Det er derfor ikke så rart at teorier som tar utgangspunkt i tilsynelatende upraktiske problemer allikvel har vist seg å få stor praktisk betydning.

I slutten av 1600-tallet ble det løst en rekke problemer fra hasardspill etter de prinsipper som Pascal og Fermat hadde grunnlagt. I dag kan man vel si at alle spørsmål omkring beregning av sjangser ved hasardspill i prinsippet er løst. Det er bare å sette seg ned å regne etter kjente prinsipper.

For statistikerne er det alltid en nyttig tanketrening å pusle litt med slike problemer. I hasardspill har man jo denslags sjangser og tilfeldigheter, som man også møter i statistikken, i rendyrket form.

La meg ta noen eksempler på hva man kan regne ut om spill ved å bruke sannsynlighetsregning. Når man deler ut kortene i bridge er sannsynligheten forat en av spillerne skal få alle 13 kort i samme farge, lik et tall som kan skrives slik. Det er null-komma og så må man skrive 10 nuller etter komma og dernest sifrene 2 5. Hvis 1000 bridgeklubber spilte 100 spill om dagen ville det allikevel være sjeldnere enn hvert 1000-år at en spiller fikk alle kort i en farge. Forutsetningen er naturligvis at kortene er skikkelig stokket.

Som et annet eksempel la oss tenke oss at alle innbyggere i en storby er på fest og så beslutter seg til å gå til sengs helt og holdent på måfå. Sannsynligheten forat ingen innbygger skal havne i sin egen seng viser seg å være forbundet med en kjent matematisk konstant, nemlig grunntallet i det naturlige logaritmesystem. Denne konstant er like fundamental i matematikken som tallet $\pi$ . Skrevet med to desimaler er tallet lik omtrent 2.72. Sannsynligheten for at ingen personer skal havne i sin egen seng er nå lik 1 dividert med dette tall, d.v.s. omtrent 37 o/o. Det er altså faktisk en sannsynlighet på 63 o/o, eller nesten $\frac{2}{3}$ forat minst en person ved en tilfeldighet skal havne i sin egen seng.

Fra begynnelsen av 1700-tallet begynte en rask utvikling innen sannsynlighetsregningen. I 1713 utkom et verk av Jakob Bernoulli med titelen "Ars conjectandi", d.v.s. "kunsten å gjette". Jakob Bernoulli var et av medlemmene av den berømte sveitsiske matematikerfamilie, hvor matematisk begavelse synes å ha gått i arv. Hos ham finner vi presist formulert antagelig for første gang det som vanligvis kalles de store talls lov. Dermed var noe grunnleggende nytt kommet til i sannsynlighetsregningen som ga perspektiver langt utover anvendelsene på hasardspill.

Jeg skal først illustrere loven med at eksempel fra hasardspill. Vi tenker oss at vi kaster med et pengestykke og noterer resultatene av de enkelte kast, d.v.s. vi noterer om resultatene er krone eller mynt. Vi vil da se at rekkefølgen av krone og mynt er helt kaotisk, det er ikke mulig å finne noen enkle lovmessigheter i rekkefølgen. Hvis man derfor vil forsøke å forutsi resultatet av et enkelt kast, blir dette vill gjetning, Men hvis man vil forsøke å forutsi hvor mange ganger vi vil få krone ved 100 kast, er vi ikke lenger henvist til å gjette. Anta nemlig at vi gjorde f.eks. 10 000 kast og inndelte disse i 100 serier hver på 100 kast. Da vil man oppdage at antall ganger krone i hver serie

på 100 vil være grovt regnet det samme og svinge rundt 50 forutsatt at pengestykke ikke er skjevt. Serier hvor det er mer enn 60 eller ferre enn 40 ganger krone vil være relativt sjeldne. Hvis vi nå istedenfor serier på 100 kast hadde betraktet serier på 1000 kast ville det prosentvise XXX antall ganger vi fikk krone, svinge mello snevrere grenser, d.v.s. de ville gjennomgående ligge nærmere 50 o/ Jo lengere serier vi betrakter desto sikrere kan vi forutsi det prosentvise antall ganger vi vil få krone. Det kommer altså tilsyne en regelmessighet i observasjonene når antall observasjoner er stort. Man pleier å kalle denne regelmessighet eller lovmessighet de store talls lov.

Slike reglemessigheter finner vi igjen i statistiske materialer hentet fra andre felter enn hasardspill. Vi vet at det er ren gjetning å forsøke å forutsi om det i et konkret filfelle vil fødes en gutt eller pike, men vi kan med meget stor nøyaktighet forutsi hvor stor prosent av fødslene i 1955 vil være guttefødsler. Sjangsen for guttefødsel har nemlig holdt seg forbausende konstant gjennom tidene. I 5-årsperioden 1906-1910 var 51.43 o/o av fødslene i Norge guttefødsler. I perioden 1946-1950 var samme tall 51.67 o/o. altså 51.43 mot 51.67. 30 år senere.

I sannsynlighetsregningen formulerer man de store talls lov som en rent matematisk setning som kan utledes utfra gitte forutsetninger. Men de statistiske fenomener som jeg nettopp har omtalt kan naturligvis like lite som de fleste andre empiriske lover utledes med logisk nødvendighet. Men paral lellen mellom de empiriske fenomener og den matematiske setning gir oss en pekepinn om at statistiske materialer med fordel kan behandles ved hjelp av sannsynlighetsregning.

De store talls lov har beskjeftiget matematikere helt opp til våre dager. I 1920 og 1930-årene ble således loven generalisert på forskjellige måter av en gruppe matematikere i Moskva. 23 mm.

Helt opp til dette århundre har det värt almindelig å anta at anvendelsen av sannsynlighetsregningen til å tolke et statistisk materiale forutsetter at vi har et stort antall målinger, for, pleide man å si, det er de store talls lov som gjör det berettiget å snakke om sannsynlighet. Nå vilde naturligvis dette bety at mulighetene for å bruke sannsynlighetsregningen i statistikk var sterkt begrenset. Det er nemlig klart at et lite statistisk materiale innsamlet ved omhyggelig planlegging ofte kan si adskillig mer enn et stort materiale mer vilkårlig innsamlet. Det er en av de betydelige nyskapninger i de siste 20-30 år i statistisk tenkning at man ~~er blitt~~ blev klar over at sannsynlighetsregningen er et uhyre nyttig hjelpemiddel også når små tallmaterialer skal bringes til veie og tolkes. I virkeligheten betöd dette en ~~xxxxx~~ en hel revolusjon i statistisk teori, en revolusjon som fremfor alt skyldes den store britiske statistikker R.A. Fisher. Vi skal komme tilbake til dette senere i disse fordrag, men först skal vi se litt närmere på de store talls lov.

Historisk har de store talls lov spilt en betydelig rolle og ved mange anvendelser er den natuligvis fremdeles viktig. Det blir ~~emne~~ for neste fordrag å se litt på dens ~~historiske~~ betydning.

## *Chicago: de levde i rus og blues*

Håkon K. Gjessing

Antagelig med tanke på eventuell fordelaktig psykoanalytisk effekt har redaktøren oppfordret meg til ikke å fortrenge minnene fra mitt Chicago-opphold men snarere legge dem ut for allmennheten å skue. De ytre omstendigheter var forsåvidt anstendige nok. Per Mykland (http://galton.uchicago.edu/~mykland/) (som selv nærmest ukentlig forlater Chicago for utenlandsopphold av lang varighet, en meget anbefalelsesverdig strategi) spurte om jeg ville søke om et år eller to ved Department of Statistics, University of Chicago (ikke å forveksle med University of Illinois at Chicago). Dette gjorde jeg, og kunne kort etter i min uvitenhet glede meg over et positivt svar. Avtalen var grei nok, med 4 timer undervisning pr. uke (de fleste "faste" har 3 timer men administrasjon og studentveiledning i tillegg).
Så i september .98 slepte konemor og jeg våre 30kg kofferter til vår "nye" leilighet i Hyde Park, ca. 10km sør for sentrum, men bare 10min gange fra universitetet. Hyde Park er (etter forholdene!) et rolig boområde med universitetet i sydenden. Mye av universitetet er bygget i engelsk gotisk stil, i kalkstein, og er vel verdt en rundtur. Da det ble bygget ble det sagt at det var en imponerende kopi av ærverdige engelske universiteter, men at vinduene var for rene og interiøret virket for moderne. Dette er nå ordnet.
Downtown Chicago er noe for seg selv, flott beliggende ved Lake Michigan, og med skyskraperne i sterk konkurranse om .... å skrape skyer. Sears tower er ikke lengre verdens høyeste (i strukturell høyde), men man har funnet ut at øverste etasje (110ende) er verdens høyeste "bebodde" etasje. Så da så. John Hancock Center er med sine 100 etasjer og spesielle reisverk "the World's most Renowned Building", ihvertfall i følge dem selv. Spor etter tradisjonell norsk arkitektur kan man også finne. I Chicago Tribune Tower er det i veggene murt inn biter av kjente bygninger fra hele verden, biter som ble stjålet av avisens utenriksreportere gjennom årene. Blant mye annet har en bit av Nidarosdomen funnet veien hit (!). Forøvrig er det selvsagt et mylder av restauranter, bluesklubber, butikker, museer, parker, (sand)strender og mye annet som kan være til glede for besøkende til alle årstider. Chicago ble i stor grad grunnlagt på slakteindustri, og det ble sagt at det eneste på grisen som ikke ble benyttet kommersielt var skrikene. Dette var neppe helt riktig, for en god del slakteriavfall ble dumpet i Chicago river som renner ut i Lake Michigan. Siden dette er stedet de også henter sitt drikkevann skapte jo det visse problemer, inntil noen fant ut at man kunne snu strømretningen på elven og sende det inn i Mississippi i stedet. Til stor glede for dem som bor *der*.

Det har gjennom tidene myldret av Nobelprisvinnere ved UC. Merton Miller, som selv har fått en, har uttalt at dette måtte skyldes vannet, det kunne ihvertfall ikke være kaffen som gjorde det. Etter å ha smakt vannet vil jeg vel betvile det. Men mange dyktige folk jobber her, det står ikke til å nekte. Ved Dept. of Statistics er vel kanskje Peter McCullagh det navnet de fleste kjenner. Michael Stein (romlig statistikk) er for tiden Chairman. Ellers knirker fortsatt gamle Bill Kruskal rundt i gangene, en hyggelig pratefelle. Og av andre hyggelige pensjonister må selvsagt nevnes Patrick Billingsley, som gjerne forteller om sine opplevelser da han spilte usympatisk fagforeningsleder i "The Dollmaker", hvor Jane Fonda forøvrig var den store publikumsavskrekkeren. Stephen Stigler er et navn man gjerne kjenner fra hans spennende bok "The History of Statistics" (selv om tittelen kanskje ikke automatisk gjør den til en internasjonal

bestselger.....). Han er en fargerik person med et utall av gode historier, og jeg fulgte hans forelesninger om statistikkens historie, en sann fornøyelse. Det er gøy å få levendegjort de "gamle gutta" som man ellers bare ser nevnt i forbifarten. Stigler selv har en klar favorittstatistiker (som kanskje overrasket meg noe), nummerskiltet på hans Lexus lyder "GALTON". Et åpenbart tap for instituttet var at den muntre kineser Augie Kong (statistikk & genetikk) forlot dem for å jobbe i et stort genetikkfirma (Decode Genetics) på Island, hvor de etter beste evne prøver å kartlegge (rappe?) viking-gener. Hans landsmann Xiao-Li Meng er blant dem som anbefaler folk å samle så lite data som mulig, hans favorittfelt er "missing data" og dataimputering..... . Ronald Thisted har blant annet skrevet boken "Elements of statistical computing" og er mye opptatt av biostatistikk/medisinsk statistikk. Foruten en del andre er det jo så også "vår egen" Per, som driver mye med finansrelatert statistikk og stokastikk. Etter køen foran kontordøren å dømme er han en populær veileder (ikke minst blant kinesiske studiner...), det er imponerende at han kan få tid til egen forskning innimellom, hvilket han faktisk gjør. Men det er ikke rart at han er populær, endel av de andre ansatte kan nok virke temmelig skremmende på unge og uskyldige studenter, selv om de i bunn og grunn (sannsynligvis) er anstendige mennesker. Apropos uskyldige studenter, Steins kontor ble brannbombet kort tid før jeg forlot Chicago. Kontoret ble fullstendig utbrent, og røyken trakk inn på andre kontorer via ventilasjonsanlegget. Det eneste han hadde igjen var noen bøker jeg hadde lånt, som sto på mitt kontor. Det var nærliggende å mistenke misfornøyde studenter, men jeg tror ikke saken er oppklart.

For mitt eget vedkommende vet jeg ikke hvor vellykket jeg kan si oppholdet var. Å være lønnet derfra gjorde nok at jeg fikk mer innsikt i hvordan ting fungerer, og på den positive siden at jeg fikk bedre kontakt med studenter og ansatte. På den annen side opplevde jeg universitetet som svært byråkratisk, det var en lang prosess bare å få låne bøker på biblioteket. Instituttets sekretær, som var pitbull i tidligere liv, hadde som hovedoppgave å sørge for at midlertidig tilsatte gledet seg til avreisedatoen. Til annet var hun liten hjelp med mindre du gjerne ville ha foretatt en amputasjon el.l. På hjemmefronten ble husets fasade revet ned og gjenoppbygget mens vi bodde der, dette holdt de på med under *hele* oppholdet, bråket og støvplagene var ubeskrivelige.

Undervisningen var lærerik (ihvertfall for meg), og studentene entusiastiske. I vinterkvartalet foreleste jeg i UC's downtown lokaler, om opsjonspriser, rentemodeller etc. for studenter som skulle ut og bli rike (via UC's Financial mathematics program www-finmath.uchicago.edu/). I vårkvartalet holdt jeg et kurs i eksperimentdesign for graduate-studenter ved instituttet. Undervisningen var preget av at man står meget fritt i valg av innhold. Den er imidlertid nokså arbeidskrevende pr. forelesningstime siden det gis mange obligatoriske innleveringer, hjemmeeksamener etc. underveis. Selv med assistenter til rettingen blir det mye arbeid å lage opplegget. På den klart positive siden kom instituttets "consulting", som tar imot henvendelser fra andre institutter om statistiske problemstillinger. Disse var jeg med å sortere og fordele til studenter, som så jobber med dem og gir en presentasjon av prosjektet på et felles møte. En meget god ordning som gir studentene verdifull erfaring! (og mindre administrasjon enn man kanskje skulle forvente).

Siden jeg valgte å "bare" være ett år gikk alt for mye av tiden med til praktisk organisering og meningsløse ergrelser. Men da jeg reiste virket ikke akkurat tanken på ett år til spesielt fristende. "Don't throw good money after bad", som de sier. Men om ikke annet føltes det jo svært godt å komme til Oslo etterpå, selv for en Bergenser!

# Et reisende sabbatsår

**Nils Lid Hjort**

Blant det akademiske livs store privilegier er muligheten til å kunne begi seg ut i en lang, statlig støttet arbeidslørdag sånn omtrent hvert syvende år. Man får (i prinsippet) fri fra undervisning og administrasjon, og kan altså fordype og fordoble seg i ens forskningsprosjekter. Sabbaten er også en kjærkommen anledning til å lade batterier og la seg inspirere av nye temaer, oppgaver og kolleger. Vår generøse stat lar også sine universitetsforskere slippe ut til interessante steder utenlands, om man har ordnet seg med nettverk og invitasjoner.

Selv har jeg hatt stort utbytte og varige gleder av slike opphold tidligere, til henholdsvis Stanford og Oxford, medbringende felle og barn. Denne gangen var imidlertid familiebeslutningen at vi ikke skulle ut i samlet flokk, men at vi og jeg fortsatt skulle ha Oslo som base. Fra denne base planla jeg derfor diverse faglige avstikkere.

Altså har dette vært et annerledes sabbatsår for meg. For det meste har jeg hatt mitt virke på Blindern. Jeg tilhører dem som faktisk liker å undervise, mesteparten av tiden, men det *er* behagelig, og inngyder en følelse av tidsluksus, å høre kollegene i korridoren skynde seg til forelesninger, mens man selv kan sysle videre med egne saker. Likevel blir ikke lørdag på Blindern det samme som sabbat i annen by; studentene vet hvor du er og man kan ikke la være å engasjere seg i noe av møteglamet. Og suset i palmene høres ikke så godt.

Det er derfor de faglige ekskursjoner i ulike himmelretninger jeg vil minnes mest fra det akademiske år sommer '99 til sommer '00. Man skal ivareta nettverk, påskynde samarbeidsprosjekter, og finne på nye ting.

Min liste starter med konferanser i Reading i juli og i Pavia i september, der det også ble tid til noen arbeidstimer med forskningsfeller Stephen Walker og Sonia Petrone (begge har vist seg på Blindern). Så bar det til Механико-математический факультет Московского государственного университета i oktober, der mitt stotrende russisk får kjørt seg (de snakker veldig *fort* i det landet). Nikolai Ushakov besøkte Blindern sommeren 1998, noe som ledet til prosjekter sammen med Ingrid Glad og meg selv.

Det er tankevekkende å se hvilke arbeidskår våre kolleger og studenter har i dette vårt store naboland. Det er usedvanlig vanskelig å komme inn som student ved prestisje-universitetet MGU; i tillegg til absolutte toppkarakterer fra skolen skal man gjennom en hard opptaksprøve, der under 20% slipper gjennom. Men så aner man at det skjer noe etter de første par år med disse i utgangspunktet toppbegavede, hardtarbeidende studenter. Arbeidsforholdene er dårlige, og undervisningen må også lide ved at de ansattes kår er enda semre. Da svikter vel motivasjonen ofte på begge sider av kateteret. Lærerlønnen er skral og ville ikke vært godtatt av en norsk 18-åring; prøv å lev på tusen kroner måneden når mange varer er like dyre som i Oslo (og kaffe og bleier dobbelt så dyre). Ting man tar for gitt hjemme glimrer kanskje med sitt vedvarende fravær. Mange har tilgang til datamaskin og e-post en gang i uken, på delt kontor. Selv tavler og kritt er av lav 1950-talls standard, og det var såvidt denne tyveetasjes eliteuniversitetsbygnings eneste overhead-prosjektor ble fremskaffet til mitt foredrag (som hadde den betimelige tittel «Forandrer verden seg?»). Lærerne ordner seg på ulike måter, etter evne og mulighet og nød. Mange gir undervisning på kveldstid for tenåringer av betalingsdyktige familier, for at disse skal komme inn de riktige steder senere, og noen klarer å ordne seg utenlandsopphold der man sparer på saltet og tar med pengene hjem. Og naturligvis blir det en trist «brain drain», der noen av de beste forlater det hele. Как жаль.

Likevel, til tross for kår som *kanskje* burde få noen norske streikende til å rødme, det blomstrer her og der. Det var en opplevelse å følge Nikolais interaktive undervisning av 15 hovedfagsstudenter i geostatistikk, inne på hans knøttlille kontor og på hans elendige tavle, klokken 18:30. Et par kolleger er av første skuffe innen jordskjelvsprediksjon, og en begeistret arbeidslunsj vedtok at vi måtte få

en doktorgradsstipendiat i kryssfeltet mellom musikk og statistikk til å jobbe med å lytte seg frem til naturens melodi via utbruddenes amplituder og posisjoner. (Her har jeg en hovedfagsoppgave.) Слава труду.

I desember var jeg en uke ved Erasmus Universiteit i Rotterdam, der jeg arbeider med Alex Koning, som har besøkt Blindern flere ganger. Det er interessant å se fellesskapet mellom økonometrikere og statistikere, der det også arbeides meget med modellbygging og sannsynlighetsteoretiske aspekter. Nobelprisvinneren Nikolaas Tinbergen henger i gangen og maner strengt til innsats.

Det er alltid interessant å sammenligne «sånn gjør vi det på Blindern» med regler, tradisjoner og nye innfallsvinkler annetsteds. Jeg tilbragte en regnfull uke i Glasgow i februar, og lærte for eksempel om karakterkortets tiltagende betydning for britiske universiteter – og da snakker jeg ikke om studentene, men om avdelingene og instituttene. Hver avdeling får sitt pass grundig påskrevet omtrent hvert fjerde år, på skalaen 2, 3, 4–, 4, 4+, 5, 5*. Konsekvensen av en dårlig karakter er at ditt universitet kan radere ut avdelingen du jobber ved, og hvis instituttet ditt gjør det riktig bra kan man bli belønnet med friske penger til flere stipendiater og kanskje en ekstra professorstilling. Noe for Blindern?

Og så bar det rett fra silkeføre i Nordmarka til 28 graders sommervarme Down Under, der vannet renner ut av vasken den gale veien og selv fotgjengerne heller mot venstre. Jeg hadde en hyggelig invitasjon fra Australian National University i Canberra og grep sjansen. Her tilbragte jeg mars og april på School of Mathematical Sciences, der statistikere, probabilister, renmatematikere og appliedniks lever om hverandre i en noe uoversiktlig blanding. Jeg nøler herved ikke med å karakterisere mitt arbeidsliv der nede som herlig – langt fra ståk og måk, tomt innbydende kontor, behagelige, befordrende og oppmuntrende omgivelser og mennesker, og naturligvis et sterkt innslag av det magisk-eksotiske: lenger vekk fra Blindern kan man da ikke komme. Kakaduene kurrer, evkalyptusen eimer og maurene maurer inne på campus, kenguruene hopper foran deg under morgenjoggen (vel, dette skjedde riktignok bare en gang), og papegøyene tiltaler deg i kontorvinduet.

Den bærende person i statistikkgruppen er naturligvis Peter Hall. Det er delvis ubegripelig at en person kan være så *flink* og allvitende og lynrask i tankesvingene, og kanskje enda mer ubegripelig at det går an å være så utenfor alle rimelige grenser *produktiv*. Plukk frem omtrent hvilken som helst årgang av hvilken som helst journal i teoretisk statistikk eller sannsynlighetsteori, og den har en artikkel av Peter. Det hviler en astralbølge av fattet, effektiv energi rundt ham, og et utstrakt følge av dr. gradsstudenter og post-doc-er, stipendiater og gjester fra mange verdensdeler duver i hans slep. Gjennom det hele er han mild og høflig, hjelpsom og beskjeden, men altså fantastisk effektiv. – For meg har det vært et frydefullt privilegium å arbeide sammen med og ved siden av denne Wonderul Wizard of Oz, med lærerike seanser med tavle og to kritt. Jeg knyttet også varige kontakter med andre av gjestene og er i gang med samarbeidsartikler.

Det ble også tid til noe ekstracampusalt liv, og jeg er riktig stolt over å ha vært med på den sydlige halvkules aller første fremførelse av Sjnittkes «Konsert for kammerkor», et bedrøvelig og intenst betagende tre kvarters verk. Kone og et stk. barn kom ned til en utvidet og uvanlig påsketur, der vi fikk med oss slanger i regnskogen, krokodiller på floden og snorkling på revene, og min tidligere student Manjula hilser fra Sydney.

Fra australske troper og høstfarger har jeg surfet hjemover til vår og sommer, og har ikke så mange reiser igjen i min bevegelige sabbatskalender: Madeira i mai, Lillesand i juni, Sankt-Peterburg i juli og Bath i august.

Jeg er en varm tilhenger av sabbatsordningen, på egne og andres vegne, og tror også at våre usynlige arbeidsgivere tjener på det. Jeg er meget tilfreds med mitt eget år, som dertil ble riktig billig for den norske stat, ettersom de fleste av mine nøkterne utgifter ble betalt av dem som inviterte meg. Og det er bare syv år til 2007.

# Greit nok med statistikk, men hvilke metoder er det som etterspørres?

Lars H. Vik
SINTEF, Økonomi og Logistikk

Jeg leste med interesse innleggene til Heuch og Frigessi angående statistikkfaget foran et nytt årtusen. Selv er jeg ikke statistiker, men har en bakgrunn fra økonometri. Jeg tror det gjør meg rimelig kvalifisert til å komme med noen kommenterer, men ut fra et brukerperspektiv. Det jeg driver med er oppdragsforskning for offentlig forvaltning og privat næringsliv, og i den sammenheng har man gjort seg noen personlige refleksjoner angående bruk av statistiske metoder (og da primært rettet mot ulike økonomiske problemstillinger). Forhåpentligvis er refleksjonene av interesse for statistikere av et mer edelt opphav.

Det første man oppdager når man anvender statistiske metoder er at den "statistiske leseferdigheten" hos oppdragsgiver er normalt sett veldig lav. Samtidig ønsker oppdragsgiver å vite hvordan man har kommet frem til svaret. I praksis medfører det at man som oftest benytter seg av svært enkle teknikker, teknikker som er *transparente* for brukeren. Teknikker som oppdragsgiver kan fatte intuitivt. Gode metoder for *visuell presentasjon* blir også sentrale. Mer avanserte metoder er noe man bruker på "bakrommet" for å etterprøve at de enkle teknikkene gir et plausibelt resultat.

Verden kan nok være kompleks, men store komplekse modeller er noe oppdragsgiver ofte finner det vanskelig å sette seg inn i. Det oppdragsgiver ønsker seg er *små* enkle (og dermed lett forståelige) modeller som fanger de essensielle aspektene. Det at oppdragsgiver forstår hva modellene fanger opp og hva de ikke fanger opp gjør det mulig for oppdragsgiver å bruke svarene til å fatte beslutninger. For oppdragsgiver er det helt greit at man ikke kommer til et helt presist resultat så lenge man forstår hvorfor. Forståelsen av metode og modell gjør at bruker kan overstyre de svar modellene gir når det trengs.

Som oftest er de data man får inn beheftet med grove feil og mangler. Det ha gjort meg til en ihuga tilhenger av *robuste* metoder og jeg er stadig på jakt etter nye teknikker for å avsløre svakheter i datasettet ("outliere", feilkoding og annet griseri). I tillegg er de datasett man får inn også ofte forbausende små. Jeg kan derfor ikke annet enn applaudere Heuch's fokus på teknikker for å analysere små datasett, og her må jeg innrømme at økonometrikere nok er noen av de verste synderene. Skremmende ofte er økonometriske metoder basert på asymptotiske argumenter.

Mine erfaringer kan oppsummeres i et slagord som jeg gjentar til kjedsommelighet, brukerene vil ha små robuste transparente modeller. Men så melder mitt store spørsmål seg: Hvilke metoder er i bruk? Hvilke metoder finner brukerene nyttige? Hvilke metoder er det brukerene virkelig ønsker seg?

Innenfor operasjonsanalyse har man gjort et omfattende arbeid med å kartlegge hvilke metoder som er i bruk [1], [2]. Undersøkelsene er nok ikke helt gode når det kommer til hvordan de er gjennomført, men de gir i det minste en indikasjon på hvilke metoder

som anvendes der ute i den store stygge verden. Det disse undersøkelsene viser er at statistikk, sannsynlighetsregning, lineær programmering og simulering er de metoder som konsekvent er hyppigst benyttet. *Tre av de hyppigst brukte metodene ligger innenfor statistikerenes domene*! Mitt store ønske er nå at noen undersøker hvilke sannsylighetsteoretiske modeller er det som brukes, hvilke simuleringsteknikker er det som anvendes, hvilke statistiske metoder er det som brukes. Jeg kan derfor ikke komme med annet enn en sterk oppfordring om at man gjennomfører tilsvarende spørreundersøkelser innenfor statistikk for å kartlegge hvilke metoder som ulike brukere har kjennskap til, hvilke som er i bruk til forskjellige formål, og hvor fornøyd brukerene er med de ulike metodene.

Innenfor ett felt har man undersøkt hvilke metoder som er i bruk, og det er innenfor utarbeidelse av prognoser [3]. Undersøkelsene tyder på at det brukeren kjenner til, anvender seg av, og er mest fornøyd med, de enkleste metodene. Men gjelder dette generelt for statistiske metoder? Ut fra mine erfaringer tror jeg det, men burde man ikke ha undersøkt dette nærmere?

La meg gå ett skritt videre. Anta at brukeren virkelig er på jakt, som jeg tror, etter små robuste transparente modeller. Når det gjelder robusthet har man kommet langt på teorifeltet, men metodenes kompleksitet gjør at de ofte kommer i strid med kravet til transparens. Robusthet kan man håndtere elegant matematisk, men hvilke metoder er det som er lettfattelige? Kan man designe eksperimenter som gjør at man, på empirisk grunnlag, kan avgjøre om en metode et lett å forstå (korrekt)? Er Tukey's trimean, som har gode robusthetsegenskaper, også lettfattelig som et deskriptivt mål? Er det slik at konfidensintervall er lettere å forstå enn Neyman-Pearson hypotesetesting? Er det slik at Jackknife estimatorer er lettforståelige? Jeg tror svaret er ja på alle spørsmålene (muligens med unntak av det første), men har noen undersøkt?

Den tålmodige leser har vel nå forstått hvor jeg vil hen. Jeg har noen hypoteser angående hva brukerene av statistikk (innen næringsliv og forvaltning) ønsker seg. Mitt håp er at noen kunne teste hypotesene. Svarene tror jeg vil være av nytte for statistikkfaget. Kunnskap om hvilke metoder som anvendes gjør at man vil kunne tilby utdannelser innenfor statistikk ut fra de ferdigheter som etterspørres, og dermed også bidra til å øke den "statistiske leseferdigheten". Å vite hva som brukes vil vel heller ikke være noen ulempe når det kommer til statistikk som skolefag. Kunnskap om hvilke egenskaper som brukerene ønsker seg ved de ulike metodene kan lede til interessante problemstillinger innen forskning. Statistikere burde her ha de beste forutsetninger for å foreta undersøkelsene jeg etterspør.

[1]. Harpell, John L., Lane, Michael S., Mansour Ali H.: Operations Research in Practice: A Longitudinal Study, Interfaces 19:3, May-June 1989, pp. 65-74

[2]. Morgan, Catherine L.: A Survey of MS/OR Surveys, Interfaces 19:6, November-December 1989, pp. 95-103

[3]. Sanders, Nada R., Manrodt, Karl B.: Forecasting Practices in US Corporations: Survey Results, Interfaces 24:2, March-April 1994, pp. 92-100

# *Referat fra debattmøte i Statistisk Forening Oslo 6.4.2000*

## Tema: Statistikkfagets sjel og perspektiver - lider faget av et rekrutterings- og kommunikasjonsproblem?

referent: Inge Helland, Universitetet i Oslo

Etter en invitasjon der det figurerte 7 foredragsholdere, alle med meningsfylte, men til dels ganske ambisiøse titler, fryktet noen av oss litt for hvordan denne kvelden ville forløpe. Vi ble gledelig overrasket. Det ble en meget interessant kveld med gode innspill og en debatt med tildels høy temperatur.

Bakgrunnen var delvis en bok av forskningsjournalist Bjørn Vassnes (`Myten om fornuften: og den nye vitenskapen om oss selv, Oslo - Pax, 1997.') som SFO-formann Ulf Indahl i sin tid hadde fått i gave til disputasen sin. Vassnes var invitert som hovedforedragsholder.

Ulf ga først en innledning der han motiverte temaet for møtet og blant annet snakket om rekrutteringsproblematikken og om uoppdagete faglige slektsbånd.

Vassnes kåserte om statistikkfagets posisjon blant utenforstående, og begynte med å si at det var mange av disse som ikke en gang kjente faget, og enda flere som var temmelig ukjente med hva det egentlig sto for. Arne Scheie ble nevnt som en typisk statistiker blant lekfolk. Vassnes snakket også endel om relasjoner til andre fag som kvanteteori, partikkelfysikk og biologi. Den genetiske evolusjonen har fungert statistisk, og dette har gitt en menneskelig hjerne som er en effektiv beslutningstaker. Læringsprosessen hos det enkelte mennesket fungerer også statistisk, og dette har ført til at også ulogisk problemløsning kan være effektiv. Kategorisering, og det å finne mønstere er viktig.

Blant utfordringene til statistikerne med sikte på å gjøre faget kjent, er å arbeide aktivt for at pressen skulle bli mer edruelig i statistiske spørsmål.

Tormod Næs hadde tittel `Yrkeslivets behov', og brukte det meste av tiden til å argumentere for at statistikeren bør beherske hele den prosessen som en statistisk undersøkelse svarer til: Problemformulering, design, datainnsamling, modellering, analyse, beregninger, tolkning, og eventuelt en ny runde med alt dette. Alt dette skal beherskes filosofisk, metodisk, praktisk og kommunikasjonsmessig. Utdanningen de fleste steder legger for entydig vekt på analyse-fasen. Metodeutvikling som sådan er utviklet til det ekstreme, og de matematiske aspektene står ikke i forhold til de praktiske problemenes grovhet.

Odd Aalen sa i diskusjonen at til tross for kritiske bemerkninger av denne typen, måtte statistikk regnes som et vellykket fag.

Arnoldo Frigessi sa at det er lett å bedømme matematikk som unyttig - etterpå. Men det er også mange eksempler på at matematikk som mange opprinnelig så på som unyttig, senere har viste seg å gi meget nyttig metodikk. Eksempler fra telekommunikasjon ble nevnt. For noen problemer er det ikke nok med enkle metoder. Han diskuterte også kritisk myten om at vitenskapen nødvendigvis beveger seg framover.

Tore Schweder fremhevet i diskusjonen at det tar tid å lære matematikk, og han understreket faren for avmatematisering.

Erik Bølviken var opptatt av IT-aspektet ved statistiker-utdanningen. Han brukte uttrykket `lost generation' om dem blant oss som ikke hadde hatt tilstrekkelig informasjons-teknologi i utdanningen sin. Ellers var han som Tormod kritisk til matematiseringen av statistikk-faget. Yrkesmessig er det etterspørsel etter kombinasjonen statistikk/IT.

Arnoldo mente at statistikk-utdanningen ved norske universiteter var god.

André Teigland snakket om statistikeren som `den ydmyke tjeneren', og mente med dette at det er viktig å løse problemer på brukerens premisser. Dette må på ingen måte ses på som mindreverdig. Han understreket også hovedpunktet til Tormod, nemlig viktigheten av å beherske hele prosessen i en undersøkelse.

Tore Schweder mente at rekruttering ikke er et IT-problem. Han var meget kritisk til mye av det som lanseres under datamining-paraplyen. Ellers gikk han inn på flere hovedpunkter i fagets historie, med en start i Bradley Efron's spissformulering om at faget i forrige århundre hadde utviklet seg `from a small obscure field to a big obscure field.' Tore nevnte tre paradigmeskifter: Laplace (1774; Bayes metoder), Gauss (1809-12; minste kvadraters metoder) og Fisher (1912; maximum likelihood).

Tobias Dahl kom med to forslag i diskusjonen: 1. Konkrete utspill mot misbruk av statistikk i avisene. 2. Nye eksamensformer knyttet til presentasjon av datasett.

Til slutt ga Ulf en demonstrasjon av dataminingsprogrammet Clementine, der et hovedpoeng er visualisering. Etterpå sa bl. a. Geir Storvik at han var blitt skuffet over dataminings-begrepet.

Odd minnet om at det finnes en yrkesrettledningsside for statistikere på internett.

Tore siterte Rao om at `Statistics is the technology of the 20$^{th}$ century.'

Altså i alt et meget rikt og mangesidig debatt-møte. Men hjalp møtet med å finne noe i retning `statistikk-fagets sjel'? Helt sikkert litt for en del av oss. La meg avslutte med følgende personlige bemerkninger; de to første punktene inspirert av at `sjel' i utgangspunktet er et religiøst begrep:

1. For å ta et ekstremt utgangspunkt: Vår lokale verdensreligion ble grunnlagt av en mann som ofte kalte seg selv tjener, og oppfordret andre til å tjene. I vår egen tid har viktige personer, som Mahatma Gandhi og Nelson Mandela, lagt vinn på å følge samme tankegang. Uten at noen av dem har lagt seg flate for eventuelle motstandere. I denne sammenhengen er André's bilde av statistikeren som hovmester absolutt i meget godt selskap. Og jeg er helt enig i hans presiseringer. Men for mange av oss kan kanskje den ideale fordring her være i tøffeste laget.

2. Den samme verdensreligion er for mange blitt ødelagt av alle de ulike sekter og kirkesamfunn som fokuserer altfor sterkt på egne spesifikke dogmer (les: meninger/kultur). Overført til vår sammenheng, blir spørsmålet da: Kan vi klare å fortsette en god utvikling uten å være for dogmatiske i noen retning? Etter mitt syn bør dette være mulig, og er også ønskelig

både for undervisningen, for leting etter løsninger på rekrutteringsproblemene og for forskningen, som også er en letingsprosess. Og som kjent, når man leter etter noe, kan det ofte være lurt å se flere steder.

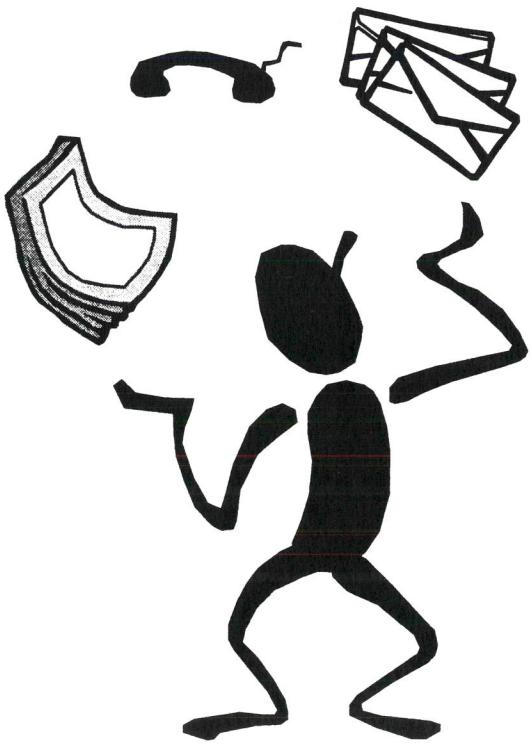
3. Tormod's krav om at man skal lære seg å beherske hele prosessen i en statistisk undersøkelse er riktig og viktig. Det finnes gyldige motforestillinger mot å la dette få en for dominerende plass i universitetsstudiene, f.eks: (i) Noe av dette læres best etter studietiden, (ii) Nåværende universitetslærere har ikke kompetanse til å undervise i alle disse aspektene. Ikke desto mindre vil det være riktig å ta dette punktet opp til en grundig diskusjon i forbindelse med statistikk-undervisningen.

4. Teoretisk statistikk har og bør også ha flere sider enn det å se etter gode metoder.

Takk for en meget fin kveld, Ulf!

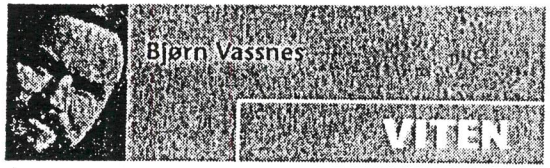
PS 1.
Tore Schweder sendte etter møtet ut en mail der han viser til Jonas Ellenberg's helt ferske artikkel `Statistician's Significance' i JASA **95**, 1-8 (2000). Artikkelen handler om i hvilken grad statistikeren får tilfredsstilt sitt behov for `recognition' - anerkjennelse, et legitimt behov som er viktig for oss alle sammen.

PS 2.
Bjørn Vassnes fulgte opp temaet for møtet med en kronikk i Morgenbladet 14.4.00: `Statistikkfaget fortjener et bedre rykte'. Her går han også inn på fagets historie: Quetelet og normalkurven, Galton og korrelasjonsbegrepet; foruten at han på flere områder argumenterer for at statistikk er viktig som metode for moderne vitenskap.



Morgenbladet, 14. april 2000

# Statistikkfaget fortjener et bedre rykte




De fleste av oss forbinder vel *statistikk* først og fremst med Arne Scheie, og hans analyser av hvorfor det er Manchester Uniteds tur til å vinne akkurat i dag. Eller med undersøkelser referert i dagspressen som forteller at finnmarkinger er kåtest og sogninger lykkeligst her i landet. Det er derfor ikke merkelig at statistikkfaget har en status på linje med falskmynteri. Mye av dette er selvfølgelig urettferdig (Arne Scheie, f.eks., har neppe så mye som et grunnfag i statistikk). Men vi ser heller ikke altfor ofte at statistikere går ut mot den vettløse bruken av «statistikk» i media, ikke minst i forbindelse med meningsundersøkelser.

Samtidig har det aldri vært så mye bruk for statistikernes kunnskap som akkurat i dag. Grunnen er de enorme datamengdene som produseres overalt i samfunnet, ikke minst innen forskningen, og som bare venter på å bli tolket seriøst. Her ligger mye potensiell kunnskap om svært viktige saksforhold, men uten mennesker som kan tolke disse dataene er vi like langt. Eller enda verre: da kan dataene brukes til å fabrikkere løgner.

Statistikkfaget har, som navnet forteller, noe med fremveksten av den moderne staten å gjøre. Det var de nye europeiske statsbyråkratiene som fra 1700-tallet begynte å samle inn data av ymse slag om sine borgere. Preusserne var først ute, men etter Napoleonskrigene ble dette vanlig i mange land. Og det var spesielt data som gikk på forbrytelser, selvmord og andre lite hyggelige ting. Tanken var at for å ha kontroll over disse unormalitetene måtte man først tallfeste dem.

Noen oppdaget så at disse tingene viste en merkelig regularitet. Antallet forbrytelser, f.eks., var ofte svært likt fra år til år. Noen mente at man her var på sporet av samfunnsmessige lover, på linje med de fysiske naturlovene. Man så her en mulighet til endelig å utvikle en samfunnsvitenskap som var like «vitenskapelig» som fysikken.

**En av de** fremste pionerene innen statistikken var Adolphe Quetelet, som hadde sin bakgrunn i astronomien (kongelig astronom i Belgia). Og det var faktisk fra astronomien at Quetelet hentet verktøyet som skulle bli så viktig i statistikken: det vi kaller «normalkurven». Dette var egentlig en kurve astronomene brukte til å justere for feilobservasjoner. Astronomene hadde lagt merke til at dersom de foretok et større antall observasjoner av et fenomen, og satte opp disse på en liste, ville resultatene forme en klokkeliknende kurve med en topp i midten. Denne spredningen (deviasjonen) skyldtes rett og slett den menneskelige tendensen til å gjøre feil.

Det Quetelet oppdaget var at han fant den samme kurven i statistikker over menneskelige og sosiale forhold, som f.eks. høyde, vekt, eller antall barn. Ut fra dette utviklet han en idé om det «normale» menneske og det «normale» samfunn. De som befant seg i kurvens ytterpunkter ble «unormale».

Det som hadde vært et praktisk hjelpemiddel for astronomene for å bestemme hva som var korrekte observasjoner, ble nå en måte å normere befolkningen på. Og «normalfordelingen» ble en naturlov: Det ble slik at det *skulle* være et bestemt antall selvmord eller drap pr. år, ellers var det «mot normalt». Denne deterministiske måten å bruke samfunnsstatistikk ble utviklet av andre statistikere på 1800-tallet, som briten Francis Galton. Han var spesielt glad i *korrelasjoner*, statistiske sammenhenger, som langt på vei kom til å erstatte behovet for årsaksforklaringer.

Denne forestillingen om at det er nok å vise tallmessige sammenhenger er en av grunnene til at statistikk kan fremstå mer som astrologi enn som seriøs vitenskap. Odd Børretzen har et godt eksempel på en slik korrelasjon: Av alle dem som omkommer i flyulykker, har et flertall drukket kaffe like før ulykken. Korrelasjon, ja – men årsak?

**Korrelasjoner** *kan* imidlertid signalisere årsakssammenheng, og *noen* regulariteter har nok en tendens til å holde seg over noen år. Derfor har myndighetene utvilsomt hatt nytte av statistikerne. Men fordi menneskelig adferd styres av flere forhold enn noen kan ha oversikt over, vil samfunnsstatistikk alltid være misvisende, i større eller mindre grad. Og i mange tilfelle: fullstendig meningsløs.

> Det har aldri vært så mye bruk for statistikk som i dag, men uten mennesker til å tolke den er vi like langt.

Men de statistiske metodene ble også tatt i bruk i statistikkens opprinnelige «hjemland», naturvitenskapen, og med stor suksess. Et typisk eksempel er partikkelfysikken, der man ut fra et stort antall eksperimentresultater bruker statistiske metoder til å finne betydningsfulle mønstre. I partikkelfysikken er jo usikkerheten innebygget – en partikkel er egentlig *sannsynligheten* for en slik partikkel på et gitt sted og en gitt tid – og virkeligheten er på dette nivået nødvendigvis av statistisk karakter.

I dag har vi imidlertid en trend der kunnskapen innenfor de fleste naturvitenskapelige fag blir mer og mer «statistiske». Klimaforskning er et typisk

forts. neste side

eksempel: Også her dreier det seg om å tolke en mengde data, fra ulike målemetoder, og ut fra disse ofte sprikende resultatene å finne hva som er «signifikante signaler» (om f.eks. temperaturutviklingen de siste tusen år). Eller med andre ord: Å finne betydningsfulle mønstre i kaoset av informasjon.

Et annet felt der statistikken nå kommer inn for fullt, er i såkalt «bioinformatikk» (se notis), sammensmeltingen av biologi og computer science. Her tenker jeg f.eks. analyser av DNA og proteiner, i begge tilfelle ufattelig komplekse molekyler, og der man må kjøre mange tester for å være sikre. Nytt verktøy har skapt en enorm dataflom, som man må prøve å finne mønstre i.

Man kan faktisk si at den viktigste nye kunnskapen i årene fremover vil komme nettopp fra analyse av store datamengder, ved hjelp av ulike metoder, men ikke minst statistikk. Det er derfor et tankekors at faget har så lav status som det har, spesielt i Norge, der rekrutteringen nå er nede på et lavmål. Vi er allerede begynt å importere indere til å jobbe med programvare, må vi begynne å importere kinesere til å jobbe med statistikk?

Eller må vi gjøre noe med Arne Scheie?



## Biologi + datamaskiner = penger

De to store vekstnæringene for tiden, både innen forskning og næringsliv, er informatikk og bioteknologi. Og spesielt eksplosiv blir veksten når disse to kobles sammen. Det kanskje heteste eksemplet er Craig Venters Celera Genomics, bioteknologifirmaet som nå truer med å ta innersvingen på det store internasjonale HUGO-prosjektet i kartleggingen av det menneskelige genomenet.

Celera har tidligere i år lagt frem genomet til bananfluen, *Drosophila melanogaster*, men nå er det noe mer langt mer (20 ganger mer) komplisert de har gjennomført. Nå nylig (6. april) annonserte Venter at hans firma faktisk har kartlagt alle baseparene (bokstavene) i det menneskelige genomet, med en metode Venter har funnet opp. Det betyr ikke at «Livets Bok» er ferdig, fordi rekkefølgen på disse bokstavene ennå ikke er klarlagt. Venters metode kartlegger nemlig sekvenser av DNA-et, men uten at man vet hvor disse sekvensene hører hjemme. Dette må man finne frem til ved å studere hvor sekvensene overlapper hverandre, et arbeid som krever mye dataprossesering.

Selv om andre genforskere forholdt seg skeptiske til Celeras annonsering, økte børsverdien til selskapet med 19 prosent samme dag som dette ble annonsert. Og noen mener det var nettopp dette som var hensikten, og at dette signaliserer en bekymringsfull trend i den bioteknologiske forskningen, henimot mer kommersialisering og hemmelighold..

Etter det menneskelige genomet venter en rekke andre genomer, på dyr og planter vi er avhengige av, og bakterier og virus som vi prøver å bekjempe. Alt dette vil kreve enorme ressurser innen bioinformatikk. Og da har vi ikke engang nevnt det NESTE store arbeidsfeltet: å kartlegge og konstruere *proteiner*, «proteomics». Dersom du vil sikre dine barn arbeid i fremtiden kan du utdanne dem til bioinformatikere. Problemet er at de da kanskje må reise utenlands.

Aftenposten
26. april 2000

## Donald i 100 000

Gamle Donald-blader i innbundet utgave selger som bare det. Rett før jul 1998 kom den første boken, med blad fra 1948 og 49-årgangen. Med den sjette utgaven, som inneholder siste halvdel av 1952-årgangen, har salget passert 100 000 bøker totalt. De aller første Donald-bladene ble for noen år tilbake utgitt på ny, parallelt med de vanlige ukebladene. Salget av de innbundne bøkene dokumenterer igjen nordmenns unike forhold til den berømte anden fra Andeby. Det er matematikkprofessor Nils Lid Hjort som i siste bok har skrevet forordet. Tidligere har historiker og Donald-ekspert Øystein Sørensen og sosialantropolog Thomas Hylland Eriksen skrevet forord til årgangsbøkene.

## Om statistikeres image i norske aviser

Etter å ha lest "Statisticians' Significance" av J. H. Ellenberg, JASA, March 2000, 1–8, ble undertegnede inspirert til å gjennomføre et nett-søk på ordet "statistiker" i et utvalg av Norges største aviser. Under følger en oversikt over resultatet, samt de utsagn som på et eller annet vis beskriver "statistikere" som gruppe. Resultatet er *ikke* spesielt oppløftende, og dersom denne lille useriøse undersøkelsen kan sies å avspeile tendenser, har vi (ikke uventet) et image problem.

**VG** Ingen treff.

**Dagbladet** 4 treff. Disse er henholdsvis en kronikk om Geir Helljesen, en artikkel om værstatistikere, et innlegg fra Nils L. Hjort fra 1996 om skøyter (500 meter), samt en kronikk av Kristin Clemet. Kronikken om Geir Helljesen i inneholdt frasen "Fra å være en relativt tørr statistiker har han med årene utviklet seg mot entertaineren." I Clemets kronikk er ikke omtalen mer rosende: "Men tar du gjennomsnittet fra en statistiker, tar du lykken fra ham med det samme, og han havner i rennestenen."

**Aftenposten** 7 treff (fra og med 1995). Disse treffene omfatter 3 artikler fra statistiske undersøkelser (misdannelser hos nyfødte i Rana, røyking og sammenheng mellom foreldres yrker og misdannelser hos barn), to med statistikk brukt i sammenheng med sport (jmf. Bjørn Vassnes og Arne Scheie), samt en artikkel som beskriver utdeling av en pris til Helge Seip (i denne sammenhengen "statistikeren"). I en leder fra 7. juni 1998, "Angst, statistikk og bekymringer" finner vi følgende utsagn: "Det gjelder, som i så mange andre sammenhenger, å finne en fornuftig balanse, og å huske at det ikke er alt som kan eller bør illustreres med statistikk. Det burde statistikerne, de av dem som tror at det aller meste kan og bør måles, bekymre seg om."

**Dagsavisen** 6 treff. Treffene i denne avisen omhandler borger nummer 500.000 i Oslo, røyking, likestilling, jubilant Grevstad, samt et innlegg med tittel "Statistikk og klar erindring". Sistnevnte er den eneste med subjetiv vurdering av statistikere, ei heller her særlig flatterende: "Det rare er kanskje ikke at tiden går, men at vi har så stort behov for å markere det. Er det bare statistikere som trenger nyttår for å oppsummere, eller er det en naturlig holdeplass for ettertanke også for folk privat?"

Xeni K. Dimakos

# Er tiden moden for nytenkende og mer kreativ rekruttering til Statistikkfaget?

Andelen studenter på begynnerkurset i matematikk (MA 100) ved Universitetet i Oslo har siden midten av 90-tallet falt fra drøye 600 til 375. Trenden for ST 101 – begynnerkurset i statistikk – er enda mer fryktinngytende. Under 40 studenter møtte til eksamen i '98 og '99 mot ca 100 midt på 90-tallet. De nye studentkullene ankommer dessuten universitetet med svakere bakgrunnskunnskaper i matematikk enn tidligere.

Den videregående skolen makter med andre ord i enda mindre grad enn tidligere å produsere universitetets etterspørsel etter studenter som besitter et tilstrekkelig kompetansegrunnlag innen matematikk. Det vi opplever er blant annet resultatet av et offentlig skolesystem som gjennom en årrekke har utviklet seg slik at dyktige realister bare unntaksvis finner det attraktivt å ta seg jobb som lærer.

Dette står i sterk kontrast til behovene i moderne industri, forskning og næringsliv som internasjonalt mer enn noen gang etterspør kompetanse i forbindelse med strategisk arbeide og beslutningsprosesser. Informasjons-grunnlagene det arbeides med og kravet til presisjonsnivå er ofte av en slik art at statistisk kompetanse og analysemetodikk med fordel bør benyttes.

For en person med bakgrunn som statistiker eller matematiker er det innlysende at lærerne i den videregående skolen må beherske de faglige forventningene elevene møtes med som studenter. Pussig nok er det ikke like åpenbart i fagmiljøet ved universitetet, at man også der bør holde seg orientert/oppdatert om utfordringene og mulighetene i det arbeidsmarkedet som møter studentene etter endt utdannelse. Dagens praksis harmonerer ikke spesielt godt med forfektelse av en statistikerimage som "ydmyk tjener" for potensielle brukere/kunder.

Man skal ikke se bort ifra at fagets "dårlige tider" også kan skyldes slett engasjement i egne rekker. Det er neppe noen tvil om at et hvilket som helst fag blir vanskelig å selge både til studenter og i arbeidsmarkedet dersom fagets innhold, nytteverdi og de naturlige tverrfaglige relasjonene ikke formidles på en tydelig måte. Begrepet "statistikk" vekker blant folk flest verken begeistring eller representative assosiasjoner. Det forteller eksempelvis lite om at framtidas statistikere i stor utstrekning kommer til å jobbe med en mengde integrerte IT-baserte analyseverktøy, og at kravet om informasjonskvalitet vil medføre et sterkt behov for koordinering og samarbeide mellom database- og analysekompetanse.

Forrige sommer var Trevor Hastie, professor i statistikk fra Stanford University, på 3 ukers faglig opphold i Oslo. Invitasjonen ble initiert av Statistisk Forening i Oslo (SFO), oppholdet i hovedsak finansiert av Institutt for Informatikk (IFI), Biologisk Institutt og SFO, med mindre bidrag fra MATFORSK og avdeling for statistikk (Avdeling C) ved universitetet. Det "tverrfaglige" samarbeidet var nødvendig for å realisere invitasjonen (og røper litt av appellen en statistiker av Hasties kaliber har selv i "lille Oslo"). Ved Stanford dekkes lønna til de ansatte for kun 10 av årets måneder. Dette sporer de ansatte til å oppsøke nye miljøer (inkl. næringslivet) og bli kjent med de behov og problemstillinger som råder der.

Mangel på denne typen erfaringer snyter vårt hjemlige statistikkmiljø for viktig praktisk kunnskap og nyttige tverrfaglige synergier med beslektede fagområder. Man kan spørre seg om det er mulig å tilby et troverdig & attraktivt studietilbud, for studenter som ikke planlegger en forskerkarriere, uten at slik innsikt og erfaring er solid forankret i det faglige miljøet? Dessuten vil en forankring i praktisk innsikt og erfaringer trolig også fungere utmerket til å opprettholde en stabil intern akademisk rekruttering.

Vår hjemlige akademiske kompetanse innen statistikkfaget er uten tvil høy nok, men bredden er mangelfull. Mye av statistikkfagets framtidige suksess vil trolig bero på hvor dyktig man er til å etablere nødvendig bredde ved inngåelse av tverrfaglige allianser. Nytten av tverrfaglige langtidsprosjekter også internt ved universitetene bør ikke undervurderes. Slike allianser vil gi fagmiljøene verdifull erfaring i hvordan statistikere og øvrige fagpersoner kan og bør samarbeide. Per i dag er man som statistiker nærmest identitetsløs utenfor egne rekker. Jobben med å gi statistikkfaget en god og tidsaktuell definisjon er ekstremt viktig, men vil trolig forutsette de faglige kontrastene man kun finner ved å ta utfordringen om seriøse tverrfaglige samarbeidsprosjekter på alvor.

Ulf Indahl
Leder Statistisk Forening i Oslo

# Yrkeslivets behov: Et bidrag til diskusjonen om statistikkfaget

Tormod Næs
MATFORSK, Ås og Universitetet i Oslo, Blinden.

## Innledning

Artikkelen er en kort sammenfatning av hovedmomentene i undertegnedes innlegg på diskusjonsmøtet om statistikkfaget den 6. april i regi av Statistisk forening i Oslo.

Først vil jeg si at jeg er svært glad for at denne diskusjonen ble tatt opp i Statistisk forening. Ethvert seriøst fag som er relevant for både næringsliv og forskning, bør med visse mellomrom ha grunnleggende fagdiskusjoner. Timingen var også meget god i dette tilfellet fordi

- faget opplever i dag et betydelig rekrutteringsproblem
- statistikerne er i for stor grad en marginalisert gruppe
- de praktiske behovene for statistisk kompetanse i næringslivet er svært store

Det første punktet er et faktum, det siste er de fleste innen faget trolig enige om, mens det midterste er kanskje mer kontroversielt. Min erfaring tilsier at det er rett.

## En viktig premiss

Min tenkning om faget statistikk tar alltid utgangspunkt i følgende premiss:

- Statistikkfaget er primært et hjelpefag for andre fag

"Alle" er interessert i medisin, økonomi etc. fordi det angår dem, men det er bare statistikerne som er genuint interessert i *faget* statistikk. Faget har derfor sin vesentlige interesse og samfunnsmessige betydning først når det kobles opp mot konkrete problemer og anvendelser.

I det følgende vil jeg diskutere noen momenter som kan belyse fagets situasjon i dag, nevne noen av de problemene jeg mener vi sliter med og avslutningsvis komme med noen forslag til dreining av fagets profil og praksis.

## Statistikkfaget bør være en integrert del av *hele* forskningsprosessen

Hvis vi tenker oss en utviklingsprosess eller et prosjekt der man ønsker å undersøke et problem, enten det er et forskningsproblem eller et mer dagligdags problem i en bedrift, vil det være typisk å gå gjennom følgende faser:

- Problemformulering/operasjonalisering
- Design/forsøkplanlegging
- Datainnsamling/kvaltets-sjekking/korrigering av feil
- Modellering/modellsjekking
- Statistisk analyse, finne gode estimatorer etc.

- Beregninger/simuleringer
- Tolkning/Validering
- Ny runde?

I praksis vil rekkefølgen ikke alltid være så stringent som her, men de fleste av disse punktene vil inngå på en eller annen måte.

Jeg mener at alle som utdannes og arbeider innen faget statistikk bør beherske de sentrale delene av denne forskningsprosessen ikke bare metodisk, men også filosofisk og praktisk. Vi bør for eksempel ikke sende studenter ut i den virkelige verden uten at de har fått solide kunnskaper om hvilke elementer som inngår i en forskningsprosess og hvordan de er relatert til hverandre og omverdenen.

Min erfaring tilsier at vi, som profesjon, har svært mye å bidra med i alle fasene av forskningsprosessen. Til og med i første og siste fase, problemformulering og beslutning om ny runde, kan vi spille en avgjørende rolle. Det at vi er et hjelpefag for andre fag, behøver slett ikke bety at vi "setter oss i baksetet" og venter på at ting skal skje. Tvert imot, som regel kan vi også bidra betydelig i førersetet. Særlig der statistikeren samarbeider med andre fagfolk over lengre tid, vil det være mulig å integrere statistisk tenkning i alle fasene av prosessen.

**Statistikere er for opptatt av egne perspektiv og matematiske aspekter**

Mitt klare inntrykk er imidlertid at vi både i undervisning, forskning og praksis har lagt særlig stor vekt kun på et par av disse punktene. Det gjelder særlig matematisk og statistisk analyse av gitte modeller. De fleste statistikerne har i løpet av sin utdanning blitt kjent med et utall modeller og kriterier og lært å utvikle optimale estimatorer og tester. Dette er viktig, men i mange praktiske sammenhenger ikke det viktigste. Etter min vurdering har denne delen av prosessen vært overfokusert på bekostning av de andre elementene.

Mange av de matematiske aspektene og metodene som er utviklet står ofte ikke i et naturlig forhold til grovheten og kompleksiteten som finnes i den praktiske verden. Jeg er redd en god del av disse avanserte metodene lever og diskuteres en kort stund av en liten gruppe forskere, for så å bli relativt fort glemt. Det er lett å glemme prinsippet om enkelhet (parsimony), som sier at blant flere alternative tilnærminger eller forklaringer, velger man alltid den enkleste og mest forståelige.

Nå er det slik at man ikke kan forutsi eksakt hvilke metoder som vil bli viktige/nyttige i framtiden, derfor er det nødvendig å opprettholde teoretisk statistisk forskning. Mitt poeng er imidlertid at denne del av faget er blitt overfokusert på bekostning av en bredere orientering.

Konsekvensen av et slikt innadvendt og spesifikt fokus har etter min mening bidratt til frustrasjon blant brukere, marginalisering av profesjonen og langt fra full utnyttelse av fagets muligheter.

**Statistikere er ofte for beskjedne og forsiktige**

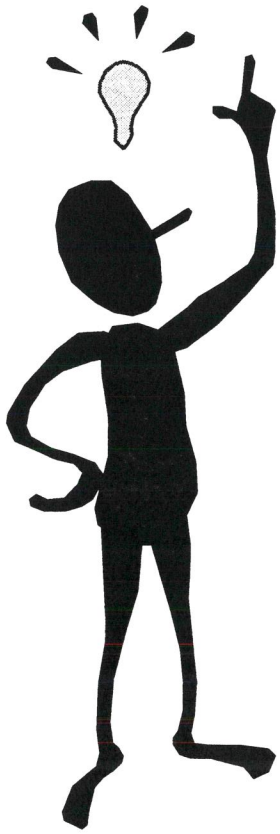
Jeg tror at sterkt fokus på et begrenset område også har vært med på å utforme en kultur som har vært for forsiktig og for lite offensiv.

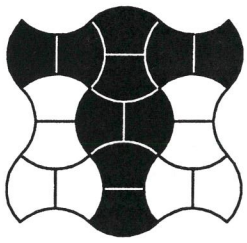
For eksempel mener jeg at vi, som gruppe, er for sterkt preget av det jeg kaller etterrasjonalisering. Vi er for opptatthet av å vite alt om metodenes riktighet og egenskaper framfor å ligge i forkant av en utvikling. Kanskje har vi for lavt "signifikansnivå" på en del av våre aktiviteter? Vi gjør nok få feil (type I), men vi utnytter heller ikke alltid fagets potensial og muligheter (type II). Vi må ikke sitte bak en brevsprekk og vente på oppgaver, regne litt på dem og levere noen "signifikanser" tilbake gjennom den samme sprekken. Slik blir det opp til andre å gjøre de viktige valgene.

Vi er etter min mening heller ikke offensive nok med hensyn til synliggjøring av fagets muligheter og potensiale overfor yrkeslivet og i rekrutteringsøyemed. Vi må prøve å gjøre statistikkfaget mer attraktivt. Dette kan blant annet gjøres ved å fremheve suksesshistorier, vise fram fagets bredde slik den er beskrevet over, vise fram yrkesmuligheter etc.

**Behov for grunnlagsarbeid**

Det er viktig å merke seg at synspunktene over ikke primært er en argumentasjon langs aksen teori-anvendelser. Det er snarere en argumentasjon for en dreining av fagets fokus - *fra* en fokusering utover det interessante på noen få punkter *til* en bredere orientering mot hele forskningsprosessen, fra problemformulering til tolkning av resultater. Jeg er helt sikker på at svært mange av disse elementene har grunnleggende aspekter som kan og bør videreutvikles, både innen forskning, undervisning og praksis. Her er mye upløyd mark.





***Teknometri AS** er et heleid datterselskap av Software Innovation (SI). SI er ledende CRM-leverandør i Skandinavia. Selskapet leverer tjenester innen dataanalyse, datavarehus og informasjonsteknologi generelt. Teknometri besitter forøvrig bred kompetanse på å bearbeide, analysere og foredle informasjonen som ligger lagret i interne eller eksterne datakilder. I tillegg tilbys næringslivet opplæring i bruk av statistisk verktøy og analysemetodikk. Analysemiljøet i Teknometri består av 4 erfarne heltidsanalytikere og 1 nyutdannet. Alle er på Dr.Scient. & Cand.Scient. nivå innen statistikk/matematikk. Selskapet ble stiftet i 1994 og har totalt 25 ansatte.*

# Datavarehus og datamining – hukommelse og intelligens?

*Av Knut Opdal og Ulf Indahl - Teknometri AS*

*Den trauste statistiker var fortrolig med å få overlevert sitt datasett på flat fil som ble lest inn i passende software, med overlevering av analyserapporten tidligst 14 dager senere. Framtida blir ganske annerledes. Det er allerede i dag mulig å gi en hel organisasjon effektiv hukommelse og intelligens ved hjelp av moderne informasjonsteknologi og analysemetodikk. Dette vil for mange bedrifter kunne bidra til vesentlig forenkling innen administrasjon, optimalisering av forretningsvirksomheten samt gjøre bedriften mer robust overfor endringer i bemanningen. I dag er det rimelig å anslå at det er et titalls norske bedrifter med "fullblods" datavarehusløsninger installert. Det er imidlertid svært realistisk å anta at flertallet av mellomstore og store norske bedrifter vil ha etablert datavarehusløsninger i løpet av de nærmeste 3-4 år. Statistikere som skal spille en rolle i denne utviklingen behøver ikke nødvendigvis noen rehabilitering, men de må trolig igjennom en viss "oppussing" hva perspektivene rundt egen kompetanse angår.*

## Datavarehus – bedriftens kollektive hukommelse

Et godt utgangspunkt for å forstå tankegangen bak datavarehus konseptet er å ta for seg mennesket selv. Som beslutningstager, leder og forretningsmessig innovatør er det viktig å:

(1) *vite/huske* (lagre informasjon om) så vel tidligere erfaringer som hva det som til enhver tid foregår.
(2) *bearbeide* og *analysere* erfaringer og fakta (lagret informasjon) for å *tolke* og *forstå* relevante sammenhenger.

Ved å gjenkjenne og forutse situasjoner hvor de sammenhenger opptrer vil kunnskapen man besitter bli verdifull. Man har innsikt om hvordan forskjellige handlinger eller tiltak vil virke. Dette er ikke noe annet enn en skjematisk beskrivelse av de læringsprosesser som foregår kontinuerlig i den menneskelige hjerne.

*En god leder på overordnet nivå er per definisjon en dyktig forvalter av så vel egen som bedriftens hukommelse og intellekt!*

Det beste beslutningsgrunnlaget for en gitt problemstilling får man ved å ha mest mulig *relevant* informasjon tilgjengelig. Fordi mer og mer informasjon registreres og lagres elektronisk, har man i mange sammenhenger store mengder informasjon knyttet til en aktuell problemstilling. Så vel interne som eksterne datakilder kan være aktuelle. Det er imidlertid urimelig å forvente at en ledelse skal kunne ha oversikt over all relevant informasjon og langt mindre kunne analysere og bearbeide informasjonen manuelt. Det er rett og slett praktiske begrensninger for hvor mye informasjon som kan lagres og analyseres i en menneskelig hjerne. Man trenger hjelp til å trekke ut struktur og essens.

Behovet for struktur og organisering av informasjonen blir viktig med voksende datamengder. Det en leder/beslutningstager behøver er *beslutningsrelevant* informasjon. Informasjonen må nødvendigvis tilrettelegges og tilpasses lederens behov, de verktøy/assistenter han har tilgjengelig og den kompetanse som besittes.

Lettere tilgang til beslutningsrelevant informasjon er årsaken til at flere og flere bedrifter i dag strukturerer, organiserer og bearbeider sine data i et *datavarehus*. Hvis man planlegger på riktig måte vil datavarehuset kunne spille en viktig rolle som bedriftens *hukommelse*. Datavarehus og ikke minst intelligent bruk av dem vil derfor bli et viktig konkurransemoment for mange bedrifter i tiden fremover.

### Datamining/CRM/dataanalyse – bedriftens kollektive intelligens

Det som står mest sentralt i datavarehuset er bedriftens beslutningsrelevante styringsparametere med tilhørende nøkkeltall. Et nøkkeltall kan være et sammendrag av flere viktige informasjonskilder og aggregert fra ulike nivåer. Fordi bedriften styres etter disse parametrene, må de være lett tilgjengelige. En av datavarehusets viktigste fordeler at det kan levere *ferske* nøkkeltall når beslutningstageren har behov for dem.

Hvis for eksempel ledelsen i en bredrift ønsker å benytte seg av *balansert målstyring* kan det bygge opp et såkalt "balanced scorecard". Dette er et styringsverktøy som, når det er etablert, forenkler ledelsens arbeid med å styre og holde øye med bedriftens kostnadsdrivende og inntektsbringende aspekter og årsakforholdene mellom dem. En del bedrifter har etter hvert gått bort fra den tradisjonelle budsjetteringen og styrer kun etter et "balanced scorecard". Bare sparte arbeidsutgifter til budsjetteringsarbeidet kan utgjøre betydelige beløp.

Enda ett hakk mer avansert blir bruken av datavarehuset hvis man iverksetter såkalt *datamining* (les: inngående og kombinert bruk av eksplorative og prediktive analyseteknikker). En analytiker med egnede verktøy og den rette forretningsinnsikten kan ved hjelp av slik metodikk oppdage "gullkorn" i form av tidligere skjulte sammenhenger og årsaksforhold i datavarehuset som direkte kan relateres til effektivitet og lønnsomhet. CRM – *Customer Relationship Management* - er et samlebegrep på metoder som er beregnet på å samle inn, organisere, analysere og bruke all relevant informasjon om kundene til en bedrift. Eksempler på slike metoder: segmenteringsteknikker, logistisk regresjon, scoring, nevrale nett, forsøksplanlegging og variansanalyse. Med bakgrunn i kunnskaper og opparbeidet erfaring er det heller ingenting i veien for at automatiske analyserutiner kan defineres til å besørge deler av de nøkkeltallene som datavarehuset rapporterer videre til beslutningstagerne.

### Det finnes som regel mer informasjon i bedriftens data enn man er klar over

Det er forløpig relativt få bedrifter som benytter seg av avansert dataanalyse i sine beslutningsrutiner. Men det begynner så smått å bli bedre allment kjent hvor mye verdifull informasjon og kunnskap det potensielt er mulig å få ut av dataene i en organisasjon. Noen eksempler på spørsmål som kan besvares ved hjelp av passende analysemetodikk er: Hvilke tiltak som er gjennomført det siste året har gitt størst effekt på omsetningen? Hva har de forskjellige tiltak kostet? Finnes det kombinasjoner av tiltak som pga synergieffekter har gitt et spesielt gode effekter ? Hvilke tiltak markedstiltak har vært mest effektive? Tenkemåter fra statistisk prosesskontroll er heller ikke uten relevans i slike sammenhenger.

Man kan i skrivende stund si det er viktig å bevisstgjøre og tilgjengeliggjøre og formidle denne typen kunnskaper til lederne i en organisasjon. Ved bruk av stolpediagrammer og tradisjonell deskriptiv dataanalyse får man riktignok generell og overordnet informasjon om virksomheten. Med velegnet *datamining* er det imidlertid mulig fravriste dataene viktige "hemmeligheter" skjult i en kompleksitet som ikke kan fanges opp ved enkle deskriptive metoder.

### Moderne teknologi kan gjøre bedrifter mindre kritisk avhengighet av enkeltpersoner

Man kan dog ikke forvente så god design av datavarehus og datamining at en bedrift oppnår selvdrevet hukommelse og intelligens. Bedriften vil nødvendigvis alltid være avhengig av å ha gode medarbeidere. Moderne teknologi kan imidlertid sørge for at det blir enkelt å drifte et datavarehus med integrerte styringsverktøy som er tilgjengelige for ledelsen. Dette kan bidra til at organisasjonen styres på en mer optimal måte. I tillegg kan rutinene dokumenteres og legges opp slik at bedriftene lettere kommer i normal drift etter at sentrale personer i administrasjonen slutter. På denne måten bidrar datavarehuset og rutinene rundt det til at bedriften blir mer uavhengig av enkeltpersoner.

### Hva slags kompetanse er nødvendig?

Hvem er det man alliere seg med ved implementasjon av "hukommelse og intelligens" i en bedrift?
For å implementere et datavarehus med dataminingverktøy tilpasset bedriftens forretningsvirksomhet, trenger man forskjellige typer kompetanse. En prosjektgruppe med mandat til å etablere et datavarehus bør bestå av personer med:

- Kunnskap om bedriftens kjernevirksomhet og kunnskap om hvordan bedriftens styringsparametere bør beregnes.
- Kunnskap om databaser og datavarehusteknologi.
- Kunnskap om hvordan dataene skal bearbeides og analyseres og hvordan resultatene skal rapporteres og presenteres for ledelsen.

Kunnskap om bedriftens kjernevirksomhet er det viktigste elementet. Det er viktig at interne personer står sentralt i slike prosjekter. Det er også viktig at man har med personer som har kompetanse på bearbeiding og analysering av data, og som har erfaring i utforming av presentasjon- og rapporteringsrutiner. Betydningen av å trekke inn personer med kvalifisert kompetanse på dataanalyse blir fortsatt lett undervurdert. Ved implementasjon av et datavarehusprosjekt er det viktig at minst en person har tilfredsstillende bakgrunn innen statistikk/dataanalyse.

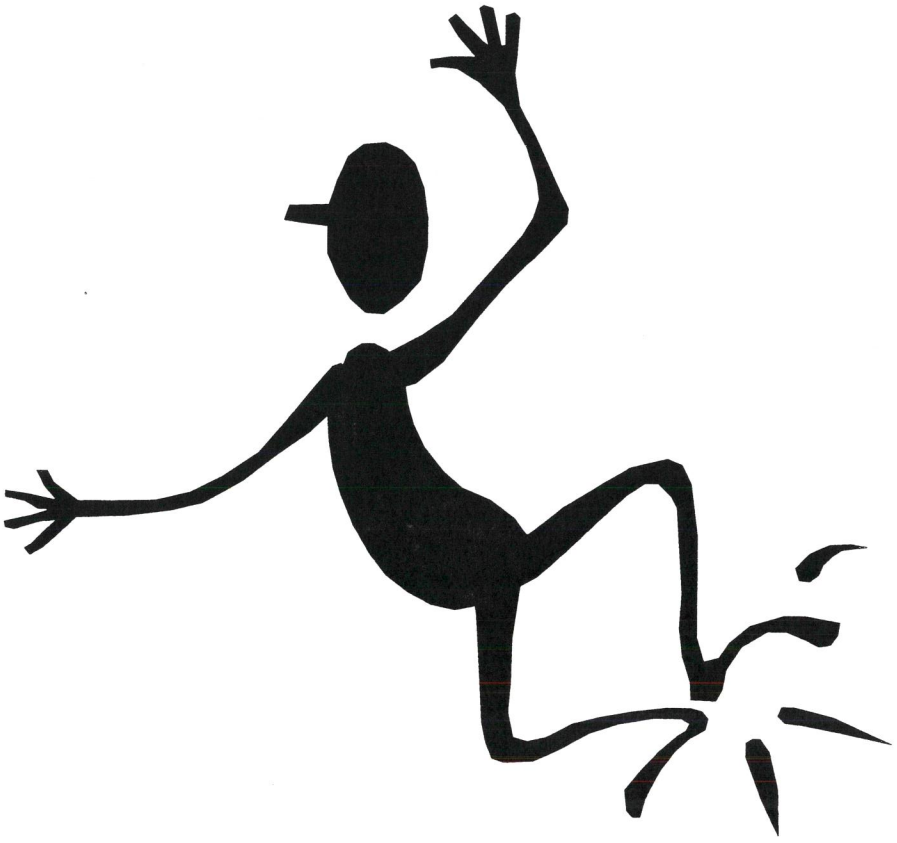
### I hvor stor grad brukes avansert dataanalyse i dag?

Dataanalyse er lite brukt i norsk næringsliv. I USA er bruken av slike metoder utstrakt i næringslivet. Det finnes ca 27.000 rene statistikere i USA, i Norge ca 300. På landsbasis utdannes det bare mellom 10 og 30 hvert år. I tillegg kommer de som har sin hovedvekt på et annet fag, men som likevel har opparbeidet tung kompetanse innen dataanalyse. Blant annet finnes matematikere, sosialøkonomer og biologer som er gode i dataanalyse. I Norge er det bare innenfor noen få bransjer hvor avansert dataanalyse i stor utstrekning blir brukt.

- Bank, finans og forsikring
- Forskning og utvikling

Det er imidlertid ingenting i veien for å bruke metodikken i hvilken som helst bransje. Alle bransjer trenger informasjon som grunnlag for å ta beslutninger. Systematisk bruk av dataanalyse gir bedre beslutningsgrunnlag og dermed en mer optimal forretningsvirksomhet.

Undertegnerne arbeider som konsulenter i Teknometri AS. Selskapet har spesialisert seg på å *organisere, bearbeide, analysere* og *foredle* den informasjon som finnes i de tilgjengelige data og presentere resultatene på en måte som gjør at beslutningstakere både forstår og kan gjøre bruk av dem.



# *Puslerispalten*

## Pusleri nr. 3

### Hvorfor går bussene oftest i motsatt retning?

Jeg ankommer bussholdeplassen nokså tilfeldig hver morgen, og registrerer hvor mange busser som går i motsatt retning før første buss i min retning kommer. Over en periode på 2 uker var det:

2 1 1 0 0 2 3 1 2 0

I gjennomsnitt noe over 1. Det var mystisk! Det må jo være sannsynlighet ½ for at min buss kommer først og ½ for at en buss i motsatt retning kommer først. Burde ikke derfor gjennomsnittlig antall busser i motsatt retning før min kommer være ca. ½ ?

Problemstillingen kan sikkert belyses på ulike måter, og denne gangen ønsker vi gode løsninger fra lesere for presentasjon i neste TG, fra den enkle og intuitive til den matematisk sofistikerte under ulike antakelser.

Send løsninger innen 4 uker fra mottatt TG til jostein.lillestol@nhh.no.

**Jostein Lillestøl**

### Løsning av Pusleri nr. 2

*Vi minner om oppgaven:* En viss type amøbe har egenskapen at hvert minutt vil den enten dø, forbli som den er, dele seg i to eller dele seg i tre. Anta at de fire mulighetene er like sannsynlige. En slik amøbe puttes i en krukke. Hva er, med rimelige tilleggsantakelser, sannsynligheten for at amøbe(kolonien) i krukken dør ut?

*Løsning:* Vi skal finne sannsynligheten p for at en amøbe og dens etterkommere dør ut (uten bruk av teori for Branching Processes). Sannsynligheten for at en koloni med k amøber dør ut er, med uavhengighetsantakelse, lik $p^k$. Ved å betinge mhp. tilstanden etter et minutt får vi "likevektsligningen"

$$p = (p^0 + p^1 + p^2 + p^3)/4$$

Dette gir oss en tredjegradsligning, $p^3 + p^2 - 3p + 1 = 0$, som har to aktuelle røtter, $p = 1$ og $p = \sqrt{2} - 1$. Et mulig argument for å velge den siste roten er:
Anta at sannsynligheten for at en amøbe dør/forblir som før isteden er hhv. $1/4 - x$ og $1/4 + x$, der x kan være fra $-¼$ til ¼. Den tilsvarende ligning har da løsning $p = 1$ og $p = \sqrt{2 - 4x} - 1$. Siden $x = 1/4$ må gi $p = 0$, er den første roten uaktuell.

# 500 skoleklasser får besøk av matematikere og statistikere!

År 2000 er verdens matematikkår. Dette gir en anledning til å sette matematikk på "dagsorden" og for å arbeide for økt interesse for matematikk og matematiske fag. Ved Matematisk institutt ved Universitetet i Oslo har dette initiert flere aktiviteter (se www.math.uio.no). Den viktigste av disse retter seg mot elever i alderen 10-19 år, og er et samarbeid med Norsk Regnesentral, Institutt for lærerutdanning og skoleutvikling (UiO) og Matematikkseksjonen ved allmenlærerutdanningen på Høgskolen i Oslo.

Rett over nyttår ble det sendt et tilbud til alle skoler i Oslo og Akershus om besøk av fagfolk som kunne holde populærfaglige foredrag eller lage andre klasseaktiviteter knyttet til matematikk og matematikkrelaterte emner. Skolene ble lovet at "felles for foredragsholderne er faglig ekspertise og engasjement, formidlingsiver og vilje til å møte elevene på deres nivå med spennende innblikk i matematikk og matematikkens historie, sannsynlighetsregning og statistikk og morsomme og nyttige anvendelser av matematikk."

De av de ansatte som sa seg interessert i å bli med på klassebesøkene, forpliktet seg til å stille opp 2-3 ganger i løpet av ett semester.

Listen av emner som ble tilbudt var omfattende. Noen eksempler:

* Talldjevelen kommer på besøk.
  Med utgangspunkt i boka "Talldjevelen" av Hans Magnus Enzensberger tar vi for oss et av emnene: Tallrekker, primtallene, kombinatorikk.

* Spillteori.
  Vi tar for oss noen forskjellige spill og ser om vi kan regne ut hvem som vinner og hva slags strategier som det lønner seg å bruke.

* Matematikk og musikk.
  Hvorfor har skalaen 12 halvtoner? Og kan musikken bli bedre med 19?

* Hemmelige koder.
  Hvordan matematikk brukes til å lage hemmelige koder.

* Kan du løse likningen $x^2 = -1$ ?
  Enkel innføring i de komplekse tallene med utgangspunkt i Caspar Wessels metode for multiplikasjon av rette linjestykker.

* Risikovurderinger, nyttig for flere enn luft- og finansakrobater?
  Anvendelser av Bayes Teorem på beslutninger under usikkerhet.

* Å se på ting vi ikke kan se.
  Bruk av statistikk og matematikk til å se inn i menneskekroppen, studere jordoverflaten eller kjenne igjen et fingeravtrykk.

* Den som leter skal finne - hypotesetesting til besvær.
  Ukritisk bruk av hypotesetesting kan føre galt av sted. Dette vil bli illustrert ved eksempler og beregninger.

Responsen på tilbudet var over all forventning. Over 500 skoleklasser fra 150 skoler meldte sin interesse. Hovedtyngden av disse var fra mellomtrinnet (4-7. klasse) og ungdomstrinnet (8-10. klasse), men også flere videregående skoler var interessert.

Matematikerne har tatt den største jobben i forbindelse med klassebesøkene. Så overskriften som likestiller matematikere og statistikere er i denne sammenhengen nokså tendensiøs. Men også statistikerne ved Avdeling C har gjort en innsats. Vi har så langt vært på besøk på videregående skoler (og i en ungdomsskoleklasse) med de tre siste temaene nevnt over. Samlet klarer å avvikle besøk i 200 klasser (delvis flere på en gang) i løpet av vårsemesteret. Resten må vente til høsten.

Hva har reaksjonen fra elevene vært? Vi har ikke noe systematisk informasjon om dette. Men mange av oss som har vært ute på skolen, har opplevd at elevene følger svært interessert med. Responsen fra lærerne har også vært veldig positiv.

Et spørsmål som har vært reist, er om slike klassebesøk er noe de ansatte bør bruke tid på i en travel hverdag. Et annet spørsmål er om aktiviteten bør fortsette etter at verdens matematikkår er over. For undertegnede er svaret på begge spørsmålene et klart ja! Dette er i en viss grad motivert ved en "folkeopplysningstankegang". Men mest er det en "egoistisk" motivasjon: Rekrutteringen til statistikk og andre matematiske fag er i dag svært dårlig. Dette må vi arbeide systematisk og på mange fronter for å endre. En slik "front" er skolen, hvor vi finner de potensielle studentene for de neste 10-20 år.

Ørnulf Borgan

# *Alarm!*
# Statistikk i ny 3-årig økonomisk-administrativ utdanning

ved Jostein Lillestøl

Den økonomisk-administrative utdanning ved de statlige høyskoler er i støpeskjeen. Den tilbys i dag ved 20-30 statlige høyskoler, og er i utgangspunktet 2-årig med enkelte påbygningsmuligheter. Første studieår har til nå inneholdt metodekurs i matematikk, statistikk og EDB, hvert av omfang tilsvarende 2 vekttall. Norsk råd for økonomisk-administrativ utdanning har arbeidet med et nytt 3-årig opplegg som skal erstatte det 2-årige. Her ser det ut til at metodefagene blir fullstendig marginalisert. Det er foreslått å begrense det obligatoriske metodepensumet til et tovekttalls kurs kombinert matematikk/statistikk. Eventuell mer metode som trengs vil så bli undervist i tilknytning til de økonomiske fag av økonomer. Bakgrunn for dette er i hovedsak markedsorientert og ikke faglig. Få krav til forkunnskaper, større marked, færre strykere, muligheter for brevundervisning, internett og etterutdanning.

Forslaget innebærer en betydelig fare for fagets fremtidige stilling rundt om ved høyskolene, og Norsk Statistisk Forening ser med bekymring på det som er i ferd med å skje. Undervisningsutvalget i NSF fikk derfor i oppdrag å forfatte en uttalelse som er sendt til de rette instanser i den videre beslutningsprosess. Denne uttalelse er gjengitt nedenfor:

**Norsk Statistisk Forening** **27. april 2000**

## Statistikkfaget i forslag til nytt 3 årig øk-adm studium

Forslaget innebærer at metodefag har fått en langt svakere stilling i det 3 årige studium enn i det 2 årige. **Det er med vantro og undring vi ser at statistiske metoder ikke har fått en større og mer sentral plass i studieplanen**. Det er å underkjenne den selvstendige rolle statistisk kunnskap og ferdigheter har i økonomisk-administrativt virke.

Det dreier seg om langt mer enn begreper og metoder som kun har sin legitimitet ved anvendelser i andre fag. **Evnen til å skaffe og utnytte data blir i dag sett på som en viktig, men ofte oversett, konkurransefaktor i bedriftene**. Samtidig møter bedriftene oftere krav til dokumentasjon knyttet til kvalitet, helse, miljø og risiko. Slike lar seg vanskelig oppfylle uten grunnlag i statistiske metoder, se vedlegg.

Det kan nok diskuteres om ikke innholdet i metodekurset i statistikk i dagens 2-årige øk.-adm. utdanning kan tilpasses bedre til det nye studiets og samfunnets behov, og faglærere i statistikk vil sikkert kunne bidra konstruktivt i en slik diskusjon. Det er nokså drøyt å fjerne istedenfor å reformere! Det har skjedd mye spennende innen statistikkfaget de senere år, som nettopp tilrettelegger for praktisk bruk i næringsliv, i stor grad kombinert med moderene informasjons- og kommunikasjonsteknologi.

Statistikk fikk for få år siden plass i matematikkfaget i videregående skoles 2.- og 3.klasse (innholdet er nå under justering). Innslaget av statistikk i det foreslåtte obligatoriske 2 vekttalls kombinerte kurset i matematikk/statistikk i bachelor-studiet rekker knapt å fylle eventuelle hull for de som har valgt bort matematikk i videregående skole. Samtidig blir det et nokså meningsløst gjentak uten videreføring for studenter som har mer enn minimum i

matematikk fra videregående. skole. Med utvidelse av øk.-adm. utdanningen til 3 år, ligger forholdene til rette for å styrke innslaget av statistisk metode. Problemet med statistikk i den 2-årige utdanningen har muligens vært teori med for knapp tid til anvendelser. Dette kan nå rettes på!

**Det foreslåtte opplegg vil neppe være tilstrekkelig metodegrunnlag for anvendelser på en rekke områder innen studiet selv og heller ikke hva en venter i praksis etter en treårig utdanning.**

Det foreslåtte opplegg virker langt svakere mht. metode enn tilsvarende utdanninger i utlandet. En sammenligning med business studier i USA kan være på sin plass: Den typiske lærebok i business statistics er ganske omfattende, med nok stoff til 2 + 2 norske vekttall. Her integreres teori, eksempler, cases og bruk av programvare.

I de senere år har det skjedd en utvikling der oppmerksomheten i større grad rettes mot bedrifters og organisasjoners forbedring av produkter og prosesser. Her er systematisk metode og statistikk sentrale redskaper. Lite tyder på at bevisstheten om statistikkfagets rolle innen business og management er svekket. Det holdes regelmessige konferanser om "Teaching statistics at business schools", men synspunkter og erfaringer derfra ser ut til å ha gått hus forbi i Norge. I Europa er det nylig etablert et nettverk for Business and Industrial Statistics (ENBIS) som tar sikte på å utvikle samspillet mellom næringslivet og undervisningsinstitusjoner på feltet.

I tråd med begrunnelsene ovenfor gir vi følgende anbefaling:

1. **Statistikkfaget må beholde sin plass i planene for den nye 3-årige økonomisk-administrative utdanningen ved de statlige høyskolene. Faget må ha sin egen identitet, og høyskolene må ha lærekrefter med spesielt ansvar for å utvikle faget over tid og tilpasse det en økonomisk-administrativ kontekst.**

2. **Et kurs i statistikk må inn på timeplanen i 2. studieår. Vi foreslår:**

**MS2. Statistisk dataanalyse i bedriften**

**med målsetning:**

**"Å utvikle evnen til å utnytte data og statistisk tenkning, med formål å øke bedriftens verdiskapning"**

Dette innebærer at vi foreslår at dette kurs inngår i rammeplanen istedenfor SMS2. Samfunnsvitenskapelig metode. Dette kurs er i rammeplan foreslått som et obligatorisk 2-vekttalls kurs i 2.studieår på bekostning av de obligatoriske metodefagene i dagens 2.årige studium. En god begrunnelse for dette mangler. Samfunnsvitenskapelig metode kan **ikke** i samme grad som statistisk metode begrunnes ut fra formålet med studiet og bedriftenes behov. Et slikt kurs kan nok inneholde enkelte nyttige statistiske metoder som det ikke er plass til i 1.studieår, men neppe ut fra organisasjonens perspektiv: å avdekke kundebehov, og å produsere produkter og tjenester som oppfyller disse. Ovennevnte forslag kan oppfylle et slikt krav, og innebærer en spennende utfordring for fagmiljøene. Ved at kurset kommer i annet studieår kan en i større grad enn i dag ta utgangspunkt i problemstillinger fra økonomi og administrasjon, og også bidra til den praktiske bruken av ITK-verktøy.

Med bedriften i fokus vil en ikke i samme grad som nå være bundet av matematikkens krav til stringens og formelle oppbygning. Et slikt kurs vil også være verdifullt som frittstående kurs for næringslivet.

Statlige høgskoler som tar sikte på en bestemt profilering og spesialisering, vil kan hende ha metodebehov ut over dette.. Det er i slik sammenheng at samfunnsvitenskapelig metode kunne få plass i planen. Dette kurs kommer i samme stilling som kurs i statistiske metoder med mer analytisk profil, som f.eks. kan støtte programmer med vekt på operations management (herunder logistikk) og finans. Et slikt kurs i statistiske metoder kan vektlegge stokastiske modeller som dagens 2.vekttalls kurs, men med et sterkere innslag av stoff hentet fra den aktuelle profilering. Studenter som vurderer å gå videre utover den 3 årige utdanningen, f.eks. til siviløkonomstudiet, vil være tjent med at slike metodetilbud fins.

Det tales i mange land om "numerical illiteracy", og at dette er noe utdanningssystemet må gjøre noe med. Forslaget til rammeplan for den nye 3-årige utdanningen i øk.-adm. fag, der både matematikk og statistikk er svekket , vil uten tvil bidra til å forsterke den negative utviklingen også i Norge. **Man kan ikke la ønsker om salgbarhet og studentgjennomstrømning være førende for det faglige opplegg, på bekostning av hva som kan være med på å sikre konkurransekraft for Norge i fremtiden!**

**Vedlegg: Statistikk og kvalitet**

Bedrifter og organisasjoner som ønsker å ha et oppegående kvalitets- og HMS-system vil nå og i fremtiden måtte forholde seg til statistikk og statistiske metoder. De mange svakheter ved slike systemer i dag går nettopp på manglende evner til skaffe relevante data og utnytte de data som fins. I Norge er dette bl.a. erkjent ved at Norsk forening for kvalitet og lederskap (NFKL) har et eget fagstyre for statistiske metoder (som har skaffet nedenstående opplysninger).

Statistiske metoder var sentrale i den japanske tilnærming til kvalitet allerede fra 1960-årene. Oppmerksomheten var rettet mot styring og forbedring av alle organisasjonens prosesser, i produksjon såvel som i administrasjon. Dette ble fanget opp i Vesten først i 1980-1990-årene. I dag krever mange foretak av sine leverandører at de skal dokumentere styring og forbedring, noe som er utenkelig uten statistisk metode på alle nivåer. Dette fremkommer bl.a.i den reviderte standarden *IS09000:2000 Quality management systems – Requirements*, som bedrifter og organisasjoner må forholde seg til. I paragraf 8.Measurement, analysis and improvement avsnitt 8.1. Planning:

"*The organization shall define, plan and implement the measurement and monitoring activities needed to assure conformity and improvement. This shall include the determination of the need for, and use of, applicable methodologies including statistical techniques*".

Dette følges opp i avsnittet 8.4 Analysis of data og i 8.5 Improvement. Disse temaer følges opp i den nye *ISO9004: Quality management systems – Guidelines for performance improvements*, der det bl.a. heter:

"*The organization should analyse data from various sources to assess performance against plans and goals and to identify areas for improvement. The organization should plan to use statistical methodologies for data analysis, which can help in assessing, controlling, and improving performance of processes.*"

**Norsk statistisk forening,**
**ved formann Geir Egil Eide,**
**Seksjon for matematikk og statistikk,**
**NHH, Bergen**

2000-04-18

## INNKALLING TIL EKSTRAORDINÆRT HOVEDMØTE

Etter vedtak i styret i henhold til § 4 i foreningens vedtekter innkalles herved alle medlemmer til ekstraordinært hovedmøte i Norsk statistisk forening:

Tidspunkt: Mandag 5.juni 2000 kl.18.15

Sted: Sørlandet kurs- og hotellsenter, Grimstad

Møtet arrangeres i tilknytning til 18th Nordic Conference on Mathematical Statistics.

**DAGSORDEN**

**1. Godkjenning av innkalling og saksliste.**

**2. Orienteringssaker.**

**3. Endringer i foreningens vedtekter.**

Forslag nr. 1. Endringer i NSFs vedtekter § 3 med kursiv.

§ 3 LOKALE FORENINGER *OG FAGSEKSJONER*

Medlemmer innen et geografisk område kan danne en lokalforening. Lokalforeningene har egne vedtekter og egen økonomi. Vedtektene godkjennes av styret *i NSF.*

*Medlemmer innen et faglig spesialområde kan danne en fagseksjon. Fagseksjonene har egne vedtekter og egen økonomi. Vedtektene skal godkjennes av styret i NSF.*

Lokalforeningene *og fagseksjonene* kan fastsette egen årskontingent, som kommer i tillegg til kontingenten til NSF. *Lokalforeningene og fagseksjonene kan søke styret i NSF om økonomisk tilskudd.*

Forslag nr. 2. Formannen endrer tittel til 'leder'. Alle henvisninger til 'formann' i vedtektene endres tilsvarende.

**4. Endringer i vedtektene for foreningens fond.**

Forslag. Endring av fondsvedtektenes § 1: Følgende punkt tilføyes:

- *stipend til hovedfagsstudenter i statistikk for deltakelse på faglige konferanser i Norden.*
- *til andre formål styret måtte beslutte med inntil kr. 30 000 pr. år.*

**5. Eventuelt.**

## Programkomite for det 11. norske statistikermøtet i år 2001

Programkomite for det neste norske statistikermøtet er nå oppnevnt og representerer såvel næringsliv som forskning og undervisning. Komiteen består av

Bent Natvig (Mat. inst. avd. C, UiO),
Alfhild Lien Eide (Statoil, Trondheim),
Ole Lauvskar (Vital, Bergen),
Hans A. Karlsen (Mat. inst, UiB),
Grethe Albrektsen (SMiS,UiB)
Knut Korsbrekke (Havforskningsinstituttet, UiB).

Knut Korsbrekkke er leder av komiteen og kan kontaktes på Knut.Korsbrekke@imr.no.

Arrangementskomite denne gang er styret i Statistisk forening i Bergen med leder Grethe Albrektsen.

## Etterutdanningsseminar

Vi arbeider for tiden også med å arrangere et etterutdanningsseminar muligens i tilknytning til neste års statistikermøte. Ideer til tema og undervisere kan sendes så snart som mulig til undertegnede på Geir.Eide@NHH.No.

Geir Egil Eide, Bergen 25.april 2000

# Resultater fra spørreundersøkelse om reklame i regi av Norsk statistisk forening (NSF)

I forbindelse med styremøtet i NSF i februar i år ble det diskutert om foreningen skulle kunne sende ut reklame til sine medlemmer mot betaling for å øke foreningens inntekter. Styret fant det riktig å foreta en spørreundersøkelse blant medlemmene før dette eventuelt ble iverksatt.

Følgende spørsmål ble sendt via e-post til medlemmer med kjent e-postadresse i slutten av februar 2000. Spørsmål skulle besvares med: ja, nei, ingen mening, (tja).

*1. Jeg mener at foreningen kan sende ut aktuell reklame med epost til sine medlemmer og ta honorar for dette.*

*2. Jeg mener at foreningen kan selge sin epostadresseliste til kommersielle interessenter under forutsetning av at detaljert avtale om bruken inngås mellom styret og kjøper av adresselisten.*

*3. Jeg mener at foreningen kan selge sin ordinære adresseliste til kommersielle interessenter under forutsetning av at detaljert avtale om bruken inngås mellom styret og kjøper av adresselisten.*

NSF har 253 medlemmer. Spørsmålene ble sendt til de 181 medlemmer vi har oppført med epost-adresse. Seks mailer kom i retur pga feil adresse, slik at vi hadde 175 mulige respondenter. Svarprosent: 42,9%

Spørsmål 1

| | antall | prosent |
|---|---|---|
| ja | 56 | 75,7 |
| nei | 17 | 23,0 |
| ingen mening | 1 | 1,4 |
| | 74 | 100,0 |

Spørsmål 2

| | antall | prosent |
|---|---|---|
| ja | 35 | 47,3 |
| nei | 37 | 50,0 |
| ingen mening | 2 | 2,7 |
| | 74 | 100,0 |

Spørsmål 3

| | antall | prosent |
|---|---|---|
| ja | 35 | 47,9 |
| nei | 33 | 44,6 |
| ingen mening | 3 | 4,1 |
| tja | 2 | 2,7 |
| | 73 | 100,0 |

## Statistics in Business and Industry (SBI)

The International Statistical Institute's SBI Committee aims to support statisticians working in business and industry by monitoring their needs on an ongoing basis, promoting research and applications and best current practices, facilitating technology transfer, and fostering communication among members.

The SBI Committee seeks to promote (both to statisticians and to industry) the value and importance of Statistics in Business and Industry, and to support suitable activities, particularly in lesser-developed countries.

A Web site and electronic discussion group have recently been established. We invite you to make use of these, so that we can facilitate international cooperation in business and industrial statistics.

The Web site can be found at: www.public.iastate.edu/~sbi .

To join the sbi-net Newsgroup, send a message to

sbi-net-request@iastate.edu

with the single word

subscribe

in the main body of the message. (This can be also done via the Web page).

IMPORTANT NOTE: you will receive two automatically generated reply from majordomo@iastate.edu. One is an authentification key; the other requests that you submit the key. The purpose of this is to avoid other people adding your name to the list.

To send a message to subscribers to sbi-net, send mail to sbi-net@iastate.edu

The SBI Committee looks forward to hearing from you.

Nicholas I Fisher, CSIRO Math. & Inf. Sciences
Locked Bag 17, North Ryde NSW 1670 Australia
Tel +61 2 9325-3258 Mobile 0407 017 016 (Int. +61 407 017 016)
Fax +61 2 9325-3243 URL www.cmis.csiro.au/Nick.Fisher

# Medlemsskap i Norsk statistisk Forening

Kontingent:
200 kr/år
50 kr/år for studenter (til og med hovedfag).
Kontonummer: 0530 2773299

Nummeret som er angitt på adresselappen på TG angir siste år kontingenten ble betalt.
Ved innbetaling må man opplyse om endret adresse (f.eks endret postnummer).

Korrespondanse om medlemsskap og innbetaling av kontingent skal skje til foreningens kasserer:

Norsk Statistisk forening v/Ågot Irgens
Kalfarlien 6A
5018 Bergen
Emailadresse: agot.irgens@haukeland.no

Nye medlemmer melder seg til kasserer og får så automatisk tilsendt Tilfeldig Gang fra første påfølgende nummer av.

***Mailadresser***:
Vi ønsker mailadresser til alle medlemmer i NSF:
Oppgi e-mailadressen din på merknadsfeltet på postgiroen, eller send en e-mail til agot.irgens@haukeland.no om mailadresse og evt adresseendring.

Annen korrespondanse med foreningen skal foregå med formannen:

Geir Egil Eide
Norges Handelshøyskole
Helleveien 30
N-5045 Bergen
Emailadresse: Geir.Eide@nhh.no

B

**RETUR**

Norsk Statistisk Forening
v/Ågot Irgens
Kalfarveien 6 A
5018 BERGEN
Email: agot.irgens@haukeland.no