ISSN 0803-8953

TILFELDIG GANG
Nr. 3, årgang 17
September 2000

Utgitt av Norsk Statistisk Forening

INNHOLD

	Side
Redaksjon:	1
Fra redaktøren. *Odd Aalen*	2
Fra leder i NSF. *Geir Egil Eide*	3
Statistisk arbeidsliv	
Statistikk på Fiskeriforskning. *Alf Harbitz*	4
Litt om virksomheten ved Seksjon for medisinsk statistikk i Bergen. *Stein Emil Vollset*	8
Etablering av KKF ved Haukeland sykehus i Bergen. *Geir Egil Eide*	11
Statistisk reisebrev	
På faglig besøk i Berkeley. *Espen Norberg*	12
Inntrykk fra Nordisk møte i Grimstad	
Reisebrev fra Grimstat!!!. *Henning Omre*	14
Som lærlinger til Grimstad. *Gunnhildur Högnadóttir Steinbakk, Birgitte Ringstad Vartdal og Ane Seierstad*	16
Puslerispalten	
Pusleri nr. 4. *Jostein Lillestøl*	18
Løsningsforslag til Pusleri nr. 3. *Jostein Lillestøl, Geir Arne Kjønstad og Trond Haider*	19
Fra fagets historie	
Hva kan vi slutte av statistikk? Andet foredrag. *Erling Sverdrup* ...	22
Intervju med Olav Reiersøl. *Yngve Willassen*	33
Meldinger	
Referat fra ekstraordinært hovedmøte i NSF. *Stian Lydersen*	47
Marit Bragelien Veierød har tatt Dr. Philos.-graden	48
Nytt fra Universitetet i Oslo	49
Foreløpig annonsering av neste norske statistikermøte.	51
International School on Mathematical and Statistical Applications in Economics.	52
Medlemsskap i Norsk statistisk forening	54

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 15.11.2000.

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:

Når dette nummeret leses er allerede det sagnomsuste år 2000 på hell. Etter hvert virker det som et helt alminnelig år, ingen store katastrofer, ingen vedvarende følelse av noe særlig spesielt.

Fortsettelsen av professor Erling Sverdrups radioforedrag fra 1950-tallet, som gjengis i dette nummer, bringer oss likevel tilbake til en annen tid. En tid da mange mennesker land og strand rundt, benket seg rundt radioapparatene for å høre på et lengre foredrag av en professor om et akademisk tema. I dagens høytempo medie- og internettverden fremstår dette som en anakronisme. Likevel, hvis man tar seg tid til å lese foredraget, vil man oppdage mye til faglig ettertanke. Sverdrup var en tenker, opptatt av de grunnleggende idéene i faget vårt. Han så ikke statistikken primært som en teknisk-matematisk disiplin, men som et begrepsapparat og en tenkemåte. Han hadde ofte en filosofisk tilnærming til faget.

Også en annen av fagets gamle nestorer figurerer i dette nummeret. Professor Olav Reiersøl var, ved siden av Sverdrup, med å bygge opp den akademiske utdannelse i statistikk i Norge. Reiersøl har sine røtter i den ærerike økometriske tradisjon, et felt der Norge har gjort seg meget bemerket gjennom to Nobelpriser (Frisch og Haavelmo). Fremfor alt fremstår Reiersøl som den grundige og originale matematiker.

I dette nummeret starter vi ellers en ny serie. Vi vil la forskjellige fagmiljøer og fagpersoner få anledning til å presentere seg og sin virksomhet. Hva er det egentlig statistikere driver med i Norge anno år 2000? Alf Harbitz skriver om sin virksomhet ved Fiskeriforskningsinstituttet i Tromsø. Stein Emil Vollset presenterer Seksjon for medisinsk statistikk i Bergen og forteller litt om prosjekter der. Begge disse arbeidsstedene fremstår som spennende og fremtidsrettede, og med svært varierte og utfordrende arbeidsoppgaver. Både kartlegging og utnyttelse av fiskeriressurser og arbeidet med folkehelse er helt sentrale samfunnsoppgaver også i et nytt århundre.

Jostein Lillestøl fortsetter med å sprite opp bladet med sine underholdende og utfordrende Puslerier. Leserne oppfordres til å løse dem, til å svare, og til å sende idéer om nye puslerier til Jostein.

Vi vil stadig gjerne ha reiserapporter til TG, og presenterer en i dette bladet. Vi vil også svært gjerne ha omtaler av runde år og andre store begivenheter. De lokale miljøer må selv ta ansvar for å markere sine kolleger når de store runde dager passeres.

Rapporter fra et meget vellykket Grimstadmøte figurerer også i bladet. For redaktøren var det et av de mest givende møter jeg har vært på. Det skyldes ikke minst de utmerkede oversiktforedragene som ble gitt på flere krevende felter. Både organisasjons- og programkomité gjorde en usedvanlig bra jobb, og skapte nytt håp for de nordiske møtene.

Odd Aalen

Fra lederen:

Høsten er kommet på oss og de fleste av oss er travelt opptatt med undervisning, forskning og statistisk analyse. Perioden fra mai til september er vanligvis høysesong for faglige møter og konferanser og forhåpentligvis har mange av foreningens medlemmer deltatt og bidratt på slike aktiviteter og startet høsten med ny giv og nye ideer! Personlig deltok jeg bl.a. på det nordiske møtet i matematisk statistikk i Grimstad og det var kjekt å oppleve at det tekniske og det faglige opplegget var tatt utmerket vare på av de norske arrangørene. All honnør til alle som var involvert fra NSF. Vi hadde også et konstruktivt ekstraordinært hovedmøte der nødvendige justeringer i vedtektene ble vedtatt (se hjemmesiden til NSF). Nå er det igrunnen bare opp til foreningen selv og dens medlemmer å benytte seg av de nye mulighetene vedtektsendringene gir. - Det var også positivt å oppleve deltakelsen i Grimstad av de NSF-støttete studentene, ikke minst på det sosiale området, og tiltaket med studentstipend til slike møter bør absolutt videreføres. Verdt å nevne er også initiativene omkring undervisning og undervisningsplaner som ble tatt på Grimstad-møtet og som vi forhåpentligvis vil se frukter av i det lange løp. - Etter Grimstad-møtet deltok også noen av oss på biometri-konferanse i Berkeley, California og det var greit å se at arrangementsmessig så sto ikke vårt nordiske møte noe tilbake for det internasjonale på et verdensberømt universitet, snarere tvert om. (Faglig var det overveldende å oppleve at på tross av 7 parallelle sesjoner hver dag i Berkeley, så var der alltid en sesjon på levetidsanalyse....).

Nåja, dette var et lite tilbakeblikk. Fremover ser vi nå bl.a. mot neste års norske statistikermøte i Bergen. Arrangementskomiteen er allerede i sving med planleggingen og i tilknytning til det 3 dager lange møtet vil det også bli tilbudt et 2 dagers etter- og videreutdanningskurs. Tema for kurset vil sannsynligvis bli eksakte metoder (trenger godkjennelse av styret) og det er gjort avtale med prof. Cyrus Mehta om at han vil komme å gi kurset. Målet er at kurset skal omfatte halvt om halvt med teori og dataøvelser spesielt med bruk av StatXact og LogXact-pakkene. Eksakte metoder kjennetegnes ved at de fort blir svært beregningskrevende (tenk bare på å beregne Fishers eksakte test for 2x2 tabeller manuelt), men nye algoritmer og stadig raskere datamaskiner har ført til økende interesse for slike metoder når de asymptotiske resultatene ikke kan brukes. Kurset skulle forhåpentligvis være av interesse både for praktikerne og teoretikerne blant oss. Så merk allerede nå av for dagene fra 18. til 22. juni 2001 i kalenderen!

Til sist vil jeg bemerke at det er kjekt å lese i media at tilstrømningen til realfagene igjen er på oppadgående. Forhåpentligvis er dette en trend som vil vedvare og få marginaleffekter på vårt lille fag. Det skulle i hvert fall ikke være noen grunn til å ikke anbefale statistikkfaget av arbeidsmarkedsmessige hensyn slik situasjonen er for tiden.

Vennlig hilsen

Geir Egil Eide
Leder i Norsk Statistisk Forening

Statistikk på Fiskeriforskning

av Alf Harbitz

Fiskeriforskning

Fiskeriforskning er et oppdragsforskningsinstitutt i Tromsø. Instituttet er et aksjeselskap eid av NORUT Gruppen (51 prosent) og Fiskeridepartementet (49 prosent). Instituttet ligger sentralt plassert på universitetsområdet i Breivika, mellom Norges fiskerihøgskole og bygget der statistikkavdelingen ved universitetet holder til. Fiskeriforskning har ca. 100 ansatte, og har som formål å drive anvendt forskning til nytte for fiskerinæringen. Fagfeltene er svært vide og dekkes av 5 ulike sentre: Økonomi og marked, industriell foredling, marin bioteknologi, marine ressurser og havbruk. Populært sier vi at arbeidsoppgavene omfatter alt "fra bunn til munn". Som eksempler på spennvidden kan nevnes et prosjekt som nå er kommersialisert, der chitosan ekstrahert fra rekeskall har vist seg ypperlig til bruk i kosmetiske kremer, og et annet prosjekt der det ble påvist at rekelarver spiser torskelarver (hvem spiser hvem er et aktuelt spørsmål). Sistnevnte prosjekt ble presentert på Schrødingers katt tidligere i år.

Fiskeriforskning har omfattende eksperimentell virksomhet, blant annet ved en moderne havbruksstasjon i verdensklasse i Kårvika på Ringvassøya utenfor Tromsø. Forsøksplanlegging er derfor ett av flere sentrale fagområder. Virksomheten ved Fiskeriforskning bibringer store datamengder og statistiske problemstillinger av svært varierende karakter, og instituttet er således et eldorado for en statistiker. Fiskeriforskning bedriver ikke undervisning, men har en del veilederoppgaver for fiskerikandidater, diplomkandidater, hovedfags- og doktorgrads-studenter.

Faglig bakgrunn og noen betraktninger

Jeg ble ansatt ved instituttet i 1996 som forsker ved Senter for marine ressurser, i den eneste statistikkstillingen ved Fiskeriforskning. Jeg er utdannet fysiker med dr.scient. grad fra Tromsø i 1990 (videobasert sjøbølgeanalyse). Imidlertid har jeg arbeidet mye med statistikk anvendt på konkrete ingeniør-relaterte problemstillinger fra slutten av 70-tallet, avlagt statistikk-eksamener på til sammen ca. 30 vt, og min avhandling inneholder betydelige islett av statistisk modellering og analyse. Min noe brokete bakgrunn har blant annet gjort meg svært ydmyk overfor nye fagfelt, og generelt når det gjelder kopling mellom teori og praksis. Dette er en ballast som etter min mening er viktig når man med en relativt sterk teoretisk bakgrunn fra "harde" realfag skal "jumpe inn" i en komplisert marinbiologisk verden der forskningen historisk sett er svært deskriptiv. Innen marinbiologien er det få, om noen, allmene (deterministiske) lovmessigheter ala Newtons lover i fysikken. Dessto viktigere blir da statistikk som verktøy, ikke minst grunnet den raske utviklingen innen måleteknologi og data. Trolig er kravet til omfattende litteraturstudier større enn innen mer matematisk orienterte fag, fordi oversikten over fagfeltet betinger at man møysommelig setter sammen en rekke, ofte overraskende, detaljer for å få en helhetsforståelse.

Senter for marine ressurser

Senter for marine ressurser utfører forvaltningsmessig forskning på sjøpattedyr (sel og kval), fisk (kysttorsk, blåkveite, gapeflyndre) og skalldyr (reke, kongekrabbe) med Fiskeridepartementet som oppdragsgiver. Mine to største arbeidsoppgaver har sålangt vært konsentrert om optimalisering av toktrute for et årlig reke- og flatfisktokt i Barentshavet, samt analyse av sjøpattedyrs diettsammensetning og sammenligning med artssammensetningen av byttedyr i havet. Optimalisering av toktrute er ikke minst viktig fordi toktene er svært dyre (ca. 100 000 pr. dag i båtleie). Diettanalyse av sjøpattedyr er viktig blant annet fordi de er mange, spiser mye, og

det er viktig å vite hvor mye de beskatter av viktige kommersielle bestander som for eksempel torsk. Eksempelvis er det anslagsvis ca. 2 millioner grønlandssel i våre nordlige havområder, som grovt anslått spiser 2 millioner tonn av havets ressurser pr. år.

Toktvirksomhet

Jeg er med på to tokt til sjøs pr. år, til sammen ca. 30 døgn, der jeg foretar prøvetaking på lik linje med de andre. Dette gir en unik mulighet til å få nært og praktisk kjennskap til de prosesser som skal modelleres og analyseres, inklusive detaljkunnskap om datainnsamlingen og inspirasjon til eventuelt å forbedre denne. Det viser seg for eksempel at lengdefordelingen til reka i enkelte situasjoner kan være signifikant forskjellig først og sist på transportbåndet, hvilket tilsier at en sammensatt rekeprøve fra hele båndet er fornuftig. Dette er viktig, fordi aldersbestemmingen av reka baseres på modalanalyse av rekelengdefordelingen. Derfor er det også viktig å ha mål på hvor nøyaktige rekelengdemålingene er. Vi har foretatt en rekke statistiske undersøkelser, der flere rekemålere måler lengden til de samme rekene, slik at vi kan sammenligne målerne parvis etterpå. I de tilfeller der nybegynnerne har klemt for hardt, har de justert seg inn iløpet av toktet. Generelt ligger standardavviket til målefeilen på mellom 0.1 og 0.2 mm, mens lengdefordelingen inndeles i intervaller på 0.5 mm.

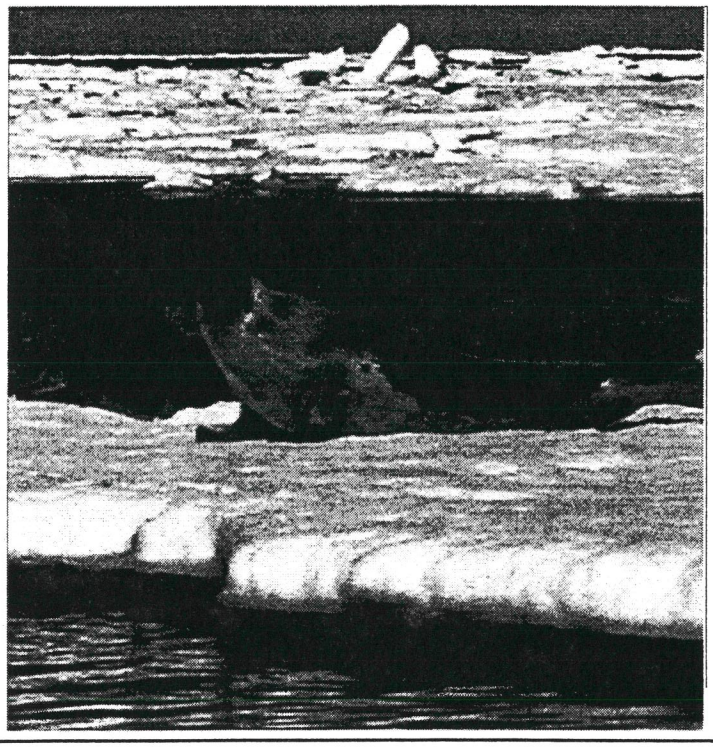
Optimalisering av toktrute

Problemstillingen har vært å bestemme den toktrute som gir best presisjon (minst cv-verdi) av totalt rekebiomasse-estimat for Barentshavet, med den tid båten disponeres som bibetingelse. Området er inndelt i 6 hovedstrata. Flaskehalsen her er fartstiden mellom trålstasjonene og at den tid som går med i et stratum ikke er proporsjonal med antall trålstasjoner (ikke-lineær kostnad). For et regulært kvadratisk grid der trålstasjonene er plassert i nodene, har vi teoretisk påvist hvordan de optimale løsningene blir. Arbeidet er publisert i Fisheries Research (1998), og deler av resultatene er implementert med hell på Barentshav-toktene, der for øvrig de estimerte cv-verdiene typisk varierer i området 7-10 %.

Det mest orginale vi har gjort er trolig innenfor subjektiv prediksjon, der vi blant annet gir en oppskrift på hvordan subjektive data formelt kan anvendes for en optimal allokering av trålstasjoner. Kort fortalt fyller både forskere, skipper og styrmann ut et prediksjonsskjema der de angir en minimumsverdi, en mest sannsynlig verdi og en maksimumsverdi for hvor mange kg reke de tror vi får opp i neste tråltrekk. I tillegg angir vår rekeforsker tilsvarende mål for midlere rekebiomasse i hvert stratum før toktet. Filosofien er at en og samme person antas å ha et "subjektivt konfidensnivå" som er rimelig konstant, dvs. en og samme person vil ligge på ca. samme score (andel prediksjoner der fasitsvaret fra tråltrekket faller innenfor intervallgrensene) fra tokt til tokt. En forsiktig person vil predikere brede intervaller, mens en "risiko-person" vil predikere smale intervaller. I tillegg forutsettes at intervallbredden justeres etter hvor god deres subjektive innsikt er. Prediksjonsforsøkene foran hvert enkelt trålhal tjener til å undersøke disse forutsetninger. Når det subjektive konfidensnivået er kjent, er neste trinn å konstruere subjektive prediksjonsfordelinger tilpasset intervallprediksjonene for hvert stratum. Disse kan så vektes med prediksjoner basert på historiske tråldata. I motsetning til aprioriforddelinger i Bayesiansk forstand, har disse prediksjonsfordelingene en klassisk, frekventistisk tolkning. Foreløpige resultater er interessante og viser et klart forbedringspotensiale ved å anvende subjektive data. De ble presentert på et internasjonalt rekesymposium i Halifax i september 1999, og vil forhåpentligvis bli publisert i rimelig nær fremtid (bidraget er akseptert).

Diettanalyse av sjøpattedyr

Multivariat analyse og romlig statistikk er nøkkelord når det gjelder diettanalyse av sjøpattedyr. Diett-dataene er preget av svært skeive fordelinger, både hva totalinnhold pr. sel eller hval og fordeling av de ulike byttedyr angår. Den relative sammensetningen blir bestemt ved å akkumulere rekonstruert mageinnhold (i kg) av hver art i teller, med akkumulert totalinnhold i nevner. Presisjonen til dette vektor-estimatet bestemmes ved bootstrap av sjøpattedyrene, slik at den innebygde samvariasjons-strukturen mellom byttedyrene bevares. De viktigste byttedyrene studeres særskilt, og vi har utviklet en generell bootstrap-basert teknikk for å konstruere multivariate konfidensellipsoider for den ukjente diettsammensetningen. Vi savner imidlertid en bedre teknikk som både reparerer for skeivhet og akselerasjon, slik Efron har utviklet i én dimensjon. Parallelt med diettdata fra kommersiell hvalfangst, har vi synoptiske ressursdata fra akustikk- og tråltokt i samme området (Øst-Finnmark 1998 og 1999) som hvalen ble fanget. Her benytter vi både geostatistikk (variogram, kriging, simulering) og bootstrap som viktige og nyttige verktøy for blant annet å oppnå presisjonsmål på arts-sammensetning. Noen resultater ble presentert på en workshop i marinrelatert, romlig analyse i Bergen våren 2000, og vil forhåpentligvis bli publisert i neste utgave av SARSIA.



Mitt mest eksotiske tokt var et seltokt rundt Spitsbergen i august 1998. Vi fartet da i gummibåt blant is, sel og isbjørn. Jeg returnerte velberget til min kone og to barn med mer enn 30 min. videoopptak av nærgående isbjørn. Blant annet har jeg en scene der en binne sitter på isen og ammer sine to unger. Hva selfangsten angår var min sønn Isak, som da var 6 år, ikke særlig begeistret for at vi skjøt disse søte dyrene. Under toktet hadde jeg følgende telefonsamtale med han:

- Jeg: Isak, kor det går med dokker?
- Isak: Har dokker skjøtte sæl?
- Jeg: Ja, men dem dør momentant og har ikkje ondt.
- Isak: Ha det, pappa.

Andre aktiviteter
Av andre aktiviteter som er utført kan nevnes 3D glatt (20 bilder pr. sek) dataanimasjon av reketetthet i Barentshavet over flere år, presentert på den internasjonale økologikonferansen INTECOL i Firenze 1998, samt en rekke andre marinrelaterte animasjoner. På trappene står blant annet mulig automatisering av otolittlesing for aldersbestemmelse og bestands-klassifisering av fisk ved hjelp av avansert billedbehandling (otolitt: ørestein med årringer). Wavelet og nevrale nett vil her trolig bli viktige verktøy.

Samarbeid med Norsk Regnesentral
Fiskeriforskning har inngått en STAT-partner kontrakt med Norsk Regnesentral (NR). Denne består i korthet i at vårt institutt betaler NR for å bistå med statistisk konsulenthjelp på timebasis, samt seminarvirksomhet. Vi har organisert det slik at statistikere fra NR kommer til oss tidlig på høsten, gjør seg kjent med de ulike sentrene, og plukker opp aktuelle data og problemstillinger å jobbe videre med. Deretter avvikles et seminar på senhøsten, hittil med simulering, forsøksplanlegging og variansanalyse som tema, der tema belyses med konkrete problemstillinger ved og data fra Fiskeriforskning. Denne virksomheten startet i 1999, der behovet for statistikkhjelp fordelt på statistikkemner, dataverktøy, og kompetansenivå (opplæring kontra konsulenthjelp) ble kartlagt i forkant ved bruk av spørreskjema. Responsen var temmelig overveldende, med ca. 40 som meddelte slikt behov, fordelt på alle sentrene og med betydelig fokus på behov for opplæring.

Hvis du vil, ta kontakt, jeg er statistikk-faglig ensom i en kjempespennende jobb.
alf.harbitz@fiskforsk.norut.no , tlf: 77629057
Mer info om Fiskeriforskning på www.fiskforsk.norut.no...

Litt om virksomheten ved Seksjon for medisinsk statistikk i Bergen

Kjære TG-lesere,

Jeg vil først fortelle om seksjonen og deretter om noen større prosjekter jeg har arbeidet med de senere årene.

Seksjon for medisinsk statistikk (SMiS) har sitt utspring i Det medisinske fakultets EDB-seksjon som ble opprettet for drøyt 25 år siden for å bistå ved oppbyggingen av Medisinsk fødselsregister. Seksjonen var og er en service-avdeling for statistikk og elektronisk databehandling for det medisinske fakultet. På slutten av 1980-tallet, stimulert av en forskningsråds-satsning ble seksjonen omorganisert og fikk to vitenskapelige stillinger, et professorat og en førsteamanuensis-stilling. Inntil 1997 var SMiS en egen enhet ved fakultetet, men er nå en seksjon i Institutt for samfunnsmedisinske fag med en seksjonsleder som utnevnes av fakultetet for å ivareta servicefunksjonen utenom instituttet.

Idag er 15 personer knyttet til seksjonen: tre professorer i full stilling, en vitenskapelig bistilling som i hovestilling er ved det nyopprettede Kompetansesenter for klinisk forskning ved Haukeland sykehus. Dette senteret og stillingen er finansiert av Norges forskningsråd. Seksjonen har en statistiker i fast konsulentstilling som skal drive service og veiledning for fakultetets forskere. Videre har SMiS en ingeniørstilling knyttet til IT-oppgaver, en halv kontorstilling og en prosjektfinansiert statistiker. Syv stipendiater har hovedveileder ved seksjonen - alle med epidemiologiske problemstillinger. Ytterligere drøyt 10 stipendiater har bi-veileder ved SMiS.

En viktig oppgave for seksjonen er undervisning. Undervisningen er i hovedsak knyttet til medisinerstudiet og ukes-kurser under Legers videre og etterutdannelse. I medisinerstudiets første år gir seksjonen et klassisk introduksjonskurs i medisinsk statistikk for 150 medisinerstudenter over ca 25 forelesningstimer. Kurset avsluttes med egen 2 timers eksamen.

Seksjonen underviser også i medisinerstudiets 5.år - i psykiatriterminen. Dette kurset har som eneste mål at studentene gruppevis skal skrive og levere inn en protokoll for et randomisert klinisk forsøk. Undervisningen er mindre formell enn i det første året, og omfatter teknikaliteter knyttet til randomisering, forsøksstørrelsesberegning og etiske aspekter knyttet til pasientforskning (Helsinki-deklarasjonen, etiske komiteer i Norge etc).

Innen Legers videre og etterutdannelse gir vi vanligvis fire kurs årlig: Et introduksjonskurs (Medisinsk statistikk I), et videregående kurs med spesiell vekt på regresjonsmodeller (Medisinsk statistikk II) og to kortere introduksjonskurs til SPSS. Kursene er mer populære enn noensinne og er stadig overtegnet. Den økte interessen skyldes i stor grad den gledelige utviklingen at andre profesjonsgrupper i helsevesenet enn legene i større grad tar fatt på forskningsprosjekter som krever statistikkunnskap.

En annen viktig funksjon for seksjonen er veiledningstjenesten som drives på to nivåer. Vi har en vaktordning hvor alle statistikere knyttet til miljøet (inkludert

Medisinsk fødselsregister, Hofteregisteret og Yrkesmedisinsk avdeling ved Haukeland sykehus) tar imot forespørsler og håndterer enkle problemstillinger som kan avsluttes innen 1-2 timer. I 1999 hadde vi drøyt 60 loggførte henvendelser til denne generelle veiledningstjenesten. Spørsmålene her dreier seg ofte om valg av statistisk metode og tolkning av utskrifter fra statistiske programpakker. I tillegg yter seksjonens faste statistiker mer langvarig og omfattende støtte til et mindre antall prosjekter.

Seksjonens forskningsportefølje er i hovedsak epidemiologisk, og ikke metodologisk, og har i de senere årene fått et betydelig omfang. Seksjonen har et omfattende samarbeid med Medisinsk fødselsregister og alle i vitenskapelig hovedstilling har bistilling ved dette registeret. Spesielt kan nevnes en stor og mangeårig befolkningsbasert pasient-kontroll studie på leppe-gane spalte som skjer i samarbeid med og mottar betydelig økonomisk støtte fra National Institute for Environmental Health i USA.

Personlig har jeg på 1990-tallet vært involvert i oppbyggingen/utnyttelsen av flere større befolkningsbaserte studier eller registre. De som foreløpig har gitt størst publikasjonsmessig uttelling er Homocystein-undersøkelsen i Hordaland (The Hordaland Homocysteine Study) og Hofteregisteret eller Nasjonalt register for leddproteser som nå er den offisielle betegnelsen. Mye tid og krefter har også gått til planlegging og forberedelse til den store norske Mor og Barn kohorten som Medisinsk fødselsregister prøver å etablere i samarbeid med Folkehelsa. Planen er å rekruttere 100,000 gravide og samle inn omfattende livsstilsinformasjon (inkludert kosthold) pluss blodprøver. Målet er å relatere denne informasjonen til senere sykdom og helseproblemer. Studien har vært møtt med skepsis og er fortsatt kontroversiell, men et pilotstudie pågår i Hordaland og slik det ser ut idag vil den gradvis utvides i årene fremover.

"The Hordaland Homocysteine Study" ble til på begynnelsen av 1990-tallet og hadde sitt utganspunkt i et fremragende laboratoriemiljø ved Farmakologisk institutt som hadde utviklet målemetoder for homocystein. Homocystein er en svovelaminosyre som kan måles i blod og som siden 1960 tallet har vært foreslått som en risikofaktor (og mulig kausal faktor) i utvikling av hjertekarsykdom. Homocystein-nivået er metabolsk knyttet til omsetning og inntak av B-vitaminer, først og fremst folat og vitamin B12 (kobalamin), men også vitamin B6 og vitamin B2 (riboflavin). Teorien var uten tilknytning til den etablerte fett/kolesterol teorien som tradisjonelt forklarer hvorfor hjertekarsykdom oppstår og møtte i tidlig fase mye motstand og skepsis. I samarbeid med Statens helseundersøkelser (SHUS) ble så den største befolkningsundersøkelse noensinne (og fortsatt) gjennomført i 1992-93. I tillegg til SHUS sitt tradisjonelle hjertekar-batteri ble homocystein målt på i overkant av 18,000 kvinner og menn i alderen 40-67 år. Dette ga oss muligheter til å studere og påvise hittil ukjente sammenhenger mellom homocystein og livsstil. Vi kunne for eksempel vise at personer som drakk mye kaffe, som røyker og som spiser lite frukt og grønnsaker eller ikke tar vitaminpreparater har betydelig høyere homocystein-nivåer enn andre. I de senere årene har vi også studert sammenhenger mellom homocystein og tidligere svangerskapsutfall hos kvinner som deltok i undersøkelsen og som hadde sine svangerskap registrert i Medisinsk fødselsregister. Her fant vi til dels sterke sammenhenger mellom homocystein og en rekke helseproblemer knyttet til svangerskap og fødsel. Dette bidrar til den økende forståelsen for at B-vitaminer er

viktige også i det som er blitt betraktet som velernærte befolkninger. Vi ser også på sammenhenger mellom dødelighet i kohorten og homocysteinnivå. Om noen år planlegger vi en kopling mot Kreftregisteret og vil da være den første befolkningsbaserte prospektive studien som ser på sammenhenger mellom homocystein og kreftinsidens. Dette vil bli spennende analyser spesielt sett i lys av at man på 1990-tallet har oppdaget at B-vitaminer og spesielt folat kan spille en større rolle i kreftutvikling enn tidligere antatt. Nylig er datainnsamlingen for fase II av Homocystein-studien i Hordaland avsluttet. Det ble da tatt en ny blodprøve og innhentet mer detaljerte opplysninger om kostholdet (i samarbeid med Institutt for ernæringsforskning ved Universitetet i Oslo) på rundt 6,000 av de personene som deltok i 1992-93.

Hofteregisteret er en annen prospektiv studie hvor SMiS har bidratt metodologisk. En statistiker ved SMiS tok for et par år tilbake sin dr.philos grad knyttet til data fra Hofteregisteret og nr. to er godt i gang med sin. På initiativ fra Norsk ortopedisk forening har hofteprotesekirurgene i Norge siden 1987 rapportert karakteristika ved det kunstige hofteleddet, sementeringsteknikk, dato og indikasjon for inngrepet. Et kunstig hofteledd skal i prinsippet kunne fungere problemfritt i 20-25 år, men hos en del pasienter må man fjerne det kunstige leddet og sette inn et nytt på et langt tidligere tidspunkt. Dette kan skyldes at det kunstige leddet løsner (hyppig sviktårsak) eller infeksjon (sjelden komplikasjon). Hofteregisteret har vist at enkelte typer proteser har langt dårligere egenskaper enn andre, og Hofteregisteret avslørte at en ny type sement (Boneloc) som ble introdusert tidlig på 1990-tallet hadde katastrofalt dårlige egenskaper i forhold til de tradisjonelle sementtypene. Denne sementtypen brukes nå hverken i Norge eller andre land. Et problem i protesekirurgien, og generelt når det gjelder medisinske implantater, er at myndighetene hverken i Norge eller andre land krever at klinisk effekt og sikkerhet dokumenteres før markedsføring. Dette står i skarp kontrast til krav som stilles til nye medikamenter.

Jeg har nå gitt en oversikt over seksjonen og noen detaljer omkring et par løpende forskningsprosjekter jeg har vært involvert i. Jeg har ikke gått inn på statistikkbruk eller statistisk håndtering av kompleksiteten i datasettene. Her er det arbeid for mange statistikere med interesse for praktisk dataanalyse. For eksempel er det avhengighetsstrukturer i dataene som kan angripes på mange måter. I fødselsregisteret bør man modellere avhengighet mellom svangerskap knyttet til samme mor. Det finnes nå også data over generasjoner og far kan trekkes inn i analysen. Her ligger det store metodologiske utfordringer. I Hofteregisteret kan samme pasient ha en historie med flere hofteimplantater både i høyre og venstre hofte. Full modellering av dette krever bruk av komplekse modeller fra forløpsanalyse-feltet.

Med hilsen

Stein Emil Vollset

Etablering av Kompetansesenter for klinisk forskning (KKF) ved Haukeland sykehus i Bergen.

Som et ledd i Norges forskningsråds satsing på medisinsk forskning er det opprettet et Kompetansenter for klinisk forskning (Centre for Clinical Research) ved Haukeland sykehus. Tilsvarende satsing gjøres også i Trondeim, Tromsø og Oslo. Sentrene vil de første årene finansieres av forskningsrådet, men skal etter hvert overtas av regionssykehusene.

I Bergen er der 5 faste stillinger:
forskningsleder Ernst Omenaas,
biostatistiker Geir Egil Eide,
statistisk konsulent (under besettelse),
IT-konsulent Gary John Allan,
og senterkoordinator Linda Stoltz Olsvik.
De 2 førstnevnte er også knyttet til Universitetet i Bergen ved bistillinger.

Som første ansatte 15. august 1999 fungerte undertegnede i startfasen som forskningsleder. Frem til april 2000 gikk mye tid med til å vurdere søkere til øvrige stillinger, vurdere lokaler, budsjettering, og bestille datautstyr og programvare med base fra mitt gamle arbeidssted på Norges handelshøyskole. Senteret flyttet inn i nyoppussete lokaler i 3.etg. i Armauer Hansens hus i april/mai 2000, og er geografisk plassert nært Seksjon for medisinsk statistikk ved Universitetet i Bergen hvilket vi også anser som svært nyttig rent faglig.

Kompetansesenteret tilbyr veiledning til alle helsefaggrupper ved Haukeland sykehus og samarbeidende sykehus i Helseregion Vest i forsøksplanlegging, datainnsamling, data- og statistisk analyse og publisering. Senteret vil også initiere og bidra til forskningsmetodisk utvikling og undervisning. Etter etableringen har det vært enorm pågang fra kliniske forskningsprosjekter som ønsker assistanse fra senteret og antallet registrerte prosjekter nærmer seg raskt 50. Senteret har også utlyst 6 kliniske forskerstipend av 3 mnds varighet hvorav 3 benyttes i høst og 3 i år 2001. Stipendiatene tilbys arbeidsplass og dataressurser ved senteret til å gjennomføre hele eller deler av egne prosjekter av begrenset varighet og omfang. For å avhjelpe situasjonen har senteret også engasjert midlertidig to dataanalytikere.

Senteret initierer og arbeider også med egne prosjekter av metodisk karakter, bl.a. vil undertegnede arbeide videre med metodikk for tilskrivbare andeler og levetidsanalyse. Forskningslederen, som er medisiner, har plass i sykehusets utvidete styre og vil ha utredende og innstillende funksjon i forhold til forskningsmidler over sykehusets og helseregionens budsjetter.

NFRs nysatsing på klinisk forskning vil trolig få positive ringvirkninger m.h.t. flere interessante arbeidsplasser for statistikere i årene fremover, ikke bare i Bergen, men også rundt de øvrige miljøene som er etablert i Oslo, Trondheim ogTromsø.

Foreløpig har senteret kun hjemmeside som kan nås via Haukeland sykehus' interne nett: http://intraweb.haukeland.no/kkf/, men det arbeides med å få til en hjemmeside tilgjengelig også for utenforstående.

Bergen 15. september 2000

Geir Egil Eide
biostatistiker/førsteamanuensis

På faglig besøk i Berkeley

Jeg er tilbake etter nærmere ti måneder som "visiting scholar" ved Department of Statistics, Berkeley. For tiden er jeg ved matematisk institutt (UiO) hvor jeg er i ferd med å avslutte som doktorstipendiat. Jeg hadde selvfølgelig hørt mye om Berkeley i forkant, så forventningene var skrudd temmelig høyt. Berkeley var først og fremst annerledes enn det jeg hadde forestilt meg. Den matematiske statistikken regnes for å stå sterkt i Berkeley, men den tid er stort sett forbi (et forbehold: Bickel og Doksum bør vel fremdeles regnes som representanter for denne retningen). Derimot hadde Berkeley et stort miljø for anvendt statistikk og et forholdsvis lite men sterkt miljø for sannsynlighetsteoretikere. Gamle helter som LeCam, Blackwell og Lehmann var tilstede, men de var ikke lenger tonangivende for avdelingen.

Berkeley ligger like ved San Francisco. Området består av en lang rekke større og mindre byer, som går under fellesnavnet Bay Area. Her er det omkring 8 millioner innbyggere, hvorav 10% er av asiatisk opprinnelse. China Town er kjent og er selvfølgelig verd noen besøk. Ellers har Bay Area et rikt kulturliv. Jeg fikk med meg noen utstillinger og konserter. Av de større begivenhetene må jeg nevne Mike Stern Quartet og en konsert med Mayhem. Sistnevnte er et norsk black metal band. Noen fæle typer med andre ord, men jeg ble ikke så veldig skremt. Jeg leide hovedleiligheten i et stort hus i El Cerrito, et typisk middelklasse strøk utenfor Berkeley. BARTen (lokaltoget) gikk bare noen få steinkast unna med direkte tilgang til Berkeley (6~min) og SF (35~min). Very convenient. Det var også et kjøpesenter i nærheten så jeg klarte meg fint uten bil.

Jeg så fram til å treffe Lucien LeCam. Jeg har vært opptatt av den asymptotiske desisjonsteorien han har utviklet. Blant annet har jeg laget en distanse for filtrerte eksperimenter som kan betraktes som en naturlig generalisering av LeCam's defisiensdistanse. LeCam har ofte brukt et litt spesielt matematisk språk som vi kaller *lattice theory*. Innenfor denne matematiske rammen finner vi løsningen på statistiske anomalier (for begrepet suffisiens) som Pitchers Patologi og Burkholders Patologi. Men det er dessverre få som virkelig behersker lattice-teorien. "Forsatt kan folk ikke tilgi meg at jeg introduserte lattice theory", sa LeCam fornøyd. På spørsmål om hvordan han hadde fått ideen til lattice-teorien, svarte han: "I should not tell this, you might think I am bad. I was trying to make a distance between experiments by means of risk functions. I could not solve it. For two years I was working on the problem. One morning I looked at it and... Ah, it was all clear! It should be considered as a linear problem!" LeCam var gammel og syk da jeg traff han. Han hadde problemer med leveren og hadde blitt sterkt redusert de to siste årene. Lucien LeCam døde 25. april 2000.

Uten en PhD blir man betraktet som faglig paria. Jeg ble f.eks. aldri spurt om å holde foredrag, så jeg fikk en følelse av å være litt utenfor det "gode selskap". En post doc. fortalte meg til og med at han ikke skulle snakke med meg fordi jeg ikke var doktor...(!). Men, det var ikke bare ille, de fleste jeg traff var sympatiske. Best kjent ble jeg med andre besøkende som jeg delte kontor med. Vårsemesteret delte jeg kontor med Jiayang Sun. Hun kommer opprinnelig fra Kina, men har etablert seg ved et universitet i Ohio. Vi gikk stort sett på de seminarene som var, og diskuterte av og til foredragene i etterkant. På slutten av semesteret delte vi også kontor med Aad van der Vaart fra Nederland. Aad er dyktig, og har god kjennskap til LeCams teori. Vi kom opp i en liten disputt om hvilke typer statistiske "løsninger" som er akseptable. Egentlig syntes jeg ikke argumentene mine var særlig overbevisende, og det fikk meg til å tenke gjennom problemene på nytt.

Det var først og fremst ved å følge seminarene at jeg fikk innblikk i den faglige virksomheten på Berkeley. Det var tre seminarer i uken: Neyman-seminaret, statistikk-seminaret og probability-

seminaret. Neyman-seminaret hadde egentlig lite med statistikk å gjøre. Her ble det invitert forelesere fra andre fagfelt som for eksempel finans, musikk, matematikk, biologi osv. Foreleserne prøvde som regel å lage en liten ``statistisk" vri på slutten av foredraget (mest for å rettferdiggjøre at man hadde snakket om helt andre ting i nesten en time). Statistikk-seminaret hadde av og til noen foredrag som var interessante. Men en stor andel av foredragene var hentet fra anvendelser, med en ordinær modell og noen p-verdier. Slike foredrag er kjedelige, spør du meg! Probability-seminaret, som ble ledet av Aldous, var det seminaret som var morsomst. Likevel, det var ikke midt i blinken for min del. Jeg er først og fremst interessert i statistikk (hvor statistikeren skal gjøre noe som f.eks. estimere en parameter, treffe en beslutning, utføre en test, osv.). Denne gjengen her var renmatematikere. Jeg lærte litt om Galton-Watson-trær av en post doc. som besøkte Aldous. Jeg tror ikke jeg får brukt den lærdommen noen gang, men man skal aldri si aldri...

Jeg traff flere nordmenn i Berkeley. Steinar Bjerve delte jeg kontor med de første ukene jeg var på Berkeley. Jeg traff Kjell Doksum. Han hadde forskningsfri mesteparten av tiden jeg var på Berkeley, og var stort sett ved et annet universitet. Bo Lindqvist stakk innom en dag, og jeg traff også Ragnar Norberg (min far) som var noen måneder ved Stanford. Jeg besøkte han et par uker i Palo Alto og fikk sett litt av Stanford, som er et praktfullt universitet.

Jeg hadde et flott opphold. Jeg føler at jeg har lært en del og fått et bedre perspektiv på ting. Ikke minst lærer man om hjemlige forhold ved å ta en reise som dette. Matematisk institutt (UiO) er av verdensklasse (fikk nylig Marie Curie pris). Berkeley er jo også veldig bra... Nevnte jeg at Berkeley har et fantastisk bibliotek?

Espen Norberg

REISEBREV FRA GRIMSTAT !!!

NORDSTAT 2000 - et program med stor spennvidde arrangert i en sørlandsby i begynnelsen av juni - det er fristende for en statistiker med røtter i Tønsberg men bosatt i Trondheim. Etter en mager vår her nord, satte jeg, sammen med en rekke andre trøndere, kursen mot Grimstad med forventninger om faglige utfordringer & inspirasjon samt om sol & sommer. La det være sagt med en gang - vi fikk mye av førstnevnte, men trønderværet fulgte med på ferden.

Den 18. nordiske konferanse i matematisk statistikk ble arrangert på Sørlandet hotel & kurssenter i Grimstad fra 4. til 8. juni år 2000. Ikke akkurat i sørlandsk arkitektonisk idyll, men i svært funksjonelle omgivelser. Det bør nevnes at det var en ypperlig gressfotballbane mellom seminarsalene og hotellet. De sagnomsuste Amasonene fra Grimstad trente forøvrig der en ettermiddag.

Jeg ankom sent søndag kveld og morgenen etter braket det løs. `Braket løs' er vel ikke helt beskrivende. Vi statistikere har ikke helt sansen for pomp & prakt. Ingen fanfarer eller kunstneriske innslag her - få paralleller til OL-åpningen. En velkomsttale av en uomtvistelig patriotisk ordfører kastet imidlertid glans over åpningen.

Arrangørene hadde gjort en grundig jobb og presenterte et mangfoldig program. Jeg merket meg spesielt skille mellom "Invited lecurer" og "Specially invited lecturer". Uklart for meg av hvilken grunn, men spekulerte litt på om sistnevnte også fikk dekket reisen. Jeg ble revet ut av mine spekulasjoner da Eilev S. Jansen fra Norges Bank presenterte deres RIMINI (få endelig med alle stavelsene) modell. Og da skjønte jeg at reisepenger og denslags bare er som småpenger å regne. Et fengende foredrag hvor statistikkens betydning i makroøkonomisk planlegging ble frenhevet. En sal fylt av statistikere lente seg selvtilfreds tilbake.Troen på statistikkens fremtid ble ikke svekket av Nils Gustafsons foredrag om værvarsling. Utfordringene synes å stå i kø. Trusler fra kaosteori og uforutsigbarhet bør ikke skremme oss statistikere.

Så den første lunsj. Jeg valgte en kaloriminimerende versjon og tok meg en løpe-sightseeing. Badestranden nedenfor ville vært innbydende hadde det vært finvær - men Grimstad by er sjarmerende selv i regn.

Ettermiddagen startet med Stratis Gavaris' foredrag om estimering av fiskebestand. Nordlendingene i salen synes noe brydd over at en Canadier med et greskklingende navn var hentet inn for å belære dem om "fæsk". Men her var det mye å lære selv for dem som var født med sjøstøvler på beina. Foredraget ble uforglemmeliggjort av begrepet "piscometrics" innstiftet av David Hinkley i den påfølgende diskusjonen. Sesjonen ebbet ut med en paneldebatt, som det alltid er vanskelig å gjøre fargerik. Intet unntak her.

Etter kaffi med velsmakende frukt tok parallellsesjonene til. Det gir alltid valgets kval. Som professor med fortid som avisbud, skulle jeg velge foredrag om "Probability and statistics in secondary school" eller om "Optimal delivery of newspapers". Jeg valgte førstnevnte, samt et fordrag til. Første dag over.

Tirsdag morgen styrket Thore Egeland selvinnsikten med sitt foredrag om genetikk. Deretter fulgte parallellsesjoner. Jeg hadde omhyggelig planlagt et slalomløp gjennom de tre sesjonene. Men når hver sesjon fremstår som en stokastisk prosess med tilfeldige start/slutt-tider og rekkefølge på

foredragene vil selv en Kjuus komme til kort. Selv statistikere burde verdsette determinisme og forutsigbarhet i visse sammenhenger. Rom for forbedringer på dette punkt.

Etter lunsj var det båttur - intet værforbehold. Menyen var reker, øl og hvitvin. Rekeforsker Alf Harbitz holdt et stuntkåseri om rekas syv stadier inkludert minst ett kjønnsskifte. Verden er mangfoldig - homofilidebatter i kirke og OL kunne vært unngått hadde vi alle vært reker. At været var slik at rekene var i sitt rette element også oppå dekk, la nok en liten demper på humøret.

Onsdagen opprant, atter en sesjon med "Specially invited lecturers". De skuffet ikke. Olle Häggstrøm doserte om Markov-stokastiske felt. Deretter var det dobbel sesjon med Murad S. Taqqu. Inspirert av sendetid for informasjon i datanettverk utledet han en modell for stokastiske skalafenomen. Erlang ser sin begrensning i telefoni, i dataalderen er det andre lover som gjelder. Et fascinerende foredrag.

Lunsj og konferansens ubestritte høydepunkt: Fotballkamp. Tradisjonelt har det vært vertsnasjonen mot resten av Norden. For få av sistnevnte møtte opp. Vi prøvde Norge mot resten av verden. Det gikk heller ikke. Løsningen ble Trondheim+ mot resten av verden. En jamn kamp, hvor en italiener, en danske samt en aldrende Oslo-statistiker med glitrende teknikk men sviktende fremdrift, tidvis satte støkk i nordboerene. Det må gjøres noe med rekrutteringen ikke bare til statistikerstudiet, men også til fotballen.

Om kvelden var det festmiddag, eller "Banquet", som det stod i programmet. Pyntet bord - feststemte statistikere. For et syn. Lang taleliste - en fra hver av de nordiske land, tradisjonen tro. Stor hyldest ble "The lone Swede", Lennart Bodeson, til del. Svenskene var dramatisk underrepresentert på konferansen, men regner å ta det tapte igjen ved neste korsvei. Da arrangeres ganske riktig konferansen i Sverige. Etter Pære "Belle Helene" og kaffi, trekker den feststemte forsamling ned på Lilly's og fortsetter festen til ut i de små timer.

Siste konferansedag: Torsdag. Parallellsesjoner frem til lunsj. Deretter foredrag av Gianpaolo Scalia Tomba, som utseendemessig ikke er ulik sin navnebror. "Statistics for infectious diseases", et besnærende foredrag. Men det burde vel være "against". Så til slutt, Nils Lid Hjort. Full sal selv på siste foredrag, bra! Et foredrag med patos, hvor de store historiske linjer ledet oss frem til en frekventistisk forsvarstale. Men husk nordiske statistiker-kolleger - det er en liten sosial-demokratisk bayesianer i oss alle.

Så var det slutt, vi reiste hjem, og alle var enige om at det hadde vært en fin konferanse.

Henning Omre

SOM "LÆRLINGER" TIL GRIMSTAD.

En dag i februar/mars ble noen studenter oppmerksomme på at det skulle avholdes en nordisk konferanse i statistikk i Grimstad begynnelsen av juni. Flere av oss snakket om å ta en tur, men i og med at de fleste nærmer seg slutten av studiene og begynner å få et litt mer anstrengt forhold til Lånekassen, ble det først fart i planen da vi fikk høre at det var mulig å søke om støtte til deltagelsen.

Slik gikk det til at seks studenter møttes på Busstерminalen i Oslo en pen søndags formiddag i juni for å ta Sørlandsekspressen til Grimstad. Den fem timer lange turen i dobbeldekkeren med vaffel- og kaffeserverende vertinne ga de fem hovedfagsstudentene fra Oslo en god mulighet til å bli kjent med det sjette, eksotiske innslaget i seksmannsstroppen - en siv.ing student fra Trondheim. Vi hadde alle våre forventninger til hvordan det ville bli å tilbringe fire dager på Sørlandet sammen med så mange dyktige og "store" statistikere. Hvor mye ville vi egentlig få med oss av det som ble sagt? Kom vi til å bli sittende alene og forskremte i en krok?

Mandag morgen troppet vi opp til første foredrag med skjerpet konsentrasjon og renvaskede ører. Heldigvis startet konferansen med eksempler fra anvendelser innen økonomi, værvarsling og fiskeri. Disse foredragene var det mulig for en stakkars student å følge relativt godt med på, så motet steg hos oss alle. Nå skulle det snart vise seg at ikke alt var like lett å få med seg, for noe ble vel utfordrende for oss. Når formlene ble mange og notasjonen vel infløkt, var det lett å ramle av. Men heldigvis fikk vi fort trøst av andre, langt mer kunnskapsrike deltakere. De forsikret oss om at det ikke var å forvente at vi skulle klare å henge med på alt, og at de til og med hadde problemer med det selv. Om alt var sant, eller om det var en pen måte å peppe oss opp på, skal være usagt. Men det hjalp uansett både på motet og selvtillien! Hele konferansen skulle foregå på engelsk, og det var en interessant utfordring for oss tilhørere. Det var tydelig at våre nordiske statistikere har et stort engelsk vokabular. Likevel var det ikke så vanskelig å gjette seg til hvilket land de ulike foreleserne kom fra, og det var morsomt å høre engelsk ispedd ulike nordisk aksenter. Ellers var foredragsholderne jevnt over veldig gode - men hvor ble det av kvinnene? Hver dag prøvde vi å holde en opptelling av hvor mange damer vi hadde fått høre prate, og det var ikke mange. Men blant oss seks som reiste sammen ned til Grimstad, var det faktisk fem jenter - og vi registrerte at det var flere jenter blant Dr. ing.-studentene fra Trondheim. Så kanskje det er håp om en bedring snart?

Konferansen i Grimstad handlet ikke bare om de faglige foredragene. De andre arrangementene ga oss en god anledning til å studere hvordan en statistiker oppfører seg i sosiale sammenhenger. Banketten ga oss et klart bevis på at også forlengst ferdigutdannede statistikere trives godt med en fest av og til. Og selv om arrangementskomiteen hadde glemt å

bestille sol og pent vær til båtturen i Grimstads skjærgård, fikk vi her testet ut de andre konferansedeltakernes evne til å prate om annet enn statistikk. Været er jo alltid en god samtaleåpner, og når et blidt båtmannskap serverte loff, reker og vin, gikk praten lett. Blant annet benyttet vi anledningen til å bli mer kjent med våre danske medstudenter. Det viste seg at danskene hadde syklet fra Danmark til Grimstad, og at de bodde i telt pga. dårlig økonomi. Vi følte oss nokså priviligerte som fikk bo på hotell!

Ellers deltok vi jo ikke bare på de "offisielle" sosiale arrangementer. Grimstad sentrums restauranter, cafeer og barer fikk noen besøk av oss i løpet av turen - både med og uten "de voksne". Alle pausene som var lagt inn i konferanseprogrammet fungerte også som "sosiale pusterom", og vi skjønte fort at vi ikke var de eneste som var takknemlige for denne sjansen til å fylle på med koffein og oksygen mellom slagene. Vi må jo tilstå at vi var litt skeptiske til tanken på de lange dagene med foredrag. Som hovedfagsstudenter er vi ikke vant til å ha så mange timer på en dag. De rause lunsjpausene ga også mulighet for å få avspilt en fotballkamp. Planen var å spille Norge mot resten av verden, men da bare en italiener og to dansker stilte opp, ble det endring i planene. Det endte opp med et Dr.ing-lag med forsterkninger mot røkla. Det må nok innrømmes at Dr. ing-laget imponerte mest med sin løpsvillighet og scoringsiver, men hardt spill og stygge tacklinger trakk ned. Et morsomt avbrekk i konferansen var det iallfall, og vi oppfordrer alle som har tenkt seg på framtidige konferanser til å huske på å ta med seg treningstøy, fotball sko - og dersom det er mange NTNU-studenter med, bør også leggskinn og nesebeskyttere vurderes!

I løpet av konferansen rakk vi også å få med oss årsmøtet i NSF. Der ble det bla. vedtatt å fortsette å gi støtte til studenter som ønsker å delta på slike konferanser. For å avslutte denne "reiseskildringen", ønsker vi til slutt å si at vi syns det er flott at NSF er villig til å støtte oss og nye studenter på denne måten. Uken i Grimstad var til inspirasjon og motivasjon for oss alle. Vi følte oss hele tiden veldig velkomne. Selv om vi langt fra forstod og kunne henge med i alt som ble sagt, så var det spennende å få være med og få et lite innblikk i hva som skjer ute i statistikkens verden. Inntrykket vi sitter igjen med er at det foregår mye - og at det er muligheter innen vårt fag. Vi syns i grunn vi lærte mye. Ikke minst var det deilig å merke at vi gjennom studiet hadde lært om, eller i det minste hørt om, en god del fra før.

Så takk for støtten til turen, og takk for en fin konferanse!

For sekserbanden,
Gunnhildur Hognadottir Steinbakk,
Birgitte Ringstad Vartdal og Ane Seierstad

Puslerispalten

Pusleri nr. 4

Auksjoner kan foregå på flere måter, bl.a.

(1) Bud høynes inntil ingen flere bud kommer. Siste budgiver betaler budt beløp ("First prize - open cry", kalt engelsk auksjon)

(2) Alle budgivere byr i lukket konvolutt. Høyeste budgiver vinner og betaler budt beløp ("Sealed bid - first prize")

(3) Auksjonarius setter pris over høyest tenkelige og senker priser inntil første budgiver slår til (kalt hollandsk auksjon).

(4) Som (2), men høyeste budgiver betaler nest høyeste bud. ("Sealed bid - second prize").

Det er nokså lett å innse at (2) og (3) er ekvivalente.

Vis (under rimelige forutsetninger) at (1) og (4) er ekvivalente.

Aha! I valget mellom de to vanlige auksjonstypene (1) og (2), må derfor (2) være bedre for selger enn (1), fordi (2) gir det høyeste lukkede bud, mens (1) gir beløp ekvivalent med nest høyeste lukkede bud. Sant eller galt?

Hint: Hvis du ikke greier å avgjøre generelt det siste spørsmålet, som ga William Vickrey nobelprisen i økonomi, kan du som statistiker få lov til å tenke deg n personer som vurderer auksjonsobjektet iht. uavhengige trekninger fra den rektangulære fordeling [0,1].

Problemstillingen kan sikkert belyses på ulike måter, og *også 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. 3

Hvorfor går bussene oftere i motsatt retning?

Vi minner om oppgaven:
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 ?

Løsninger: Det kom inn innspill fra Geir Arne Kjønstad og Trond Haider. Her er de i den rekkefølge de kom inn:

Geir Arne Kjønstad:

Intuitiv løsning: Det er klart at sannsynligheten for at det kommer ***en*** buss i motsatt retning før din kommer er 0,5, men det kan jo også komme flere busser i motsatt retning før den første kommer i din retning. Dermed blir gjennomsnittlig antall over 0,5 for antall som kommer i motsatt retning før første buss i din retning.

Matematisk løsning:
P(din buss er først) = 0,5 <=> P(null busser før din første) = 0,5.
P(en eller flere busser før din første) = 0,5

Gjennomsnittlig antall busser per dag før din første vil etter ei (lang) stund konvergere mot forventet antall,

E(antall busser før din første)
= (null busser før din første)*P(null busser før din første)
+ (en eller flere busser før din første)*P(en eller flere busser før din første)

$= 0*0,5 + (\Sigma\, xP(X = x), \forall x \in \mathbf{N})*0,5$

der X er antall busser som kommer før din og $P(X=x) = 2^{-x}$. Forventningen blir videre:

$= (1*2^{-1} + 2*2^{-2} + 3*2^{-3} + 4*2^{-4} + \ldots)*0{,}5 = 2*0{,}5 = 1$

Så ved å fortsette å registrere gjennomsnittlig antall busser før den første i din retning, vil du etter hvert nærme deg 1.

Trond Haider:

Det finnes i alle fall tre forklaringer på det observerte misforhold, med vidt forskjellige bakgrunner:

Den ene forklaringen er parallell til forklaringen på følgende pusleri, dog med noe mer tilfeldig variasjon siden det dreier seg om buss:

" En mann hadde to elskerinner. Hver kveld dro han hjemmefra for å besøke en av dem. Han lot skjæbnen, i dette tilfellet Oslo Sporveier, bestemme hvem av dem han skulle besøke, ved å ta den av de to T-banene som førte ham til sine elskerinner som kom først. Hver av disse T-banene gikk 4 ganger i timen. Likevel viste det seg at over en tid hadde han besøkt den ene av sine elskerinner 65% av kveldene, mens den andre bare ble besøkt i 35%. Denne mannen hadde dessverre ikke hatt hverken statistikk i videregående eller besøk av Ørnulf B. på ungdomsskolen, så han forsto ikke helt hvordan dette kunne skje."

Den andre forklaringen bunner rett og slett i at denne mannen drar på jobb på et tilfeldig tidspunkt i rushtiden i rushretningen. Bussen han skal ta blir mye oftere påvirket, og forsinket, enn de bussene som går andre veien. De vil ha en jevnere strøm fordi de kom seg forbi motsatte (mannens) holdeplass før rushtiden satte inn. Siden mannen går til holdeplassen nokså, men ikke helt, tilfeldig, vil det derfor være færre busser i ens retning enn i motsatt.

Den tredje forklaringen er simpelthen at det går to bussruter forbi hans stoppested, men at han bare kan ta én av dem.

Kommentar fra Lillestøl:

Jeg takker for de to innsendte forklaringene. Jeg velger (muligens feilaktig) å oppfatte Trond dithen at han søker etter mulige årsaker til et observert misforhold. Hans refleksjoner er i så måte interessante, og kan brukes i pedagogisk sammenheng. Spørsmålet er imidlertid om det observerte er et misforhold.

Geir Arnes resonnement var i hovedsak det jeg hadde i tankene, men hans oppstilling gir anledning til kommentarer som kan ha pedagogisk interesse. I beregningen av E(antall busser før din første) er den første likheten ikke formelt

riktig: (en eller flere busser før din første) er ikke en verdi (men flere) som ikke kan veies direkte med en sannsynlighet. Det er vel setningen om dobbelt-forventning det legges opp til, ved å betinge mhp om min buss kommer først eller ikke: $E(X)=E(X|B=0)P(B=0)+E(X|B>0)P(B>0)$.
Denne setningen er svært intuitiv, men ikke kjent av alle med elementære kunnskaper i sannsynlighetsregning (som 2MX/3MX i videregående skole). Man trenger den heller ikke, men kan gå rett på oppstillingen $EX=0(1/2)+1(1/4)+2(1/8)+3(1/16)+.... =1$, og så argumentere for at $P(X=x)=1/2^{x+1}$. I denne argumentasjonen vil betinget sannsynlighet og upårvirket gjenstående ventetid måtte inngå, dvs. formalisering av de upresise utsagn i oppgaven om at både bussene og jeg innfinner oss tilfeldig.
Hvis vi tillater oss å bruke setningen om dobbeltforventning, er det en enklere oppstilling $EX=0.5\times0 + 0.5\times(1 + EX)$ som løst mhp EX gir 1, der vi har brukt et "fornyelsesargument" forankret i nevnte tilfeldigheter.

Jeg kan også forestille meg en mer generell løsning, der under visse forutsetninger $EX= (\mu^2 +\sigma^2)/2\mu^2$. Her vil spesialtilfellene at alle busser er i rute uten variasjon $\sigma= 0$ gi $EX=1/2$ og ren tilfeldighet dvs. eksponensiell fordeling med $\sigma=\mu$ gi $EX=1$. Et formelt bevis for dette kan skje ved fornyelsesteori og evt. ergodiske argumenter.

ERLING SVERDRUPS RADIOFOREDRAG FOR 45 ÅR SIDEN

I serien "Universitetenes radioforedrag" bidro Erling Sverdrup med tre foredrag under tittelen "Hva kan vi slutte av statistikk?". Manuskriptene til disse foredragene ble funnet da Knut Håkon Brox ryddet i Sverdrups etterlatte papirer i fjor høst. Det første av foredragene ble trykket i forrige nummer av TG. På de neste sidene i dette nummeret gjengis det andre foredraget i sin originalform.

Hva kan vi slutte av statistikk ?

Erling Sverdrup

Andet foredrag.

Vi omtalte i siste foredrag de store talls lov. Det var en regelmessighet som kom til syne når vi gikk over fra å betrakte enkeltobservasjoner til å betrakte et stort antall observasjoner. Når vi kaster med et pengestykke er det ren gjetning å forsøke å forutsi om resultatet av et enkelt kast vil bli krone eller mynt, men vi har visse holdepunkter hvis vi forsøker å forutsi hvor mange av 100 kast vil vise krone. En ganske fornuftig spådom vil det f.eks være å si at antallet vil ligge mellom 40 og 60.

De store talls lov har funnet anvendelse på mange områder, det mest nærliggende er ~~hasardspill~~ spillebankvirksomhet. En enkeltperson med beskjedne pengemidler bør under vanlige forhold ikke gå ut fra at han kan skaffe seg en sikker inntekt til livets opphold ved hasardspill. Men muligheten for stor gevinst før han har brukt opp sine penger er der naturligvis. For en spillebank med en viss startkapital som kan regne med å spille et stort antall spill om dagen stiller saken seg anderledes. For banken kan hasardspill være en grundsolid forretning med liten sjangse for ruin og også med liten sjangse for uvanlig stor profitt.

La meg nevne (noen) konkrete tall. Anta at banken gir sine kunder anledning til å spille spill med terning. Hvert enkeltspill består i å utføre et kast. Hvis spilleren får 6 øyne vinner han 60 kr. Hvis han ikke får seks øyne vinner han intet. Spilleren må gjøre en innsats før han kaster med terningen. Hvis innsatsen er 10 kr vil, såfremt terningen er rettferdig, den totale innsats ved en rekke spill omtrent være lik de totale utbetalinger. I gjennomsnitt vinnes nemlig hvert sjette spill. Da gevinsten er 60 kr. vil det si at gevinsten pr. spill er 60 dividert med 6, altså 10 kr. Men dette er

jo det samme som innsatsen, som var 10 kr. pr. spill. Naturligvis vil ikke spilleinntektene nøyaktig være lik utgiftene ved de vunne spill. Tilfeldigheter vil virke inn. Vi vil tenke oss at spillebanken sikrer seg mot at tilfeldige uheldige spill skal ruinere dem ved å innkreve en ekstraavgift på 1 kr. pr. spill. Innsatsen blir altså 11 kr. Anta nå at det spilles 10 000 spill om året. Det kan da regnes ut at sjangsen for underskudd et år vil være 4 på en million, d.v.s. at gjennomgående 4 ganger i løpet av en million år vil banken ha underskudd. Banken vil i virkeligheten ~~som regel~~ ha et meget stabilt overskudd. Ved å bruke sannsynlighetsregning kan det regnes ut at overskuddet som regel vil ligge mellom 5 og 13 o/o av spilleinntekten. Det vil gjennomgående bare være ett av 20 år hvor dette ikke vil være tilfelle, ellers vil altså overskuddet ligge mellom 5 og 13 o/o.

Dette illustrerer hvor liten betydning det har for bankens soliditet at det enkelte spill er hasardiøst. Det ble forutsatt at det ble spilt nøyaktig 10 000 spill om året. I virkelighten vil vel usikkerheten med hensyn til kundesøkningen samt anbringelsen av fondsmidlene være de vesentlige usikkerhetsfaktorer, altså samme slags usikkerhetsfaktorer som ved en hvilkensomhelst anden forretning.

Jeg nevnte i foredraget for en uke siden et praktisk statistisk eksempel på de store talls lov. Det gjaldt sjangsen for guttefødsel som har holdt seg forbausende konstant gjennom årene.

I våre dager er vi blitt så blasert når det gjelder tall og statistikk at vi neppe kan mobilisere noen særlig begeistring fordi om vi oppdager såkalte statistiske reglemessigheter. Men for et par hundre år siden har man åpenbart sett anderledes på saken. Noen av de eldste statistiske undersøkelser vi kjenner gjaldt dødeligheten. Det er den engelske astronom Edmund Halley som har æren av å ha stillet opp de første dødelighetstabeller på rasjonelt grunnlag. Halley var forresten en nær venn av Isaac Newton. Det var han som først innså betydningen av Newtons oppdagelser om planetenes bevegelser og overtalte ham til å publisere det. Halleys dødelighetsundersøkelse stammer fra 1693 og ble etterfulgt av en rekke andre. En dødelighetstabell viser dødelighetens variasjon med alderen. Av en slik tabell kan vi f.eks. finne at dødssannsynligheten for en 20-åring er la oss si 7.4 pr. 1000. D.v.s. at det av 10 000 20-åringer vil dø ca. 74 innen de fyller 21 år. Statistikere i det 18. og 19. århundre synes å ha vært slått av forbauselse over den statistiske regelmessighet de fant i dødelighetsmaterialene. Død og ulykker synes i det hele tatt å ha interessert folk sterkt, tildels av religiøse grunner. ~~Død og~~ De enkelte ulykkes fall og ulykker opptrer jo tilsynelatende lite lovmessig, det er noe som rammer når en minst venter det og ble vel regnet som et utslag av Guds uransakelighet. At man oppdaget statistiske lovmessigheter for dødsfall ble derfor av noen tatt som bevis for at det ikke var nødvendig å regne med Guds eksistens når man skulle "forklare" dødsfall, av andre ble det tatt som bevis på en guddommelig ordning. Denne eiendommelige teori om Gud som en ordensmann finner vi hso en tysk prest fra midten av det 18. århundre som arbeidet meget med dødelighetsundersøkelser. Han skriver f.eks.: "Aldri var jeg meg en slik harmoni mellom klosterfolkene i Paris og de brandburgske bønder bevisst. Jeg stusset da jeg oppdaget den og en overordentlig glede over den guddommelige ordning var belønningen for min møye". Videre skriver han: "Slik var livsvarigheten for 3000 år siden, til og med i Orienten, nettopp

den samme som den nå er. Som menneskene i Tyskland blir født, lever og dør, nettopp slik skjer det i Finland, Sverige, England, Holland og Frankrike."

I disse uttalelser er naturligvis dødelighetens og livslengdens uavhengighet av tid, sted og sosial klasse kolossalt overdrevet. Riktignok er det slik at dødeligheten på forskjellige steder i forskjellige sosiale lag og til forskjellige tider viser visse felles trekk. Men det er, som vi er klar over i dag, ~~og som~~ mange var klar over i det 18 århundre, i virkeligheten store avvikelser.

Likevel er dødeligheten så stabil at det er mulig å drive livsforsikringsselskaper på forsvarlig finansiell basis. I livsforsikringsvirksomheten som i spillebankvirksomheten har man trukket de praktiske konsekvenser av loven om de store tall. Det er jo utvilsomt visse likhetspunkter mellom en livsforsikringskontrakt og et hasardspill. Ta en mann som har tegnet en livsforsikring som forplikter selskapet til å betale ham 10 ooo kr. ved hans død. Den forsikrede må betale, la oss si, 250 kr. årlig for denne forsikring. Hvis den forsikrede dør allerede første år, ville den forsikrede altså få utbetalt 10.ooo kr. mens han har innbetalt bare 250 kr. til selskapet, han får altså en pen gevist. Hvis han derimot lever lenge kan summen av premiene med renter og rentesrente overstige forsikringssummen 10.000 kr. og det er selskapet som går av med seieren så og si. Hvis et selskap skulle forsøke å spå om den forsikrede vil leve lenge eller ikke, vil det naturligvis være henvist til ren gjetning. Men et selskap har en stor bestand av forsikrede og kan for denne bestand som helhet danne seg et billede av hvorledes den fremtidige dødelighet vil forløpe og trekke de finansielle konsekvenser av det. Dette er forsikringsmatematikerens aktuarens oppgave. Forsikringsmatematikken bygger på sannsynlighetsregningen, den eri virkeligheten en av de eldste anvendelser av sannsynlighetsregningen.

Er hvis er viktig å være oppmerksom på.
~~Men må naturligvis ikke trekke parallellen mellom hasardspill~~

~~og forsikring for langt~~. At aktuarene anvender sannsynlighetsregning og beregner sjangser for å dø betyr ikke at det er noe spesielt hasardiøst over livsforsikringsvirksomheten. Omtrent all forretningsvirksomhet og anden menneskelig virksomhet, er jo forbunnet med risiko. I livsforsikring har man funnet frem til rasjonelle måter til å anslå denne risiko. Bruken av snnaynlighetsregning øker derfor den finansielle trygghet og er en vesentlig årsak til forsikringsselskapenes finansielle stabilitet.

~~Heller ikke fra den forsikredes synspunkt kan~~ For å unngå misforståelser bør det kanskje pekes på at parallellen med hasardspill ikke bør trekkes for langt. Hasardspilleren søker spenning og avkobling i den finansielle utrygghet. Motivet til å tegne forsikring er jo akkurat det stikk motsatt, å oppnå finansiell trygghet. Man "vinner", om man vil si det slik, når man mest trenger det.

Jeg skal nevne et annet eksempel på statistisk lovmessighet som er litt egenartet. Det gjelder undersøkelse over dødsulykker av hestespark i 10 armekorps av den prøissiske arme i 1875-94. Dødsfall av hestespark var en ytterst sjelden foreteelse, men med så mange hester og soldater som i den prøissiske armé måtte det naturligvis inntreffe. Det er en av de sjeldne ulykker som av og til rent tilfeldig vil inntreffe og som derfor tilsynelatende ikke kan være underkastet noen lovmessighet, det er de såkalte hendelige uhell. Nå kan vi imidlertid gjøre noen rimelige forutsetninger om dødsfallene. Vi kan for det første forutsette at de inntreffer uavhengig av hverandre, det er intet "epidemisk" ved slike ulykker. For det annet kan vi forutsette at det er en forsvinnende liten sjangse forat en på forhånd utpekt soldat skal bli drept av hestespark. Enedelig kan vi forutsette at det er svært mange soldater som er utsatt for å bli sparket ihjel. Disse forutsetninger, man ~~kaller det ofte en~~ sier ofte at de utgjør en "modell", kan oversettes til et stringent matematisk sprog og det er da mulig å utlede en statistisk lov om hvorledes ~~dødsfallene av~~ hvordan det vil bli med hestespark ~~vil inntreffe~~. Den går under navn av Poissons lov etter

den kjente matematiker Poisson fra begynnelsen av det 19. århundre. Ved hjelp av denne lov kan det, når gjennomsnittlig antall dødsfall av hestespark pr. armékorps i et år er kjent, regnes ut hvor mange år armékorpsene vil ha ingen, ett, to, tre o.s.v. dødsfall av hestespark. Disse regninger ble utført og de ble sammenlignet med de faktiske tall i perioden 1875 til 1894. Overensstemmelsen mellom teori og praksis viste seg å være forbløffende god og har åpenbart gjort et sterkt inntrykk på statistikere i slutten av det 19. århundr
De hendelige uhell var misforstått in ~~Hong~~ muskmelk Cor.
Poissons lov finner i dag anvendelse på de forskjelligste områder, i ulykkesforsikring, ved forsikring av sjeldne hendelser, ved bakteriologiske tellinger, ved tellinger av blodlegemer og ved analyse av overbelastning av telefonledninger.

Jeg kunne også tenke meg at loven ville komme til anvendelse ved tolkning av statistiske materialer over flyulykker, skjønt jeg har ikke sett at den virkelig er blitt anvendt. For en tid siden var det aktuelt å sammenligne statistikkene over jetflyulykker i to flyvåpner. Spørsmålet var om tallet på ulykker var påfallene stort i det ene flyvåpen eller om forskjellen var så liten at det var berettighet å snakke om ~~hendelige~~ tilfeldige uhell. Jeg ville gjette på, uten å ha undersøkt saken nærmere, at Poissons lov kunne brukes her.

La meg nevne et ganske eiendommelig eksempel på anvendelsen av de store talls lov som riktignok er uten praktisk betydning. Tenk Dem et alminnelig gammeldags lys-lakkert gulv hvor de enkelte gulvplanker kan sees. Skillet mellom plankene vises altså som parallelle linjer på gulvflaten. Nå ~~kastes det~~ mister man en nål på gulvet slik at det er vilkårlig hvor nålen faller. Sannsynligheten for at nålen skal krysse en av skillelinjene mellom plankene viser seg da å være forbundet med tallet π . π ~~er~~ kan som kjent defineres som forholdet mellom omkretsen og diameteren i en sirkel. Det er forøvrig et fundamental matematisk konstant som forekommer på mange områder i matematikken, og statistikken, uten ~~noen~~ forbindelse med sirkelen. Det viser seg nå

at hvis man dividerer det dobbelte av nålens lengde med π ganger gulvplankenes bredde så får man sjangsen for at nålen skal krysse en skillelinje på gulvet. Anta nå at vi kaster et stort antall ganger med nålen og teller hvor mange ganger nålen krysser en skillelinje. Det er nå et enkelt regnestykke å konstatere at hvis vi multipliserer to ganger nålens lengde med antall kast og dividerer dette med gulvplankenes bredde ganger antall krysninger så må vi omtrent få tallet π, altså 3,14. En astronom i Zürich forteller at han har kastet 5000 ganger og beregnet π til 3,16. I virkeligheten er dette et forbløffende godt resultat. Metoden er altså ikke særlig anbefalelsesverdig når det gjelder å beregne π.

Forutsetningen for den regneregel som jeg har gitt er for det første at gulvet er helt flatt, for det annet at alle gulvplankene er like brede, for det tredje at nålen er kortere enn plankenes bredde og for det fjerde at nålen kastes blindt og vilkårlig d.v.s. alle beliggenheter i forhold til skillelinjene og alle retninger av nålen er like sannsynlige.

Vel, dermed er jeg ferdig med det jeg har å si om de store talls lov, [når det gjelder hasardspill skal vi komme med noen tilleggsmerknader.] Sannsynlighetsregningen er åpenbart et ypperlig redskap til analyse av hasardspill, eller som man ofte pleier å uttrykke det, sannsynlighetsregningen er en realistisk matematisk modell når man skal forklare de fenomener man observerer i hasardspill. Det var den store franske matematiker og fysiker Laplace som førte sannsynlighetsregningen frem til fullkommenhet i så måte, i et verk fra 1812.

Teorien for spill har vært tatt opp til behandling også i de senere år av en av vår tids største matematikere, den ungarsk-amerikanske John von Neumann [medlem av den amerikanske atomenergikommisjon]. I 1928 publiserte han en matematisk teori for selskapsspill, han tenkte på slike spill som bridge, poker, sjakk o.s.v. Men her var det ikke lenger beregningen av sjangser som var det vesentlige. Hensikten var å utforske selve spillestrategien når to eller flere spillere spiller mot hverandre. Det som vakte

von Neumanns interesse var utvilsomt at han her har øynet noe prinsipielt nytt i tankegangen som ikke tidligere hadde opptatt matematikere. Hans teorier kan i virkeligheten anvendes på andre områder enn selskapsspill. Det har f.eks. vært sagt at han i virkeligheten har gitt en matematisk definisjon av militær strategi. Om von Neumanns teorier får noen betydning ved krigføring er det vel ennå for tidlig å si noe om, iallefall for en utenforstående.

La oss vende tilbake til Laplace og hans arbeide fra 1812. Laplace var en fremragende matematiker hvis arbeide har betydd mye for ettertiden. Innledningen til hans store verk om sannsynlighet er generelt filosoferende og kan leses av ikke-matematikere, den er forresten nettopp utgitt i en engelsk oversettelse. Her finner man et typisk uttrykk for opplysningstidens optimisme når det gjelder å forklare fenomener i natur og samfunn ved enkle lover og prinsipper. Laplace hentet naturligvis lovene og prinsippene fra sannsynlighetsregningen. Vi vil vel i dag finne fremstillingen noe kritikkløs, men lesningen av den gir et interessant tidsbillede hvis man er kulturhistorisk interessert.

Laplace grunnla sammen med den store tyske matematikeren Gauss teorien for behandling av målefeil ved fysikalske målinger. Tilsynelatende er dette noe svært teknisk, men i virkeligheten kan det være et godt utgangspunkt for refleksjoner av prinsipiell natur. Vi skal derfor se litt nærmere på denne teori.

Vi vil ta et eksempel fra tegning av karter. Ved den såkalte triangulering legger man inn merker eller varder i terrenget. Hvis man tenker seg rette linjer trukket mellom disse varder, fremkommer trekanter. Landmåleren må blant annet måle vinklene i disse trekanter. La oss nå tenke oss at han har målt de tre vinkler i en trekant, ved fra hver av de tre hjørner å sikte inn retningene til de to andre hjørner. Fra geometrien vet vi at summen av de tre vinkler skal være 180°, d.v.s. det samme som summen av to rette vinkler.

Nå vil naturligvis de vinkler som landmåleren finner ikke ha en sum på 180° på grunn av målefeil som han aldri helt kan eliminere. Det er et statistik problem å kombinere det faktum at summen av de riktige vinkler skal være 180° med de faktiske måleresultater og derved oppnå nøyaktigere verdier av vinklene. I vårt spesielle tilfelle kan vi på denne måten øke nøyaktighetsgraden med 18 o/o såfremt det bare er en måling av hver av de tre vinkler.

Her kan det være på sin plass å minne om det som jeg sa i mitt første foredrag, nemlig at statistikerens oppgave er å kombinere det som på forhånd er kjent om det som måles eller telles med de opplysninger som det statistiske materiel gir. I tilfellet med triangu-leringen var det som på forhånd var kjent at summen av vinklene skulle være 180°. Det statistiske materiale var målingene av de tre vinkler.

På de forskjelligste områder hvor statistisk teori brukes finner man samme situasjon, det gjelder landbruk, medisin, biologi, meteorologi o.s.v. Det er bare den forskjell at det i høyere grad kan være en vurderingssak å avgjøre hva man kan gå ut fra som på forhånd kjent. Legen, gårdbrukeren, bedriftslederen o.s.v. vil til daglig i sitt yrke høste en masse erfaring som det er umulig for ham å nedtegne med vitenskapelig nøyaktighet. På grunnlag av disse erfaringer vil han kunne komme med almengyldige påstander. Det kan riktignok ofte være farlig å stole på slike påstander basert på ikke systematisk nedtegnede inntrykk. Det er i virkeligheten utrolig hvor lett man kan bli bedradd. Men enhver som har arbeidet med statistiske undersøkelser vil vite at det kan føre til rene meningsløsheter å neglisjere alle erfaringer som ikke er basert på omhyggelige nedtegnelser. Det blir derfor et viktig vurderings-spørsmål å avgjøre i hvilken grad man skal ta hensyn til slike erfaringer. Når man har vurdert hva man skal gå ut fra som på for-hånd kjent må denne forhåndsviden omsettes til det sprog som sann-

synlighetsteoretikeren bruker. Når det er gjort kan man begynne å slutte noe av det statistiske materiale. Da er det et matematisk problem å finne frem til den beste måte å skaffe tilveie et statistisk materiale og tolke det.

Vi skal i neste foredrag nærmere omtale vurderingen av forutsetningene ved en statistisk undersøkelse og dessuten se litt på matematikkerens arbeidsmåte når det gjelder å komme frem til den beste statistiske metode.

OLAV REIERSØL

Da Olav Reiersøl fylte 90 år sommeren 1998, fikk professor Yngve Willassen ved Økonomisk institutt anledning til å intervjue ham i hans hjem på Nøtterøy. Intervjuet ble trykket i tidsskriftet Econometric Theory i år. Selv om det legges mest vekt på Reiersøls forskning på området økonometri i dette intervjuet, vil det ganske sikkert ha interesse også for rene statistikere.

Siden det er 25 år siden Reiersøl ble pensjonert fra sin daværende stilling som professor i statistikk ved Matematisk institutt, Universitetet i Oslo, er det sikkert mange lesere av TG som hverken kjenner ham personlig eller kjenner hans vitenskapelige produksjon. Uansett om man har kjent Reiersøl eller ikke, får man nå en glimrende anledning til enten å fornye bekjentskapet eller å bli kjent med personen og forskeren Olav Reiersøl.

Econometric Theory, **16**, 2000, 113–125. Printed in the United States of America.

THE ET INTERVIEW: PROFESSOR OLAV REIERSØL

Interviewed by Yngve Willassen
University of Oslo



Olav Reiersøl

Professor Olav Reiersøl was born on 28 June 1908 in the countryside in Norway. He enrolled at Oslo University in 1928, where he studied chemistry, mathematics, physics, and botany. Choosing mathematics as his main subject, he received the cand. real. degree in 1935. As a graduate in mathematics and mathematical statistics he was employed as an assistant by Professor Ragnar Frisch.

During the Second World War Norway was occupied by the Germans in April 1940. Later on, in November 1943, the university was closed down. In this

© 2000 Cambridge University Press 0266-4666/00 $9.50

action the Germans arrested many university teachers and all the students they got hold of. Reiersøl, who had been active in the resistance, then fled to Sweden (in December 1943) and stayed there until June 1945. In the years 1946–1949 he spent most of the time abroad, mainly in England and the United States.

From 1949 he has been permanently attached to Oslo University. He was appointed docent in mathematics in 1950 and in 1958 professor of mathematics. In 1961 he was appointed professor of mathematical statistics, a position in which he served until his retirement in 1975. With this background in econometrics, mathematics, and mathematical statistics it is quite natural that Reiersøl in his scientific writings has touched many different areas. His works include the following: identifiability of structural models in econometrics and factor analysis, general estimation by means of instrumental variables, design of experiments, logit analysis, multiple comparisons, computational methods for matrices and stochastic processes, statistical problems in genetics, and the history of probability. Reiersøl has made several substantial contributions to econometrics and statistics. His works on identifiability and instrumental variables are standard references both in econometrics and statistics. The work on genetic algebras has been reprinted in a collection of benchmarks in genetics [i]. So undoubtedly, Reiersøl has made several fine contributions to mathematical statistics. Professor Reiersøl, Ragnar Frisch, Trygve Haavelmo, and Herman Wold made up the strong Scandinavian group whose contributions definitely helped lay the statistical foundation of modern econometrics. Therefore, I think it is Reiersøl's works on econometrics that are most interesting to the econometric society. Thus in the interview that follows I will concentrate on that part of his scientific work.

I met Reiersøl in autumn 1974. I remember clearly my first glimpse of him at an underground tram station: an unexpectedly stout, ruddy, and shy figure, almost fisherman-like, with a rucksack and an old woollen hat. I had then heard quite a lot about him during my stay in Uppsala and Gøteborg with Herman Wold. Wold always spoke of him with approval and respect. The stay with Wold had aroused my interests in soft modeling so I already knew his papers on identifiability quite well. Knowing, therefore, that Reiersøl was a prominent figure on latent models I quickly got in touch with him. Since then he has been extremely generous in the amount of time he has spent on answering questions, reading, and commenting on my papers. Therefore, it is a pleasure for me to conduct the present interview.

In order to get started, can you tell us about your educational background?

I attended the local primary school, where my father was the teacher. This was a small school with only two classes, so the pupils belonging to the same group had quite different ages. I studied books on algebra, geometry which I found at home. Another interest was in botany and the identification of flowers. I read with great interest an old Danish book on astronomy. From this and other books on astronomy I got an idea of the seemingly strange movements of the planets. On dark

nights I could observe how the constellations changed positions depending on the rotation of the earth and the different seasons.

I was interested in science. In those days the education in science was quite comprehensive, and we had to study several topics. I studied botany, chemistry, mathematics, and physics at the University of Oslo. Then I chose mathematics as my main subject and decided, if possible, to specialize in probability theory and mathematical statistics. My knowledge of these subjects was at that time quite limited. I had read the greater part of E. Czuber's *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung* [iii]. I had also tried to read some of the articles appearing in *Biometrika*, but I quickly realized that I didn't have the prerequisite knowledge to understand them fully. When I was a student, there was no chair in mathematical statistics at the University of Oslo. Alf Guldberg, a professor of mathematics, gave lectures in mathematical statistics for students in actuarial science and lectures in the calculus of probability for students in economics. When I got in touch with Professor Guldberg he recommended R. von Mises's book *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik* [x]. I also read Steffensen's book *Matematisk iagttagelseslære* [xii]. I finished with the cand. real. dissertation *Undersøkelser over momenter, halvinvarianter og deres frembringende funksjoner* (A study on moments, semi-invariants, and their generating functions) in 1935.

I guess Frisch was quite active at that time. Can you tell us about your contact with Frisch and his institute?

When I started my studies in mathematical statistics, I didn't know that Frisch was working in this field. But in the course of my studies I became familiar with his "Correlation and Scatter" [v]. In this paper Frisch makes extensive use of matrix theory, and I think this is one of the earliest applications of matrices in statistics. However, I hadn't learned anything about matrices in my studies of mathematics, so I had to read up this subject during my summer holidays in 1935.

Frisch and Wederwang were running an Economic Institute which was not part of the university but was financed by the Rockefeller Foundation and Norsk Varekrigs-forsikrings Fond. Early in 1936 I was employed by Frisch as his assistant. In his work on developing new statistical methods, Frisch worked partly with real and partly with artificial data. In order to help him in carrying out the numerous numerical calculations he had a group of young assistants. When I joined him, cand. oecon. Trygve Haavelmo was the supervisor of this group. Later on, Haavelmo left the institute, and I took over this position.

In January 1938 I was appointed lecturer in theoretical statistics for students in economics. This was a newly created lectureship at the university. From 1938 until 1945 I was the only teacher in statistics for students in economics. Occasionally, Frisch himself gave lectures on special topics. He lectured on time series analysis and later on regression theory and confluence analysis.

Frisch's book on confluence analysis [vi] appears in one way or another to be the origin of most of your works in econometrics. In modern parlance this book deals with latent variable models. Can you tell us why you think these models are important to study in econometrics?

Well, in econometrics as in any science for that matter, we wish to deduce stable relations. In order to be specific I think it is best to consider a simple example. Let $X_1, \ldots, X_n$ denote the observable variables while $Y_1, \ldots, Y_n$ denote the corresponding latent variables. Then consider the standard latent variable model given by the structural equation

$$Y_1 = \alpha_{10} + \alpha_{12} Y_2 + \ldots + \alpha_{1n} Y_n + U \tag{1}$$

and the measurement equations

$$X_i = Y_i + \varepsilon_i \tag{2}$$

where U is a random disturbance and ε_i are measurement errors.

If the Y's and the ε's are independent and both are multivariate normal, then the X's are also normal. Hence, if we take the regression of X_1 on $X_2, \ldots, X_n$, the regression equation

$$X_1^P = \beta_{10} + \beta_{12} X_2 + \cdots + \beta_{1n} X_n \tag{3}$$

gives the expectation of X_1 for given values of $X_2, \ldots, X_n$ and can be used to predict X_1. Similarly we can regress Y_1 on $Y_2, \ldots, Y_n$ to obtain the regression equation

$$Y_1^P = \alpha_{10} + \alpha_{12} Y_2 + \cdots + \alpha_{1n} Y_n. \tag{4}$$

However, we can question the use of (4) since this is a regression equation in unobservable variables.

The point is that we may be interested in (4) because it is more than a regression equation; it can be a structural equation. By a structural equation I mean an equation which, compared to a regression equation, has a greater stability. Here I don't think of the stability of the estimators in repeated samples but of the stability with respect to changes in the model. The use of equation (3) for making predictions presupposes that the joint distribution of the X's doesn't change. However, equation (4) can still be valid although the distribution of observable variables changes. We could say that equation (1) or (4) has a higher degree of autonomy.

In addition we can also be interested in equation (4) although we only consider it to be a regression equation. Even in this case it will be more stable than equation (3) since it is independent of the distribution of the measurement errors $\varepsilon_1, \ldots, \varepsilon_n$, while the regression equation (3) will depend on the variances of $\varepsilon_1, \ldots, \varepsilon_n$. I assume of course, that these errors are independent.

Again it is interesting to illustrate by an example. Moran [xi] considers a model with two X's and two Y's. In Moran's study Y_1 is the total annual stream flow of

the river Kiewa, which passes the town Kiewa (in Victoria, Australia). The term Y_2 is the total annual precipitation in the catchment area. The variable X_2 is the annual precipitation measured at Bright, a town just outside the Kiewa catchment. If, instead of X_2, we had used a variable X_2^* based on the precipitation from different places within the Kiewa catchment, then we might obtain a smaller value of ε_2 and, therefore, another value of β_{12}.

An interesting result in Frisch's confluence analysis is that although the vector of structural parameters is not identifiable, one can under certain conditions deduce bounds on this vector by ordinary least square regressions. For the bivariate case the theorem is stated in Frisch [vi, p. 60], although the original result appears to go back to Gini [iv]. Nowadays, econometricians often refer to Kalman [vii] and Klepper and Leamer [viii] for the extension of this result to the multivariable case. Can you tell us about the history of this result?

Well, Gini wrote in Italian so perhaps Frisch didn't know about Gini's result. The first extension of this result to the multivariate case is given by Koopmans [ix, p. 101]. That the adjoint of the covariance matrix was positive or could be made positive by proper sign changes of the variables turned out to be an important condition for the validity of this result. In a way Koopmans proved the necessity part of this condition, and then I showed that it also was sufficient. The complete formulation and proof of this result are given in theorem 14 in my *Econometrica* article from 1941 [2].

So Koopmans is an important contributor in this connection. When did you first meet Koopmans?

Koopmans came to Oslo in autumn 1935 to study Frisch's confluence analysis. On the request of Frisch, Koopmans then gave a series of lectures on R.A. Fisher's theory of estimation and Neyman–Pearson's theory of testing hypotheses. I attended these lectures. Later on I studied the original works by Fisher, Neyman, and Pearson. That was a great experience, indeed.

You worked on your doctoral thesis during the Second World War when Norway was occupied by the Germans. I guess it was difficult to do scientific work under such circumstances. Can you tell us about your life and work in that period?

Yes, that was difficult times, indeed. In August 1941 the Germans started to confiscate people's radios. By that time we already had some illegal newspapers. But after this confiscation of radios several new illegal newspapers appeared. One of these was *Avantgarden*, to which I was connected. Together with a friend I was listening to news from London and Moscow (my friend was quite fluent in both Polish and Russian), which was afterwards published in this newspaper. I also did some work with distributing *Avantgarden*. In October 1943 two of my

illegal contacts were arrested, and I had to go into hiding with one of my cousins living outside Oslo. In November that year the Germans went into the university, arrested students and teachers, and closed the university. After this event I decided to flee to Sweden. Equipped with a false identity card and fake papers I traveled by train and a lorry to Trysil (a place not far from the Swedish border). From Trysil I went on skis to Sweden.

Having arrived in Sweden, I went to Stockholm and was lucky enough to get a room at Cramér's institute for actuarial mathematics and mathematical statistics. I had almost completed my doctoral thesis, *Confluence Analysis by Means of Instrumental Variables*, when I left Oslo in December 1943. The manuscript was typed at an office run by the Students' Association. When I had to flee to Sweden the typewriting was only half finished, so I managed to bring with me only the first part of my dissertation. Later on, someone working at the Economic Institute delivered the typed second half of this manuscript to the Norwegian editor of the journal *Skandinavisk Aktuarietidsskrift*. He in turn sent it to professor C.O. Segerdahl, the Swedish editor of this journal. In this way I managed to get the complete dissertation to Sweden, and in May 1945 I was awarded the doctorate from Stockholm University.

Staying with Cramér in Stockholm I also remember that I attended his lectures on measure and integration theory. Cramér was then writing his book *Mathematical Methods of Statistics* [ii]. When I returned to Norway that year I had a copy of the first part of this book. The book was published in Uppsala in the autumn of that year.

Your doctoral thesis was an extension of Frisch's confluence analysis?

Yes, there I introduced the idea of an instrumental variable to the confluence analysis. Herman Wold was the "first opponent." I remember that there was a summary of the thesis in a Stockholm newspaper saying that it dealt with breeding of pigs. You see, on the last two pages I illustrated some of my methods by applying them on some Danish pig production data.

So the newspapers then were as reliable as they are today?

I think you can say that.

The method of instrumental variables has a place of prominence in modern econometrics and statistics. You are considered to be its originator. Can you tell us how this idea developed?

Well, a random variable is said to be an instrumental variable relative to a stochastic equation if it is uncorrelated with the disturbance of the equation and uncorrelated with possible errors of observation of the variables occurring in the equation, while it is correlated with some of or all the true values of the variables in the equation. This idea emerged during my work on deriving consistent estimators of the slope parameters in some models in the confluence analysis. There

I saw quickly that with the usual assumptions on the random errors, one could deduce consistent estimators by using lagged values of the appropriate observable variables. The idea is used in my *Econometrica* article from 1941 [2]. Originally I tried several different terms for this variable, but Frisch was not satisfied with any of my proposals. He thought that this was a good idea and should have a name that matched its importance. One day he told me enthusiastically that he had found the proper name and this variable should be termed *Instrumental variable*! I myself used this term for the first time in my doctoral thesis in 1945.

Can you recall using the idea of instrumental variables in other applications?

I think I have touched upon it several times, e.g., in [2] and [5]. In [5] I showed that the well-known Frisch–Waugh theorem easily generalizes to the instrumental variable case.

When you returned from Sweden you resumed your work as a lecturer in statistics for students in economics. But already the next year you went abroad again. Can you tell us about your travels abroad?

Yes, in 1946 I received a three months' scholarship which I spent in Cambridge, England. There I got a small office at the Department of Genetics, of which R.A. Fisher was then the director. Most of the time I spent studying some of Fisher's works, among them his papers in genetics. In the office in which I had my desk, Fisher also kept a huge pile of reprints. In this collection I happened to see some articles on genetic algebras by Etherington. These papers were interesting, but it wasn't until 1958 that I myself started to study seriously these algebras.

After Cambridge I got a scholarship from the Rockefeller Foundation which made it possible for me to go to the United States. So, the autumn of 1946 I spent at Columbia University. There I followed lectures on design of experiments given by A. Wald and lectures on multivariate analysis by T.W. Anderson. The following spring I spent at the University of North Carolina attending lectures on probability theory by H. Robbins and on multivariate analysis by the Chinese P.L. Hsu. This summer (1947) I stayed in Chicago. Here I followed lectures on identification and estimation of systems of stochastic difference equations, and there I also participated in the weekly Cowles Commission Discussion Meetings. I left Chicago for Berkeley in January 1948. In Berkeley I participated in the weekly seminars but didn't follow any lectures. For the academic year 1948–1949 I was visiting assistant professor at Purdue University after being recommended by Neyman for this position. There I gave an introductory course in mathematical statistics and a course in difference calculus. I also gave two seminars, one on a book by Wilks and one on Thurstone's factor analysis.

In 1949 I went back to Norway and did not revisit the States until 1964. In the academic year 1964–1965 I was visiting professor at Cornell University.

During your first stay in the United States you wrote three influential papers on the identifiability of structural models [7–9]. Can you tell us how these papers came into being?

The first one I wrote together with Koopmans. I already knew Koopmans from his stay in Oslo, and collaborating with him on identification problems was quite natural. Koopmans had worked on the identifiability of econometric models in the Cowles Commission while I had studied the identifiability of factor analysis models. Anyway the paper was published in *Annals of Mathematical Statistics.*

Tell us about your interests in factor analysis.

Although I am no psychologist I have some interest in psychology. When I read Thurstone's books [xiii, xiv] on factor analysis, I realized, immediately, the similarities between factor analysis models and models in Frisch's confluence analysis. There is also a formal resemblance between the reduced form of an econometric model and the structural form of a factor analysis model. Although, it should be kept in mind that the exogenous variables are observable in the econometric model while the factors, as well as the number of factors, are unobservable in the factor model. However it was clear that the methods deduced by the Cowles Commission to study identification of the econometric model should have a certain analogy to methods appropriate for the factor analysis. On one occasion Koopmans asked me to study the identifiability of the parameters in Thurstone's multiple factor analysis from the point of view of this analogy.

I met Thurstone in Chicago (1947) and later at a meeting in Oslo. Thurstone had Swedish parents, and Thurstone is an anglicization of Torsten, which is a very common Scandinavian first name.

Now you have told us about two of four works from this period. However, the paper "Identifiability of a Linear Relation between Variables Which Are Subject to Error" [9] is possibly the one which is the best known. Can you tell us about its emergence?

In those days a lot of work was going on in deriving consistent estimators of the regression parameters in latent variable models. Since identifiability is a necessary condition for estimability it was quite natural to study the identifiability of these parameters. Using the general definition of identifiability I approached this problem through the application of characteristic functions. The theory of characteristic functions I knew quite well from my education in mathematics and from the books by von Mises [x] and Cramér [ii].

Finally, I know that you have a lifelong interest in Esperanto. Can you tell us about it?

As a young boy I happened to find a book on Esperanto at home. Nobody there seemed to know anything about Esperanto, but this little book aroused my interest in this language. When I later on went to Oslo to attend schools, I quickly

became a regular user of the public library. There I found textbooks and other written material on Esperanto, but still I didn't know anyone speaking this language. Nor did I know there was an international Esperanto Association which had a local society in Oslo.

However, in autumn 1932 a German lecturer in Esperanto came to Oslo to give an Esperanto course. I attended her lectures, and afterwards I managed to read Esperanto quite easily with a dictionary at hand. The Oslo Esperanto society regularly arranged meetings/lectures which offered good opportunities for me to practice the language. The Second World War restrained the activities of the Esperanto movement in Norway, but in Sweden the war had the opposite effect. When I arrived in 1944 the Esperanto movement there was very active. Some refugees from Central Europe staying in Sweden spoke Esperanto more easily than Swedish. I learned then that Esperanto was a good second language for people having very different native languages.

When, after the war, I went to England and later on to the States, I always took the opportunity to visit the local Esperanto societies. In 1950 I participated for the first time in "Universala Kongreso," which is held every year. That year the meeting took place in Paris. For me it was a great experience to witness that thousands of delegates coming from almost every country in the world were able to speak to each other and to understand lectures given in one language without any need for translations. In 1952 the Universala Kongreso was arranged in Oslo. I participated in the "Loka Kongresa Komitato" organizing the congress. The yearly "Universalaj Kongresoj" is also a good occasion to meet and to discuss professional questions with colleagues.

I was one of the founders of the international Esperanto association for mathematicians, "Internacia Asocio de Esperantistaj Matematikistoj." In 1976 I gave lectures on probability theory in Esperanto at a seminar in Antwerp, "Someraj Universitataj Kursoj." I have corresponded with statisticians and mathematicians from all over the world, in particular on the compilation of statistical and mathematical terminology in Esperanto. But I have also been interested in Esperanto from a linguistic point of view. I have worked on the use of the verbal forms in Esperanto and with criteria for language assessment of interlinguistics. Together with Professor Richard E. Wood I have published, in English, an article on the use of suffixes in English and Esperanto.

In 1951 I published an article, "Transiro per diferencialaj ekvacioj de probablodenso al karakteriza funkcio kaj inverse," in the Portuguese journal *Portugaliae Mathematica* [11]. After the publication I sent a reprint of this paper to a colleague in the States. He responded with kind words and remarked that he didn't know that it was so easy to read Portuguese!

REFERENCES

i. Ballonoff, P. (ed) (1974) *Genetics and Social Structure: Mathematical Structuralism in Population Genetics and Social Theory.* Benchmark Papers in Genetics, vol. 1. Stroudsburg: Dowden, Hutchinson, and Ross. (Reiersøl paper is reproduced on pp. 134–153.)

ii. Cramér, H. (1945) *Mathematical Methods of Statistics.* Uppsala, Sweden: Almquist and Wicksell.
iii. Czubers, E. (1903) *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung.* Leipzig: Druck und Verlag von B.G. Teubner.
iv. Gini, C. (1921) Sull' interpolazione di una retta quando i valori della variabile indipendente sono affetti da errori accidentali. *Metron* 1, 62–82.
v. Frisch, R. (1929) Correlation and scatter in statistical variables. *Nordisk Statistisk Tidsskrift* 8, 36–102.
vi. Frisch, R. (1934) Statistical Confluence Analysis by Means of Complete Regression Systems. Publ. 5, Universitetets Økonomiske Institutt, Oslo.
vii. Kalman, R. (1982) System identification from noisy data. In A. Bednarek & L. Cesari (eds.), *Dynamical Systems II.* New York: Academic Press.
viii. Klepper, S. & E. Leamer (1984) Consistent sets of estimates for regression with errors in all variables. *Econometrica* 52, 163–183.
ix. Koopmans, T. (1937) *Linear Regression Analysis of Economic Time Series.* Haarlem-De Erven F. Bohn N.V.
x. v. Mises, R. (1931) *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik.* Leipzig and Wien: Franz Deuticke.
xi. Moran, P.A.P. (1956) A test of significance for an unidentifiable relation. *Journal of the Royal Statistical Society, Series B* 18, 61–64.
xii. Steffensen, J. (1923) *Matematisk iagttagelseslære.* Gad: København.
xiii. Thurstone, L.L. (1935) *The Vectors of Mind.* Chicago: University of Chicago Press.
xiv. Thurstone, L.L. (1947) *Multiple-Factor Analysis.* Chicago: University of Chicago Press.

REIERSØL'S PUBLICATIONS IN ENGLISH AND ESPERANTO

1940

1. A method for recurrent computation of all the principal minors of a determinant and its application in confluence analysis. *Annals of Mathematical Statistics* 11, 193–198.

1941

2. Confluence analysis by means of lag moments and other methods of confluence analysis. *Econometrica* 9, 1–24.

1944

3. Measures of departure from symmetry. *Skandinavisk Aktuarietidsskrift* 27, 229–234.

1945

4. Confluence analysis by means of instrumental sets of variables. *Arkiv för Matematik, Astronomi och Fysik* 32A (4), 1–119.
5. Residual variables in regression and confluence analysis. *Skandinavisk Aktuarietidsskrift* 28, 201–217.

1949

6. With R.C. Marble. A comparison between word formation in Esperanto and English. *Esperantologio.* 1 (1), 53–58.

1950

7. With T.C. Koopmans. The identification of structural characteristics. *Annals of Mathematical Statistics* 21, 165–181.

8. On the identifiability of parameters in Thurstone's multiple factor analysis. *Psychometrika* 15, 121–149.
9. Identifiability of a linear relation between variables which are subject to error. *Econometrica* 18, 375–389.
10. Diferencialaj ekvacioj de specimenaraj distribuoj. Sosialøkonomisk Institutt, University of Oslo.

1951

11. Transiro per diferencialaj ekvacioj de probablodenso al karakteriza funkcio kaj inverse. *Portugaliae Mathematica* 10 (2), 71–80.

1954

12. Tests of linear hypotheses concerning binomial experiments. *Skandinavisk Aktuarietidsskrift* 37, 38–59.
13. Identification of Systems of Linear Stochastic Difference Equations. Memorandum fra Sosialøkonomisk Institutt, University of Oslo.

1955

14. A Reconsideration of Confluence Analysis. Memorandum fra Sosialøkonomisk Institutt, University of Oslo.
15. Derivation of Differential and Difference Equations for the Calculation of Integrals. Institute of Mathematics, University of Oslo.
16. Pri statistikaj miskonkludoj. *Scienca Revuo* 7, 10–16, and pp. 51–57 in *Elektitaj Prelegoj de ISU.* Muusses, Purmerend, Netherlands.

1956

17. A note on the signs of gross correlation coefficients and partial correlation coefficients. *Biometrika* 43, 480–481.

1958

18. Methods of Deriving Recurrence Formulas. Institute of Mathematics, University of Oslo.

1961

19. Linear and non-linear multiple comparisons in logit analysis. *Biometrika* 48, 359–365.

1962

20. Genetic algebras studied recursively and by means of differential operators. *Mathematica Scandinavica* 10, 25–44.
21. Path Coefficients, Moments of Moments and Expectations of Expectations in Population Genetics. Publications from Det Norske Videnskaps-Akademi in Oslo, Matematisk-Naturvitenskapelig Klasse, Ny Serie, no. 6, 23 s.
22. Applications of Markov Chains to Population Genetics. Institute of Mathematics, University of Oslo.

1963

23. Rational orthogonal matrices. *American Mathematical Monthly* 70, 63–65.
24. Identifiability, estimability, phenorestricting specifications and zero Lagrange multipliers in the analysis of variance. *Skandinavisk Aktuarietidsskrift* 46, 131–142.

25. Nonsensical Maximum Likelihood Estimators. Institute of Mathematics, University of Oslo.
26. Kelkaj konsideroj pri matematikaj terminoj en Esperanto. Institute of Mathematics, University of Oslo.

1965

27. Statistical Inference for Multinomial Distributions. Cornell University, New York.
28. Notes to Wolfowitz: Coding Theorems of Information Theory. Institute of Mathematics, University of Oslo.

1966

29. Notes to Kai Lai Chung: Markov Chains with Stationary Transition Probabilities. Institute of Mathematics, University of Oslo.
30. Notes on Mappings, Systems of Sets and Systems of Functions. Institute of Mathematics, University of Oslo.

1968

31. Note on some propositions of Huygens in the calculus of probability. *Nordisk Matematisk Tidsskrift* 16, 87–91.

1973

32. Konsideroj pri matematikaj kaj stokastikaj terminoj en Esperanto. Institute of Mathematics, University of Oslo. Republished with additions in H. Bergmann and S. Stimec (eds.), *Klerigo kaj prikalkulado*. Bamberg: Verlag Difo-Druck, 1976.

1976

33. Probablokalkulo. Elementa enkonduko. SUK-kursoteksto-KT 21, Liège.

1980

34. Participoj kaj konjugaciaj formoj en Esperanto. Institute of Mathematics, University of Oslo.
35. Kriterioj por lingvotaksado en interlingvistiko. In Istvan Szerdahelyi (ed.), *Miscellanea Interlinguistica*, pp. 216–227. Budapest.

1983

36. Matematikaj kaj statistikaj terminoj en Esperanto. Suplemento al Plena Ilustrita Vortaro. Statistical memoirs 1, Institute of Mathematics, University of Oslo. Republished in R. Eichholz (ed.), *Akademiaj Studoj*. Ontario, Canada: Bailieboro.

1984

37. Diferencialaj ekvacioj de samplaraj distribuoj. Statistical research report 6, Institute of Mathematics, University of Oslo.

1985

38. Principoj por elekto de matematikaj kaj stokastikaj terminoj en Esperanto. Statistical memoirs 2, Institute of Mathematics, University of Oslo.
39. Kritika analizo de la Akademiaj proponoj kaj decidoj pri la uzo de verbaj formoj. In. R. Eichholz (ed.), *Akademiaj Studoj*, pp. 195–204. Ontario, Canada: Bailieboro.

1986

40. Respondo al la artikolo de Lapenna: "Pri la pasivaj participoj -ata kaj -ita." In R. Eichholz (ed.), *Akademiaj Studoj*, pp. 158–159. Ontario, Canada: Bailieboro.

1987

41. Matematika kaj stokastika terminaro Esperanto. Statistical memoirs 1, Institute of Mathematics, University of Oslo.
42. Kritiko de la recenzo de Michel Duc Goninaz. In R. Eichholz (ed.), *Akademiaj Studoj*, pp. 15–18. Ontario, Canada: Bailieboro.
43. With Richard E. Wood. A comparison of the use of some suffixes in Esperanto and English. *Serta Gratulatoria in honorem Juan Regulo*, vol. 2. pp. 583–599.

1988

44. Komentario al kelkaj vortoj en Plena Ilustrita Vortaro de Esperanto. In R. Eichholz (ed.), *Akademiaj Studoj*. Ontario, Canada: Bailieboro.

Ekstraordinært hovedmøte i Norsk statistisk forening

Mandag 5 juni 2000, kl 18.15, Sørlandet kurs- og konferansesenter, Grimstad,
i tilknytning til 18th Nordic Conference on Mathematical Statistics.

Referat

Tilstede: 33 medlemmer.
Som ordstyrer og referent ble valgt Geir Egil Eide og Stian Lydersen.

1. Godkjenning av innkalling og saksliste

Innkallingen stod i Tilfeldig gang, mai 2000. Innkalling og saksliste ble godkjent.

2. Orienteringssaker

Formannen gav en kort orientering om følgende saker, se også TG mai 2000:

- Årsberetning og regnskap for siste år kommer i neste TG.
- Relevant reklame pr e-post til medlemmene: Etter en rådgivende epost spørreundersøkelse (se TG mai/2000) vil foreningen tillate dette. Det enkelte medlem kan reservere seg mot dette.
- Statistikk i den videregående skole: Ørnulf Borgan er medlem i en referansegruppe oppnevnt av Departementet. Det kommer informasjon om dette i neste TG.
- Diskusjonsgruppen NSF-forum på e-post er opprettet, men det er lite aktivitet der.

3. Endringer i foreningens vedtekter

Endringer i NSFs vedtekter med *kursiv*:

Endring nr 1:
§ 3 LOKALE FORENINGER *OG FAGSEKSJONER*

Medlemmene innen et geografisk område kan danne en lokalforening. Lokalforeningene har egne vedtekter og egen økonomi. Vedtektene *skal* 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.*

Endring nr 2:
Formannen endrer tittel til "leder". Alle henvisninger til "formann" i vedtektene endres tilsvarende.

Endring i §1 i vedtektene for foreningens fond:
Følgende punkt tilføyes:

- *stipend for hovedfagsstudenter i statistikk for deltakelse på faglige konferanser i Norden*
- *til andre formål som styret måtte beslutte å støtte med inntil kr 30000 pr år.*

Ovenstående endringer ble enstemmig vedtatt. De er identiske med forslagene i møteinnkallingen, bortsett fra et par språklige forbedringer.

5. Eventuelt

Ingen saker.

Stian Lydersen/møtereferent

Kreft er ikke fett

Hvert år får 20 000 nordmenn diagnosen kreft. Men hva skyldes det? Marit Bragelien Veierød har tatt doktogradenpå sammenhengen mellom enkelte krefttyper og fettinntaket i maten.

Av Karsten Eig og Khai Nguyen (foto)

– Har du funnet svaret på hvordan vi skal holde oss raske og friske og unngå kreft?

– Nei, så enkelt er det nok ikke. Kreft er et samlebegrep for en rekke sykdommer. Mat kan inneholde komponenter som fremmer kreftutviklingen og komponenter som beskytter mot kreft. Min avhandling er en liten brikke i et stort puslespill, sier Veierød.

– Jeg har forsket på sammehengen mellom fettintaket og forekomsten av tre forskjellige krefttyper: lungekreft, føflekkreft og prostatakreft.

Joda, røyking er farlig

Veierød har sammenlignet kostholdsdata fra 50 000 nordmenn, samlet inn i 1977-1983, og koblet dette med data fra Kreftregisteret frem til 1992.

– Vet vi ikke allerede at lungekreft skyldes røyking og at hudkreft skyldes for lite solkrem om sommeren?

– Joda, vi regner med at 80 prosent av alle lungekrefttilfeller skyldes røyking. Men ikke alle som røyker får lungekreft, og vi tror at kostholdet har en innvirkning her.

Veierød fant at de som tok tran hadde lavere risiko for lungekreft enn de som ikke tok tran. Det er første gang en slik sammenheng er undersøkt i, så hun understreker at det må gjøres flere studier om dette.

Farlig sommer

– Føflekkreft, derimot, ser ut til å være lite påvirket av kostholdet. Her er det overdreven soling som er den store stygge ulven.

Norge ligger dessverre i verdenstoppen når det gjelder hudkreft, noe som nok har sammenheng med at vi ofte utsetter vår lyse hud for brå soling.

– Og, eh, der nede...?

– Prostatakreft er en kreftform som har økt enormt, men vi vet lite om hva som forårsaker den, bortsett fra at det er store geografiske variasjoner i forekomsten. Jeg fant ikke noen sammenheng med fettinntaket i dataene mine.

Veierød sier at de viktigste rådene for å forebygge kreft er god, gammel visdom som vi alle kan:

– Spis nok frukt og grønnsaker, unngå overvekt, hold deg fysisk aktiv og stumprøyken!

Marit Bragelien Veierød har tatt Dr. Philosgraden på avhandlingen Malignant melanoma, lung cancer and prostate cancer in relation to diet, *der hun har studert sammenhengen mellom fettinntaket i kosten og forekomst av tre forskjellige kreftformer, og metodologiske problemer knyttet til slike studier. Arbeidet er finansiert av Kreftforeningen og Lægeforeningen.*

Marit Bragelien Veierød, cand. scient. med statistikk hovedfag fra Universitetet i Oslo, disputerte for Dr. Philos.-graden 1. september i år.
Dette intervjuet med henne er sakset fra Universitas, 13. september 2000.

Nytt fra Universitetet i Oslo

Nye cand. scient.-er

Jan Fredrik Nordby: "Simulering av totalskade-fordelinger."
Veileder: Arne Bang Huseby

Trond Kristian Pedersen: "Asymptotiske egenskaper til REML-estimatorer i blandede lineære modeller."
Veileder: Inge S. Helland

Thorleif Maudal: "Prisberegning med produktmotor i skadeforsikring."
Veileder: Morten Byholt

Henning Myrheim: "Beslutningsstøtte fra analyser av pålitelighetsdata med alternative probabilistiske basiser."
Veileder: Terje Aven

Mats Sollie: "Forløpsanalyse av kundeforhold i et skadeforsikringsselskap."
Veileder: Nils Lid Hjort

Trond Anders Moger: "Nøstede kasus-kontroll studier – hvor god er asymptotisk teori?"
Veileder: Ørnulf Borgan

Ingunn Fride Tvete: "Bayesiansk tidsrekkeanalyse anvendt på næringslivs-forsikringen."
Veileder: Bent Natvig

Pia Gustavsson: "Ruinsannsynligheter i en endelig og uendelig tid ved numerisk integrasjon"
Veileder: Erik Bølviken

Tor Cato Gjerdingen: "Dynamiske endringer i parametrene i skadeforsikring."
Veileder: Erik Bølviken

Gunnhildur Högnadóttir Steinbakk: "Parameterestimering i en autoregressiv modell fra kategoriske data med støy."
Veileder: Geir Storvik

Nyansettelse

Arnoldo Frigessi er ansatt som professor II fra 1. august 2000.

Statistical Research Reports

Xeni K. Dimakos & Arnoldo Frigessi: "Hierarchial Bayesian premium rating with latent structure." No.1, 2000

Joe Sexton & Anders Rygh Swensen: " ECM algorithms that converge at the rate of EM." No.2, 2000

Inge S. Helland: "Quantum theory from symmetry and reduction statistical models. The compact case." No.3, 2000

Anders Rygh Swensen: "Bootstrapping unit root tests." No.4, 2000

Statistical Memoirs

Bent Natvig: "Have the Natural Sciences and Technology Forgotten Ethical Perspectives and Become Disciplines for Unaware Nerds?" No.1, 2000

Grete Usterud Fenstad

FORELØPIG ANNONSERING!

Norsk Statistisk Forening

KONFERANSE OG årsmøte

I Bergen

20.-22. juni, 2001

Det er planlagt et prekonferansekurs 18.-19. juni (eksakte metoder), med forelesning og PC-øvelser.

Hold av datoene - og følg med i TG for videre informasjon og påmelding

Vel møtt i Bergen neste år!

Hilsen

Bergen Statistisk Forening

International School on Mathematical and Statistical Applications in Economics

January 15–19, 2001, Västerås, Sweden

First Announcement

The International School on Mathematical and Statistical Applications in Economics will take place in Västerås, Sweden from 15 January to 19 January 2001. Participation in the School is open to researchers, students, practitioners in business and industry and other specialists from all countries.

It is organised in the frame of the EU Tacis-Tempus Project "Statistical Aspects of Economics" under the auspices of Mälardalen University, Umeå University, Stockholm University, University of Helsinki and Kiev University.

The programme covers Mathematical Economics in its wide sense as the scientific area of mathematical and statistical applications in economics and it includes topics on Financial and Actuarial Mathematics; Survey Sampling in Economics; Mathematical Models in Micro- and Macro-Economics; Computer Based Modelling in Economics; and Educational Programmes and Teaching in Mathematical Economics.

The programme will consist of invited lectures (50 minutes), communications (20-25 minutes) and software demonstrations. The working language is English.

Invited lectures will be given by Håkan Andersson (FöreningsSparbanken, Stockholm), Thomas Björk (Stockholm School of Economics), Kimmo Eriksson (Mälardalen University, Västerås), Jan Grandell (Royal Institute of Technology, Stockholm), Mats Gyllenberg (University of Turku), Sune Karlsson (Stockholm School of Economics), Anders Klevmarken (University of Uppsala), Gunnar Kulldorff (Umeå University), Anders Martin-Löf (Stockholm University), Harri Nyrhinen (University of Helsinki), Lars-Erik Öller (National Institute of Economic Research, Stockholm), Bengt Rosén (Statistics Sweden), Dmitrii Silvestrov (Mälardalen University, Västerås) and Alexander Kukush (Kiev University), Bengt Swensson (University of Örebro), Imbi Traat (University of Tartu), Mikhail Yadrenko and Nadija Zinchenko (Kiev University).

Abstracts of invited lectures and communications will be distributed to participants upon arrival at the school.

The form for a preliminary registration and submission of abstracts can be obtained from the web-page of the School, which address is given below. The deadline is **December 15, 2000.**

The registration fee is $50 US.

Proceedings are planned to be published after the school. Those wishing to submit their papers are asked to do this by sending LaTex file of the paper to e-mail address of the school and by sending two paper copies to the mail address of the school given below. The deadline for submission of the papers is **December 15, 2000.** The LaTex templates for papers can be obtained from the web-page of the School.

The venue for the school will be the Mälardalen University (Västerås).

Västerås is a nice Swedish town with interesting historical and tourist sites for visitors. It is located on the Mälardalen lake, 100 km west of Stockholm. The social programme will include a welcome reception, excursions and a banquet. Accommodation will be available on different price levels.

The detailed information about scientific programme, accommodation, social programme and ways to pay the conference fee will be included in the Second Announcement. It will be available before **November 1, 2000** at the web-page of the School.

Organising Committee: Dmitrii Silvestrov (Chairman, Västerås), Clas Nordin (Västerås), Evelina Silvestrova (Secretary, Västerås), Ingrid Westerberg-Eriksson (Umeå), Nadija Zinchenko (Kiev).

Scientific Programme Committee: Dmitrii Silvestrov (Co-chairman, Västerås), Anders

Martin-Löf (Co-chairman, Stockholm), Alexander Borisenko (Kiev), Kimmo Eriksson (Västerås), Gunnar Kulldorff (Umeå), Evelina Silvestrova (Secretary, Västerås), Esko Valkeila (Helsinki), Mikhail Yadrenko (Kiev).

Address of the School:

The International School on Mathematical and Statistical Applications in Economics
Mälardalen University
Department of Mathematics and Physics
Box 883, SE-721 23 Västerås
SWEDEN

Tel: +46 - (0)21 - 151701
Fax: +46 - (0)21 - 101330
E-mail: msae2001@mdh.se
Web-page:
http://www.ima.mdh.se/_msae2001-ima.htm

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-post

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