



ISSN 0803-8953

**TILFELDIG
GANG**
Nr. 3, årgang 25
Oktober 2008

Utgitt av Norsk Statistisk Forening

INNHold

	Side
Fra lederen, <i>Magne Aarset</i>	3
Fra redaksjonen, <i>Jo Eidsvik</i>	4
Artikler	
Rapport fra et forskningsliv på siden av hovedstrømningen i det norske statistikk-miljøet, <i>Arvid Næss</i>	5
Nytt <i>Centre for Conservation Biology</i> ved NTNU, <i>Steinar Engen</i>	6
Jubileum	
Jan F. Bjørnstad 60 år, <i>Johan Heldal</i>	7
Stipendiatpresentasjoner	
<i>Tore Selland Kleppe</i>	8
<i>Haakon Michael Austad</i>	10
Møter og konferanser	
Royal Statistical Society Ordinary Meeting	12
Referat fra NORDSTAT, <i>Mette Langaas</i>	13
Rapport fra Nerdeland, <i>Eirik Skogvoll</i>	15
Referat, Temakveld i Norsk statistisk forening, Oslo, 7. mai, <i>Ingunn Fride Tvette</i>	18
Puslerispalten	
Pusleri nr. 25, <i>Jostein Lillestøl</i>	19
Løsning til pusleri nr. 24, <i>Jostein Lillestøl</i>	20
Utlysninger	
Stilling i anvendt statistikk ved Norsk Regnesentral	21
Dataanalyse i praksis: En casesamling, <i>Jostein Lillestøl</i>	22
Meldinger	
Nytt fra Norges teknisk-naturvitenskapelige universitet	23
Nytt fra Universitetet for miljø og biovitenskap	25
Nytt fra Universitetet i Oslo, Avd. C	26
Nytt fra Universitetet i Oslo, Avd. for biostatistikk	28

NSF

Som kjent er vi en liten, men skal vi si celeber forening. Jeg føler det i alle fall som en ære å være medlem. Men også i små foreninger trengs det litt organisering, og såpass gode matematikere er vi alle at vi skjønner at det gjerne blir mer arbeid per person i en liten organisasjon enn i en stor. Så derfor til dagens tema:

Finnes det noen ”frivillige” der ute?

For tiden er det de følgende oppgaver som må fordeles:

- Arrangementskomité og Programkomité til NSF konferansen i 2009.
- Arrangementskomité til NORDSTAT-møtet i 2010.
- Programkomité til NORDSTAT-møtet i 2010.

Arrangementskomité og Programkomité til NSF konferanse i 2009.

Dette er et arrangement som går på omgang blant regionforeningene, og denne gangen er det Oslo-foreningen som står for tur. De har akseptert ansvaret.

Programkomité til NORDSTAT-møtet i 2010.

Her har Håvard Rue i Trondheim sagt seg villig til å lede programkomiteen, kanskje med flere deltakere fra samme regionforening. (Av rent praktiske grunner.) Men her trengs det flere deltakere.

Arrangementskomité til NORDSTAT-møtet i 2010.

Sist Norge arrangerte NORDSTAT var det regionsforeningen i Oslo (eller var det Blindern-miljøet?) som sto for arrangementet. Siden ”Tromsø” arrangerte siste nasjonale møte for NSF, ”Oslo” arrangerte siste NORDSTAT-møte, ”Trondheim” har tatt på seg ansvaret for programkomiteen, så går mine tanker nå mot vestlandet. (Dette er jo også den flotteste delen av Norge, med de kjekkeste, peneste og mest kultiverte innbyggerne.) Bergensavdelingen har få aktive medlemmer, så per dags dato er det imidlertid ikke avklart om Bergens-miljøet kan stå for dette. Men forsterket med noen erfarne, noen ivrige og noen som er begge deler, så er det utrolig hva som går an.

Vi trenger altså støtte til de miljøene som får ansvaret for de to komiteene angående NORDSTAT-møtet. Og det er her vi trenger ”frivillige”. Både nye og ferske fjes (det er både lærerikt og morsomt, samtidig som det kan være bra å ha på CV-en at man har vært med på et slikt arrangement) og ”gamle etablerte” med erfaring, ønskes velkommen.

Ta gjerne kontakt med meg (4641 0798 eller magne.aarset@bi.no) eller noen av de andre i styret for å melde deres interesse. Det kan innrømmes at vi ikke har arrangert ekstra telefonvakt for å ta mot alle henvendelsene, men håper at noen stiller opp slik at vi slipper å ta i bruk all vår overtalingsevne.

Jeg håper og tror dette blir flotte arrangement både i 2009 og 2010! På forhånd takk til alle som kommer til å bidra!

Magne

Fra redaksjonen

Etter en deilig sommer er høsten i gang. Høstsemesteret er kort og intenst. Dagene blåser forbi – ideer og planer skrinlegges for å utføre kurs, studentveiledning og administrasjon. Vel, vel; 'It's called life!' Selv har jeg dette semesteret litt ekstra i form av et kurs i pedagogikk, noe som er obligatorisk for fast ansatte ved NTNU. Kurset omfatter noen hel- og halvdags samlinger, noen møter med kolleger for å høre på hverandres forelesninger, og et prosjektarbeid på ca 15 sider. Så kurset tar mye tid, men er også interessant på mange måter fordi man reflekterer mye over egen undervisning/veiledning og ser hvordan andre fagområder legger opp sin undervisning. En forskjell er at vi har mer tradisjonell undervisning. Spesielt i lavere grads fag er det jo hos oss fokus på 'overføring' av kunnskap, med noe 'konstruktiv' student-lærer kontakt, mens en del andre fag har en mer 'sosiokulturelt' læringsform, dvs læring gjennom deltakelse og samspill. Vår gruppe ved matematiske fag, NTNU, har ansvar for 11 grunnkurs i statistikk og sannsynlighetsregning, dvs ca 1 grunnkurs per fast ansatt per år. Her fyller vi på (fra tavla), mens studentene stort sett lytter. Studentene har imidlertid muligheten til diskusjon og interaksjon i små grupper.

Til min grunnkursforelesning i forrige uke kom det vitenskapelige tilhørere fra arkitektur, tverrfaglige kulturstudier og statsvitenskap, med oppgave å vurdere meg/vårt system. Tema for denne timen var kontinuerlige fordelinger, spesielt normalfordling. Timen var en 'gjennomsnittlig' forelesning med mye bruk av tavle, noen slides, litt datamaskin, og et par diskusjoner med salen. Etterveiledning viste at Normalfordeling var vanskeligere enn antatt: De eksterne ville ha mer relevans, og mindre av disse integraltegnene, skift /skalering og slikt. De mente videre at det ble for mye tavle, mens de likte mine eksempler på fødselsvekt til barn og måling av pH. Enig, - alt må motiveres med enkle begreper, men håndverket / regningen skal jo også læres.... Og tavla er fortsatt populær blant våre studenter.

I dette nummeret av TG har vi holdt på de faste spalter: jubileum, presentasjon fra PhD kandidater, to senterbeskrivelser, puslerier med løsninger, og to brev fra konferanser. I tillegg fortsetter klimadebatten med et referat fra temamøtet i Osloforeningen. **Lederen ber om hjelp til arrangementer:** Det første forestående er Norsk Statistikermøte 2009, det andre forestående er Nordisk Statistikk konferanse i 2010.



Foto: Jørn Adde

Vi tar gjerne imot flere bidrag til TG. Vær aktiv!

Neste nummer av TG kommer først i 2009. Frist for innsending er 15 Januar 2009.

God høst!

Redaksjonen v/Jo Eidsvik

En liten rapport fra et forskningsliv på siden av hovedstrømningen i det norske statistikk-miljøet.

Arvid Næss, professor i statistikk ved NTNU (www.math.ntnu.no/~arvidn)

I 2001 var jeg i ferd med å skifte beite. Jeg hadde vært professor i konstruksjonsteknikk ved NTNU/NTH i nesten 15 år, og hadde lyst til å bruke mer tid på en del fundamentale problemstillinger innen bl.a. ikke-lineær stokastisk dynamikk. Det var ikke så enkelt å få til det ved Institutt for konstruksjonsteknikk – man er jo tross alt avhengig av gode doktorstudenter. Jeg søkte derfor en stilling i statistikk ved Institutt for matematiske fag (IMF), hvor jeg tidligere hadde tatt en doktorgrad i matematikk. Omtrent på samme tid ble jeg dratt med i en søknad om en SFF (senter for fremragende forskning) innen skips og havkonstruksjoner. Et av de temaene man ønsket å flagge i søknaden var forskning rettet mot probabilistiske metoder for beregning av slike konstruksjoner. Dette var et område jeg hadde holdt på med i nesten to tiår, og hadde hatt en smule suksess med. Jeg ser ikke helt bort fra at det var grunnen til at jeg ble invitert med. I alle fall, i 2003 var **Centre for Ships and Ocean Structures (CeSOS – www.cesos.ntnu.no)** i full sving, og jeg har siden den gang brukt svært mye av min forskningstid der – litt for mye har kanskje enkelte i statistikkgruppa ved IMF tenkt.

Det er minst to grunner til at jeg tilbringer så mye tid på CeSOS. Den viktigste er nok innholdet i forskningsprosjektene, som passer meg veldig bra. Dessuten er ressursituasjonen upåklagelig. Med et årlig budsjett på rundt 30 mill. kr. er ikke pengemangel et problem.

For å gi et inntrykk av den forskningen ved CeSOS som kan være av interesse for en leser av TG, kan jeg kanskje starte med de tematiske områdene hvor statistikk og sannsynlighetsteori klart kommer til anvendelse: 1) Stochastic dynamic analysis of extreme load effects, 2) Fatigue damage due to nonlinear effects, 3) Uncertainty analysis and structural reliability.

Min innsats, i tett samarbeid med et par postdoc'er, er i hovedsak rettet mot problemstillinger innen 1) og 3). Vi har de siste par årene jobbet mye med å utvikle metoder for effektiv prediksjon av ekstrem respons for systemer hvor det er mulig å gjøre Monte Carlo simuleringer. Disse metodene har åpnet for mange muligheter for ekstremverdi-prediksjon som tidligere var meget besværlig eller umulig å gjennomføre. En annen erkjennelse, som er blitt forsterket gjennom dette arbeidet, er at de vanlige metodene for ekstremverdianalyse, som baserer seg på bruk av de asymptotiske fordelingene, ofte er lite hensiktsmessige. For praktisk ekstremverdianalyse er det nok ofte det subasymptotiske regimet man bør fokusere på heller enn det asymptotiske. Dette er et viktig element i våre metoder.

I tillegg til metoder for ekstremverdianalyse, har jeg heldigvis også hatt muligheten til å arbeide videre med problemstillinger innen stokastisk dynamikk. Vi har spesielt konsentrert oss om å utvikle beregningsmetoder basert på numerisk vei-integrasjon (path integration - PI). Oppmerksomme lesere av TG vil ha sett at Eirik Mo for ikke så lenge siden forsvarte sin dr.-avhandling innen dette temaet. Han gjorde en kjempejobb også med å utvikle godt dokumentert programvare for beregninger basert på PI. Dette har gjort det enkelt for andre samarbeidspartnere å bruke denne programvaren, som nylig resulterte i en artikkel i Phys Rev E om stokastisk populasjonsdynamikk. Numerisk PI har også et stort potensiale innen finans, som vi så smått har begynt å utnytte.

Nytt *Centre for Conservation Biology* ved NTNU

Ved NTNU er det opprettet et forskningssenter for bevaringsbiologi. Senteret er lokalisert ved Institutt for Biologi ved det Naturvitenskapelige Fakultetet og drives i samarbeid med Institutt for Matematiske Fag. Som det fremgår av senterets egen presentasjon (<http://www.ntnu.no/ccb>) er stokastisk modellering en grunnleggende komponent i forskningsprofilen:

'The vision behind the Centre for Conservation Biology at NTNU is to develop an interdisciplinary centre for research into dynamical changes in biological diversity at different organismic levels (genes, populations and communities), based on a common theoretical framework.

A strategic goal for the Centre will be to extend existing theory by developing stochastic dynamic models of genetic and demographic mechanisms in populations and in communities of interacting species. The assumptions and predictions of these models will be tested using a set of carefully selected empirical model systems, in both the field and laboratory.

A second strategic goal will be to use models and empirical systems to address basic questions in ecology and evolution, as well as practical problems related to conservation of biological diversity at the national and international levels'.

Senteret har finansiell støtte fra NTNU som har nedsatt et styre med representanter fra Norsk Institutt for Naturforskning, Artsdatabanken, Direktoratet for Naturforvaltning, Institutt for Biologi og Institutt for Matematiske Fag. Styreleder er Eivind Røskaft og daglig leder for senteret er Bernt-Erik Sæther.

Det er idag 39 forskere ved senteret, 15 faste vitenskaplige, 15 post doktoter og 13 stipendiater (<http://www.ntnu.no/ccb/people>). De største prosjektene som er igang er et strategisk universityetsprogram (SUP) *Strategic University Programme in Conservation Biology*, et Storforsk program *Population genetics in an ecological perspective* og et klima-prosjekt *Verteklim The effect of climatic variation on vertebrate populations*.

Jan Fredrik Bjørnstad 60 år

Jan Bjørnstad fylte 60 år den 13. juni. For dem som ikke kjenner ham er han forskningssjef ved Seksjon for Statistiske Metoder og Standarder (SSMS) i Statistisk Sentralbyrå (SSB) hvor han rår over en gruppe på 17 personer (inkl. Jan selv), hvorav 14 statistikere. Det er et av de største statistikermiljøene i Norge.

Jan er vokst opp på Strømmen rett øst for Oslo. Han tok hovedfag ved Universitetet i Oslo i 1973. Etter en sommerjobb i SSB samt vit-ass stillinger ved sosialøkonomisk institutt og matematisk avd C og militærtjeneste, dro han i 1975 til Berkeley hvor han i 1978 disserterte under Lehmann med en avhandling som bar tittelen "On Optimal Subset Selection Procedures". Deretter fikk han jobb som assistant professor i statistikk, først ved Purdue University og i 1980 ved University of Texas at Austin. I 1983 valgte han imidlertid å vende nesen mot kaldere klima og ble professor i matematisk statistikk ved Universitetet i Tromsø i 1985. Han ble der en del av det den gang svært sterke statistiske miljøet sammen med Tore Schweder og Erik Bølviken. Det var mens han var der han utviklet interesse for feltet "utvalgsundersøkelser" og som senere førte ham til SSB. Han har så langt vært den eneste universitetsstatistiker i Norge som har utviklet særskilt interesse for dette feltet som har så mange anvendelser i det moderne samfunn. Jan ikke bare interesserte seg for feltet. Han utviklet sitt eget teoretiske rammeverk for hvordan inferens basert på utvalgsundersøkelser skulle gjøres, kalt "Prediktiv Likelihood". For den interesserte henvises til artikkelen "Predictive Likelihood: A Review", Statistical Science, Vol 5 no. 2 (1990).



I 1991 flyttet han til Trondheim og Institutt for Matematikk og Statistikk. I perioden 1991-2001 hadde han en professor II stilling ved Universitet i Tromsø. Samtidig kom han til SSMS i en 20% stilling som han hadde i fire år. I 1995 begynte han som forsker 0 ved SSMS. I 1998 ble han professor på Ås for så å komme tilbake som forskningssjef i 2001 da stillingen som sjef for SSMS ble ledig.

Jan har bidradd sterkt til å heve det faglige nivået i SSMS, både ved sin egen faglige innflytelse og ved å verve noen av sine tidligere hovedfags- og doktorgradsstudenter til seksjonen. Det forhold at han var blitt sjef var medvirkende til at jeg selv vendte tilbake til SSMS etter noen års fravær.

Jan er ellers svært interessert i musikk, spesielt klassisk og gammelpop, og sport. Fra tiden i Tromsø er det fremdeles TIL som er favorittlaget i fotball.

I juni i år kjøpte han hus på Strømmen. Dermed kan vi si at ringen er sluttet også hva bopel angår. Vi i SSMS setter stor pris på ham som seksjonens leder og håper (og tror) at vi får beholde ham som sjef så lenge som mulig og at han om to år ikke vil benytte sine muligheter til avtalefestet pensjon og tre tilbake.

Johan Heldal

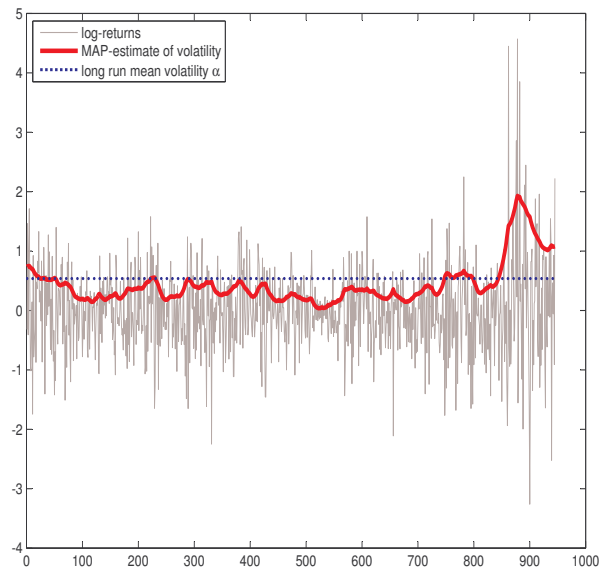
Selvpresentasjon: Tore Selland Kleppe

Planlagt disputas: høst 2010

Matematisk Institutt, Universitetet i Bergen

Veileder: Professor Hans J. Skaug

Etter oppvekst i Stavanger og militærtjeneste dro jeg til Trondheim for å studere på siv.ing. linjen fysikk og matematikk. I utgangspunktet hadde jeg tenkt å spesialisere meg innen teknisk fysikk, men etter to år falt valget til slutt på industriell matematikk. Det tre årene på industriell matematikk gikk med til en blanding av optimeringsteori, statistikk og stokastiske differensiallikninger. Det hele kumulerte med masteroppgaven «Numerical path integration for Lévy-driven stochastic differential equations» med professor Arvid Næss som veilder.



Etter Trondheim vendte jeg nesen mot Bergen og doktorgradsstudier på UiB med professor Hans J. Skaug som veileder. Den røde tråden i dr. prosjektet er likelihoodberegninger i modeller med latente variable. I hovedsak tar dette utgangspunkt i Sadelpunkt- eller Laplace approksimasjoner.

Det første arbeidet vi gjorde handler om bygging og tilpassing av latent variabel modeller ved hjelp av den momentgenererende funksjonen (MGF). I mange typer modeller er det mulig å finne enkle uttrykk for MGF selv om likelihoodet er vanskelig tilgjengelig. Eksempler er latent variabel regressjoner med ikke-Gaussiske latente variable, ikke-Gaussiske state-space modeller og stokastiske nettverk. Vi brukte en ikke-Gaussisk basert sadelpunkt approksimasjon til å estimere likelihoodet, og dette virker som god generisk metode for et ganske bredt spekter av modeller. Sammenliknet med Laplace approksimasjoner er dette som et godt alternativ i modeller med flere latente variable enn observasjoner siden det er beregningsmessig raskere.

Veien videre har blitt parameterestimering i stokastisk volatilitet modeller med ikke-Gaussisk latent prosess. I dette arbeidet utnytter vi at Laplace approksimasjonen ikke er invariant under ikke-lineære variabeltransformasjoner til å lage raskt konvergerende importance samplere. Vi har utviklet en kode som raskt kan tilpasses forskjellig modellspesifikasjoner. Spesielt er vi interessert i kontinuerlig tid modeller slik som Heston modellen. Pr. tid er denne metodikken prøvd ut på nokså enkle eksempelmodeller, men vi håper å kunne utnytte den i en mer empirisk kontekst etterhvert.

I tillegg til arbeidet med prosjektet kommer pliktarbeidet. Jeg har jobbet med

organisering og veiledning av dataøvinger i «Elementær statistikk» og vært øvingsleder «Statistisk inferensteori» de siste to årene. I tillegg har jeg laget og avholdt et kort innføringskurs i programpakken R for de nye masterstudentene. Dette har vært lærerikt og nyttig, og mye morsommere enn jeg trodde på forhånd.

Selvpresentasjon Haakon Michael Austad

Institutt for matematiske fag, NTNU.

Planlagt innlevering våren 2011.

Foreløpig tittel på avhandling: Stochastic modelling of discrete spatial variables for applications in petroleum reservoir modelling.

Høsten 2002 flyttet jeg opp fra Oslo til Trondheim for å starte på sivilingeniør-graden min på linjen for fysikk og matematikk. I løpet av de fem årene fant jeg ut at det var matematikk, og da spesielt statistikk som var mest gøy å holde på med. (Bildet er forøvrig hentet fra tredje året mitt når vi drog på klassetur til Beijing.) Jeg trivdes



også så godt i byen at jeg ikke behøvde å tenke meg så lenge om når jeg mot slutten av femte året fikk tilbud om en stipendiatstilling til høsten.

Arbeidet mitt inngår som del av Norsk Regnesentral (NR) sitt prosjekt "Multipoint Methods for Improved Reservoir Models" og er del av Petromaks forskningsprogrammet. Her på Gløshaugen jobber jeg mest med veilederen min Håkon Tjelmeland, og de andre i statistikk-gruppa. Vi er også regelmessig i kontakt med folkene på NR som blant annet inkluderer bi-veilederen min Odd Kolbjørnsen. Mye av tiden siden jeg startet for et år siden har gått med til pliktarbeid og fag, men jeg har også rukket å komme litt igang med andre ting. Fokusområdet for arbeidet mitt er gridbaserte modeller for modellering av diskrete reservoarvariabler, med tilhørende parameterestimering og simuleringsmetoder. Mer spesifikt så konsentrerer vi oss for øyeblikket om Markov Random Field (MRF) modellklassen. MRF modeller drar nytte av å være svært fleksible modeller, men benyttes lite i reservoarmodellering hovedsakelig siden simulering og parameterestimering krever store data-ressurser.

Til nå har vi sett litt på forover-bakover algoritmen, som gjør det mulig for oss å beregne likelihood og simulere på små grid. Dette lar oss igjen gjennomføre pseudo-blokk MLE og kjøre blokk-Gibbs simuleringsalgoritmer. Vi har sett på dette i kombinasjon med en multi-grid modell. Multi-grid ideen handler i prinsippet om at vi deler et grid opp i flere sub-grid eller nivåer, og så definerer en serie av modeller, en for hvert nivå. På denne måten prøver vi å fange opp interaksjoner mellom celler som ligger et stykke fra hverandre uten å benytte store naboskapsmodeller.

Nylig har vi jobbet med en måte å approksimere normaliseringskonstanten i en MRF modell. Ved å velge en alternativ representasjon av normaliseringskonstanten kan vi kjøre forover-bakover algoritmen på en approksimativ modell mye raskere og med mye lavere minnebruk enn før, som igjen lar oss takle mye større grid. Dette er veldig "work

in progress”, men de foreløpige resultatene ser ganske lovende ut.

Foreløpig er alt vi gjør i 2D, men vi ønsker å gå over til 3D etterhvert. Planen er også å begynne å betinge på data, da en av fordelene med å forholde seg til MRF modeller er at dette lar seg gjøre på en konsistent måte, i motsetning til i algoritmedefinerte modeller, hvor dette er vanskelig

Ordinary Meeting

Organized by the Research Section

“Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations”,

by Håvard Rue, Sara Martino (Norwegian University of Science and Technology) and Nicolas Chopin (CREST-LS and ENSAE, Paris)

Wednesday 15 October 2008, 5pm

RSS Headquarters, 12 Errol Street, London



Summary

By using an integrated nested Laplace approximation and its simplified version, we can directly compute very accurate approximations to the posterior marginals. The main benefit of these approximations is computational: where MCMC algorithms need hours or days to run, our approximations provide more precise estimates in minutes or seconds.

The paper can be downloaded in advance from:
<http://www.rss.org.uk/events>

Fellows and visitors are invited to make contributions of no longer than 5 minutes following the reading of the paper. Please register your interest with the Secretary of the RSS Research Section Committee (e.wit@lancaster.ac.uk) or C. Stovell at the RSS (c.stovell@rss.org.uk).

Tea is served at 4:30pm. **The meeting will be held at the RSS Headquarters in London:**

12 Errol Street
 (close to Moorgate & Farringdon Underground/Railway stations)
 London EC1Y 8LX
 Tel: +44(0)20 7638 8998

For more information, contact Charlotte Stovell at the RSS (c.stovell@rss.org.uk).

After the meeting you can meet the authors over **dinner**. Book your place in advance: Teresa Dewane (tel. 020-7614 3910, t.dewane@rss.org.uk), price £20 (£10 for student RSS members).

NORDSTAT 2008

16. - 19. juni 2008 var det duket for den 22. nordiske konferansen i matematisk statistikk (NORDSTAT). Denne gangen ble konferansen holdt utenfor Norden, i Vilnius i Litauen. De baltiske landene er nå innlemmet i NORDSTAT-samarbeidet. Derfor, som lederen for programkomiteen for NORDSTAT 2008, Vygantas Paulaskas (bildet), fortalte om i sin åpningstale, vil konferansen hvert 10. år bli avholdt på Baltikum. Det vil gå på omgang mellom Litauen, Latvia og Estland (så neste gang det er Litauens tur er trolig i 2038).



Foto: <http://www.nordstat2008.com>

Programkomiteen hadde funne frem til fire ypperlige plenums-foredragsholdere. Kjersti Aas (Norsk Regnesentral) fortalte oss hvordan vi kan jobbe multivariat med data som har

avhengighetsstrukturer med annen type

haleavhengighet enn det vi finner for multivariat normalfordling. Hennes byggestener var par-kopula konstruksjoner. Willem van Zwet (University of Amsterdam, Nederland) hadde i sitt foredrag fokus på William Gosset og Student-t fordelings viktighet i statistisk inferens. Hans Künsch (ETH Zürich, Sveits) satt klimaprediksjoner på dagsorden og Aad van der Vaart (Vrije Universitet, Nederland) tok oss med inn i uendeligdimensjonale parameterrom.

I tillegg til de inviterte plenumsforedragene var det ikke mindre enn 12 tema-seksjoner med inviterte og påmeldte foredrag. Flere detaljer om programmet kan du finne ved å klikke deg inn på konferansens www-side <http://www.nordstat2008.com/>.



Konferansen ble avholdt på et hotell like ved sentrum av Vilnius, med nydelig lunsjservering og godt ventilerte konferanserom i ulike størrelser. Tonen ved konferansen var gemytlig, og mange benyttet seg av muligheten til å stille spørsmål til foredragsholderne. Presentasjonene var generelt svært godt forberedt og profesjonelt fremført.

Konferansemiddagen ble avholdt på restauranten Belmontas, et lite stykke utenfor Vilnius, med levende folkemusikk fra et lituansk band. Konferansedeltakerne fikk mulighet til å vise frem sine ferdigheter innen ulike typer ringdans, og den norske delegasjonen viste en

imponerende innsats. Maten smakte også ypperlig. Trygve Nilsen tok under konferansemiddagen ordet på vegne av lederen av Norsk Statistisk Forening, og inviterte alle delegatene til NORDSTAT 2010 i Norge.

Konferansen var godt besøkt med 180 påmeldte deltakere. Den norske kontingenten talte rundt 30 statistikere og statistikere in spe, der en stor andel var studenter fra Oslo og Bergen. Dette tegner meget godt for rekrutteringen til fremtidige NORDSTAT konferanser.

«Alle var enige om at det hadde vært en fin tur.» Det var så fint, at flere delegater ble igjen i Vilnius i noen dager etterat konferansen var avsluttet - eller hadde sett Baltikum på vei til konferansen. Dette setter selvfølgelig lista høyt for arrangementskomiteen og programkomiteen for NORDSTAT 2010.



Foto: <http://www.nordstat2008.com>

Mette Langaas

Av Eirik Skogvoll, 1.amanuensis, Enhet for anvendt klinisk forskning, Det medisinske fakultet, NTNU

Nerdeland lå sommeren 2008 ikke i Nederland slik en kanskje kunne tro, men i Dortmund, Tyskland. Men i sommerferien? "R-user forum 2008" høres ut som den mest nerdete form for ferie en kan tenke seg!

Hva hadde jeg i vente? Redaktør Langaas i TG gav meg gode råd: "Anlegg hestehale, ta på svart t-skjorte, ha lap-topen åpen og nistirr på den slik at det er tydelig for alle at du er veldig opptatt". At dette ikke bare var tomt prat framgikk av R-help (www.r-project.org) 08. juli 08: "R groupies ...? I have this mental image of a bunch of geeks dressed in tie-dyed shirts, jeans and sandals, sitting on the corner of Haight-Ashbury listening to the Grateful Dead while coding on their laptops..." Planene om å anlegge hestehale ble fort lagt bort, men laptop'en tok jeg med. Forberedte meg på å holde samtalen under festmiddagen i gang; *do you feel comfortable with the new package ... is there really a class struggle with the new S4 ... what about some lexical scoping later tonight...*

Vel, "R-user forum" er helt klart for de spesielt interesserte, og samler R-entusiaster fra hele verden hvert år. Selv om Internett klart er den store kilden til utvikling av dette fantastiske, internasjonale, dynamiske, *open source* miljøet for dataanalyse og visualisering, er det behov for å møte folk personlig. Jeg har vært flittig bruker av R og lest R-help siden 2004. Så det var med spenning jeg så fram til både å se og kanskje veksle noen ord med sentrale (*core*) notabiliteter som Peter Dalgaard, Duncan Murdoch og Frank Harrell jr.; ja kanskje til og med den til tider nokså *strenge* Brian Ripley fra Oxford ("You didn't tell us ..."). Men han kom ikke.

En rekke halvdags kurs ble arrangert før selve samlingen. Jeg valgte to: Douglas Bates fra Madison, USA med "Linear, generalized and nonlinear mixed models" og Frank Harrell "Regression modeling strategies". Som halvstudert røver (medisiner) i medisinsk statistikk er det særlig lineære modeller, og spesielt lineære sammensatte modeller som har vist seg nyttige i praksis (hva kaller man egentlig "fixed" og "mixed" på norsk?) Bates' innlegg var tidvis ganske teoretisk og innviklet, men han ga en del verdifulle tips i praktisk bruk av programpakken *lme4* som nå foreligger i versjon 0.999375-11 (!). Her benyttes ny syntaks ift. den foregående (*nlme*), og man nå kan modellere kryssede, tilfeldige faktorer. Dette er en

verdifull utvidelse sammenliknet med en streng hierarkisk modell. Jeg grep tak i Bates i en pause og spurte om noen råd i tolkningen av en aktuell modell vi jobber med; han var både vennlig og imøtekommende om enn litt travel.

Frank Harrell jr. fra Vanderbilt i Tennessee, USA, hadde en meget rask gjennomgang av sine synspunkter på regresjonsmodellering. I korte trekk hevdet han – ikke uventet – at modeller i alminnelighet er langt mer anvendelige og nyttige enn diverse ”tester”, og at man i randomiserte studier bør søke å justere for ubalanserte kovariater. Men også at man i søken etter den rette modellen bruker opp sine frihetsgrader uten at det nødvendigvis kommer fram i det endelige resultatet. Videre anbefalte han generelt bootstrapping for å validere variansestimaterne for den modellen man til slutt ender opp med. Hans bok ”Regression modelling strategies” (Springer) ventes i revidert utgave i 2009.

Selve konferansen på 3 dager var som sagt for de spesielt interesserte, og deltakerne fikk vel størst utbytte av å høre om det de selv jobbet med. Program og abstracts ligger på nettet. Dels var sesjonene tematisk definert (biostatistikk, økonometri, økologi, robust estimering), dels mer teknisk preget (hvordan håndtere store filer eller utvikle R-pakker). Noen ganger ble diskusjoner på ”engelsk” mellom engasjerte deltakere fra den vide verden heller vanskelig tilgjengelig. Av generell interesse var bl.a. John Fox sitt innlegg om sosiologien og utviklingen av R-prosjektet. Bl.a. utviser tilfanget av nye pakker på CRAN (som både kan uttales ”C-ran”, og ”cran” fant jeg ut) tilnærmet eksponentiell vekst. En egen sesjon om læring hadde mye interessant, bl.a. ble pakken *animation* demonstrert og ga meg noen gode ideer. Videre var det interessant å høre Andrew Gelman fra USA diskutere ”weakly informative prior distributions” i bayesiansk logistisk regresjon med utgangspunkt i funksjonen *bayesglm* fra pakken *arm*. Han hevdet bl.a. at estimerte koeffisienter (på logskala) sjelden overstiger 5 i absoluttverdi (røyking ble nevnt som eksempel) og at en bør se med skepsis på større estimater.

Det sosiale fungerte bra, og det tyske vertskapet med Uwe Ligges i spissen var meget gjestfrie. Rikelig mat og drikke var tilgjengelig hele tiden! Kveldene gikk med til ”welcome mixer”, mottagelse hos ”Der Überbürgermeister” i Dortmund og festmiddag i VIP-området på Signal Iduna Park stadion; hjemmebanen til Borussia Dortmund (80 000 sitteplasser). Som det tidligere er hevdet, ”even the mean statisticians are nice people”! Jeg trengte følgelig ikke de innøvde frasene for å få samtalen til å gå. Ellers kunne Dortmund Universität by på en artig svevebane (”H-bahn”) mellom sine to campuser.

Neste år, i juli, arrangeres R-User forum i Frankrike. Bør *du* reise? Hvis du er spesielt interessert!



Frank Harrell jr. fra Vanderbilt foreleser om regresjonsmodeller. I bakgrunnen skimtes Bates' R-pakke lme4, som antas å konvergere numerisk mot 1.



Dortmund Universität har en artig svevebane!



Bildene fra konferansen ble ikke så gode. Derfor: forfatteren i vesentlig mer romantiske omgivelser ved Alanya, Tyrkia.

Referat, Temakveld i Norsk statistisk forening, Oslo, 7 mai.

Vårens møte i Osloavdelingen til Norsk statistisk forening tok for seg et populært tema, nemlig klima. I strålende solskinn og nesten sommertemperaturer stilte nærmere 30 personer opp for å høre på Terje Berntsen fra CICERO og Institutt for Geofag, UiO, og Ola Haug fra Norsk Regnesentral. Vi startet som vanlig med Pizza og forfriskninger, og Peppes Pizza leverte som før på minuttet til avtalt tid.

Første foredragsholder var Terje Berntsen. Tittel på foredraget hans var Klimaendringer, prognoser og usikkerhet. Allerede i 1896 presenterte en svensk forsker, Svante Arrhenius, beregninger som anslo at en dobling av CO₂ konsentrasjonen vil kunne øke den globale temperaturen med 5-6 grader celsius. Å diskutere mulige klimaendringer er altså ikke noe nytt. Berntsen gav en innføring i hva en klimamodell er, og hvordan de ulike elementene i modellen blir validert. Klimafølsomhet betegner endring i globalt midlet likevektstemperatur ved bakken ved en dobling av CO₂mengden i atmosfæren., og vi fikk forklart hvorfor denne er komplisert å beregne. Vi fikk videre en gjennomgang av IPPCs konklusjon som sier at ”det er meget sannsynlig at de observerte klimaendringene i hovedsak er menneskeskapte”, og usikkerhetsmomentene knyttet til disse.

Ola Haug, neste foredragsholder, snakket om klimaendringer og hvilke konsekvenser dette vil kunne ha for forsikringsselskapene. I et samarbeidsprosjekt mellom Gjensidige og Norsk Regnesentral har man sett på hvordan bygningsskader kan relateres til en rekke lokale værvariable over hele Norge. Skadestørrelser og -frekvenser på kommunenivå fra Gjensidiges portefølje og værdata fra Meteorologisk institutt har blitt analysert. Ola presenterte en statistisk modell som kobler antall skader mot værvariablene, og viste prediksjoner for skadenivået i framtida basert på ulike klimascenarioer.

Begge foredragsholdere håndterte, til tider krevende, spørsmål fra publikum elegant. Ut i fra tilbakemeldingene fra mange av tilhørerne var dette en vellykket temakveld. Jeg vil igjen benytte anledningen til å takke foredragsholderne for flotte foredrag.

Ingunn Fride Tvette.

Pusleri nr. 25

Du har to fyrstikkesker med n fyrstikker i hver eske. Hver gang du trenger en fyrstikk velger du en av eskene tilfeldig, og bruker en fyrstikk fra denne. Slik fortsetter du, uten å holde rede på hvor mange fyrstikker du har brukt, inntil den eske du åpner er tom. Hva er forventet antall fyrstikker i den andre esken?

Løsning til: jostein.lillestol@nhh.no

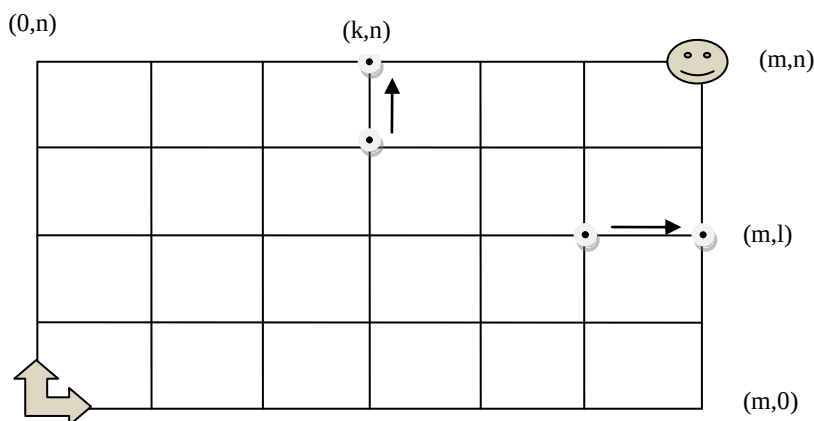
Pusleri nr. 24 (løsning)

Problemstillingen var:

Du bor i en by med kvadratisk gatestruktur, og vil besøke Per som bor m kvartaler øst og n kvartaler nord for deg. Alle gangfelt er lysregulert med vekslende grønt og rødt lys, og møter du et rødt lys i gangretningen, som er øst evt. nord, skifter du gangretning uten stopp til nord evt. øst, krysser gaten, og fortsetter i samme retning etter samme prinsipp, så lenge en slik mulighet fortsatt finnes. Anta at trafikklysene er usynkroniserte, og at rødt eller grønt i gangretningen er like sannsynlig. Hva er forventet antall nødvendige stopp for rødt lys inntil du er framme hos Per? Angi spesielt tilfellet $m=n$.

Løsning:

Du skal fra $(0,0)$ til (m,n) i diagrammet, og kan alltid skifte retning, uhindret uten stopp inntil du kommer til randen av byen i nord (k,n) eller i øst (m,l) .



Fra disse to punkter er forventet antall stopp henholdsvis $(m-k)/2$ og $(n-l)/2$. Vi betinger derfor mht. hvor en først treffer randen. Sannsynligheten for å treffe randen (k,n) langs en bestemt sti er opplagt $(1/2)^{n+k}$, og antall mulige stier av dette slaget som ikke treffer randen før, er

$$A(k,n) = \binom{k+n-1}{n-1}.$$

fordi dette må være lik antall mulige måter å komme til $(k, n-1)$ på, og derfra direkte til randen. Setningen om dobbeltforventning gir derfor forventningen lik

$$E(m,n) = \sum_{k=0}^{m-1} A(k,n) (1/2)^{n+k} (m-k)/2 + \sum_{l=0}^{n-1} A(m,l) (1/2)^{m+l} (n-l)/2$$

Tilfellet $m=n$ kan forenkles, og gir (tilnærmingen med Stirling's formel)

$$E(n,n) = \sum_{k=0}^{n-1} \binom{n+k-1}{n-1} (1/2)^{n+k-1} (n-k)/2 = n \binom{2n}{n} (1/2)^{2n} \approx \sqrt{\frac{n}{\pi}}$$

Summeringen krever litt omhu, med mindre man gjenkjenner følgende:

$$P(X = k) = \binom{n+k-1}{n-1} \left(\frac{1}{2}\right)^{n+k-1} \quad k = 0, 1, \dots, n-1$$

som en fordeling knyttet til myntkast og kjenner dens forventning. Gjør du?

Norsk Regnesentral (NR) er et ledende forskningsinstitutt innen anvendelser av statistisk-matematisk modellering og IKT. Instituttets 74 ansatte arbeider med anvendt forskning på oppdrag for næringsliv og offentlig sektor i Norge og internasjonalt. Vi skaper forskningsresultater som brukes og synes.



Bli del av Nordens største oppdragsmiljø i ANVENDT STATISTIKK

**FINANS • RÅVAREPRISER • MEDISIN • MILJØ
KLIMA • MARINE RESSURSER**

Avdelingen SAMBA søker etter både erfarne og nyutdannede statistikere med

- Solide teoretiske kunnskaper innen statistisk analyse og modellering
- Erfaring fra og/eller interesse for å jobbe med anvendte problemstillinger
- Programmeringskompetanse (C/C++, S-plus)
- Gode karakterer
- Entusiasme og gode samarbeidsevner

Vi tilbyr

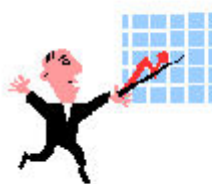
NR tilbyr spennende og varierte prosjekter og hyggelige kollegaer i et unikt fagmiljø. Mange av våre oppdragsgivere er ledende innen sine felt. Dette gir oss muligheter til å samarbeide med svært dyktige miljøer og personer. Vi har både kortsiktige løsningsorienterte prosjekter og langsiktige innovasjonsutfordringer med mye metodeutvikling.

I tillegg kan vi tilby konkurransedyktig lønn, pensjons- og forsikringsordning, bedriftshelsetjeneste og mulighet for barnehageplass. NR holder til ved Forskningsparken på Blindern i Oslo.

Kontaktpersoner

Forskningssjef André Teigland, 2285 2675, eller ass. forskningssjef Kjersti Aas, 2285 2694.

Søknad med CV, vitnemål og attester sendes HR-kontoret innen 1. november 2008, per e-post som én pdf-fil til human.resources@nr.no.

**Jostein Lillestøl : Dataanalyse i praksis: En casesamling**

som kan egne seg til bruk i økonomisk-administrative studier. Samlingen omfatter 20 cases med problembeskrivelse, datafiler og løsningsforslag, og er fritt tilgjengelig for interesserte fra internetsidene til Institutt for Foretaksøkonomi ved Norges Handelshøyskole, direkteadresse

<http://www.nhh.no/Default.aspx?ID=3221>

Nytt fra NTNU

Nye Master

Tony Gjerde: 'A heavy tailed statistical model applied in anti-collision calculations for petroleum wells'.

Veileder: Jo Eidsvik, Bjørn Bruun (StatoilHydro) og Erik Nyrnes (StatoilHydro).

Mads Herdahl: 'Linear mixed models for compressor data'.

Veileder: Jo Eidsvik.

Ingrid Østgård Jensås: 'A blockiness constraint for seismic AVA inversion'.

Veileder: Jo Eidsvik og Ulrich Theune (StatoilHydro).

Jon Gustav Vabø: 'Using the ensemble Kalman filter with kernel methods for history matching of petroleum reservoirs'.

Veileder: Jo Eidsvik og Geir Evensen (StatoilHydro).

Ranveig Niemi: 'Grafiske analysemetoder for ikke-geometriske design.'

Veileder: John Tyssedal.

Jens Morten Nilsen: 'Bayesianske metoder relevant for jernbanevedlikehold.'

Veileder: Jørn Vatn,

Fagleg ansvarleg: John Tyssedal.

Ivar Håkon Søråas: 'Kortidspredikering av tyske strømpriser.'

Veileder: John Tyssedal

Håkon Toftaker: 'Automatic parametrisations and pseudo block likelihood estimation in binary Markov random fields.'

Veileder: Håkon Tjelmeland

Eivind Grøver Aukrust: 'Fast pricing of barrier options with NIG dynamics.'

Veileder: Arvid Næss

Erik Edsberg : 'A statistical simulation-based framework for sample size considerations.'

Veileder: Mette Langaas

Maren Drange Førland : 'Rapid Seismic Inversion for Isolated Singularities.'

Veileder: Henning Omre

Even Haug : 'Methods for Extreme Value Statistics Based on Measured Time Series.'

Veileder: Arvid Næss

Håkon Kirkeby : 'In silico Investigation of Possible Checkpoint Signalling Mechanisms.'

Veileder: Henning Omre

Torbjørn Lilleheier : 'Analysis of common cause failures in complex safety instrumented systems.'

Veileder: Mette Langaas og Marvin Rausand

Marita Risberg : 'Combining the MAX test with methods for Family-wise error rate.'

Veileder: Mette Langaas

Kjartan Rimstad : 'Bayesian Seismic Lithology/Fluid Inversion Constrained by Rock Physics Depth Trends.'

Veileder: Henning Omre

Susanne Svendsen: 'Analysis of Hedging Portfolios in Turbulent Markets.'

Veileder: Jacob Laading

Torstein Tønnessen : ' Optimization Analysis of Risk/Return of Large Equity Portfolios, Applying Option Strategies.'

Veileder: Jacob Laading

Astrid Hetland Vesteraas: ' Comparison of methods and software tools for availability assessment of production systems.'

Veileder: Arvid Næss

Permisjoner

Mette Langaas og Jarle Tufto har forskningstermin høsten 2008 og vår 2009

Nyansettelser

Instituttet har tre nye stipendiater i statistikk: Håkon Toftaker, Kjartan Rimstad og Erlend Aune.

Diverse

19-20 September ble et symposium i Statistikk avholdt i Selbu. 33 stk var tilstede, derav 5 inviterte foredragsholdere: Roger Bivand (NHH), Geoff Nicholls (Oxford), Hans J Skaug (UiB), Daniel Sorensen (Danmark Agricultural University) og Geir Storvik(UiO). Samlingen var faglig interessant og samtidig en sosial begivenhet.

NYTT FRA UMB:

Master i anvendt statistikk juni 2008:

Lasse Samuelsen: *Sammenligning av metoder for analyse av path-PLS modeller.*

Veiledere: Tormod Næs og Trygve Almøy

Master i bioinformatikk juni 2008:

Katrine Damgaard: *Analyse av mikroarraydata ved bruk av blandede modeller.*

Veiledere: Guri Feten og Trygve Almøy

Nytt fra Avdeling C, Matematisk institutt, UiO

Nye PhD-kandidater

Gunnhildur Högnadóttir Steinbakk forsvarte 16. mai 2008 sin avhandling:

Conceptual and computational issues for Bayesian model Criticism

Prøveforelesning: *Bayesian model selection and model averaging*

Veiledere: **Geir Olve Storvik** og **Nils Lid Hjort**

Nye Master-kandidater:

Ingrid Schrøen Maudal: *Pensjonssystemet i endring*

Veileder: **Frank N. Proske**

Stig Korsnes: *Synthetic CDOs – An introduction, evaluation and risk estimation*

Veileder: **Xeni Dimakos**

Marie Kaas Eriksen: *Interest rate modelling with non-Gaussian Ornstein-Uhlenbeck processes*

Veileder: **Fred Espen Benth**

Gudmund Horn Hermansen: *Bayesian nonparametric modelling of covariance functions, with application to time series and spatial statistics*

Veileder: **Nils Lid Hjort**

Tekniske rapporter:

Arne Bang Huseby, Nils F. Haavardsson: *A framework for multi-reservoir production optimization.*

Pål N. Henriksen: *Pricing Barrier Options by a Regime Switching Model.*

Jørund Gåsemyr, Bent Natvig: *Extension of a Conflict Measure of Inconsistencies in Bayesian Hierarchical Models.*

Steffen Grønneberg, Nils Lid Hjort: *The copula information criterion.*

Nils F. Haavardsson, Arne Bang Huseby og Lars Holden: *A parametric class of production strategies for multi-reservoir production optimization.*

Sluttet:

Tormod Næs har sluttet som professor II etter 15 år.

Ansettelse:

Gudmund Horn Hermansen er ansatt som stipendiat fra 15. juli 2008.

Permisjoner:

Inge Helland og **Arne Bang Huseby** har forskningstermin høsten 2008 og våren 2009.

Nytt fra Avdeling for biostatistikk, UiO

Avhandlinger

Christine Louise Parr forsvarte 18. april sin avhandling *Self-reported exposure data in epidemiologic studies: measurement errors and missing values* for graden ph.d.

Tittel på prøveforelesningen var *The problem of recall bias in case-control studies of diet and disease*.

Komité: Doktor Christine M. Friedenreich, Division of Population Health and Information, Alberta Cancer Board, Canada, Professor Lauren Lissner, Sahlgrenska Akademi, Göteborgs Universitet, Forsker Nanna Lien, Universitetet i Oslo.

Veiledere har vært Marit Veierød og Petter Laake, Avd. for biostatistikk, Anette Hjartåker, Kreftregisteret, Eiliv Lund, Universitetet i Tromsø.

Joseph Andrew Sexton forsvarte 15. mai sin avhandling *Algorithmic modeling, with emphasis on measurement error and incomplete data* for graden ph.d.

Tittel på prøveforelesningen var *A survey of frequentist and Bayesian approaches to covariate measurement error in regression*.

Komité: Professor Leonard A. Stefanski, North Carolina State University, USA, Professor Geir Storvik, Universitetet i Oslo, Professor Ole Christian Lingjærde, Universitetet i Oslo.

Veileder har vært Petter Laake, Avd. for biostatistikk

Ansettelser

Christine Parr, Joe Sexton og Andrew Brown er ansatt som postdoktorer.

Diverse

Marit Veierød har forskningstermin høsten 2008 og våren 2009.

RETUR:

*Norsk Statistisk Forening
v/Hilde Grude Borgos
Schlumberger Stavanger Research
Postboks 8013
4068 Stavanger*

B

