



ISSN 0803-8953
Tilfeldig gang
Nr. 2, årgang 29
Juli 2012

Utgitt av Norsk statistisk forening

Redaksjonelt

Fra lederen	1
Fra redaksjonen	2

Leserbidrag

Jobben min som aktuar	3
Roger W. Babson – statistiker, skribent, velgjører og fantast	5
More Than Gold	9

Tankenøtter

Pusleri nummer 35	10
Løsning på pusleri nummer 34	10

Møter og konferanser

Rapport fra Second Workshop on Bayesian Inference for Latent Gaussian Models with Applications	12
Årsmøte NSF Bergen	13
29th European Meeting of Statisticians	13

Bokomtaler

Visualizing Data Patterns with Micromaps	16
--	----

Meldinger

Nytt fra Norges teknisk-naturvitenskapelige universitet	22
Nytt fra Universitetet for miljø- og biovitenskap	23
Nytt fra Universitetet i Bergen	23
Nytt fra Avdeling for biostatistikk, UiO	24
Nytt fra Avdeling C ved Matematisk institutt, UiO	24
Nytt fra Universitetet i Tromsø	25

Scandinavian Journal of Statistics

Innholdsliste juni 2012	26
-----------------------------------	----

Informasjon

Har du husket å betale kontingenten i år?	28
---	----

Fra lederen

Jo Eidsvik | Trondheim

Vårt yrke som statistiker er ett av de mest populære for tiden. En ranking gjort av *The Wall Street Journal* i fjor¹ setter statistiker på fjerde plass, etter softwareutvikler, matematiker og aktuar, foran computeranalytiker, meteorolog og biolog. Blant disse seks første er gjennomsnittslønna lavest for statistikerne. Like viktige faktorer i analysen er arbeidsmiljø, fremtidsutsikter og stressnivå. Forhåpentligvis ser dagens unge at vårt yrke kan være interessant. Det hjelper også at Googles *chief economist* kommer med utsagn som «I keep saying the sexy job in the next ten years will be statisticians»².

Arbeidet til statistikere blir også framsnakket i en artikkel i *The New York Times*³. En fortelling her omhandler butikkjeden Target, som har ansatt en statistiker. Han blir satt til å utvikle metoder kjeden kan bruke til å tiltrekke seg kundegrupper. Target sitter på store datamengder, og kan se hvilke produkter ulike grupper handler, og bruke det i markedsføring. De er ute etter barnefamilier. Ideen er å trekke mødre til Target – allerede mens de går svangre. Statistikeren utvikler en metode for å predikere om en kvinne er under «lykkelige omstendigheter», og denne blir implementert i bedriften. Hvis sannsynligheten er stor nok, får kvinnen tilbud på barneprodukter sendt til sin adresse.

Ingen tenkte så mye over det, algoritmen gjorde sine automatiske prediksjoner, og økonomisk ble den en stor suksess for Target. Men så en dag, etter at algoritmen hadde vært ett år i drift, kom en mann til Target i Minneapolis. En sur mann, som ville snakke med lederen. Hans husstand fikk tilsendt tilbud på barneklær, bleier etc. Uten grunn! Butikksjefen kunne bare beklage; han ante ikke hva som hadde skjedd. En uke etter ringte imidlertid mannen til Target-butikken. Han sa: «I had a talk with my daughter. It turns out there's been some activities in my house I haven't been completely aware of. She's due in August. I owe you an apology.»

Mange er aktive på konferanser og workshops før sommerferien: *Bayesian Inference for Latent Gaussian Models with Applications* og *Geostatistics 2012* har vært avholdt, og en del nordmenn tok turen til Nordstat i Umeå. Det er viktig at det nordiske miljøet opprettholdes. I 2014 holdes møtet i Turku, Finland. Flere norske arrangement avholdes i juli og august. Jeg håper noen tar initiativ til å lage en kort rapport for TG.

God sommer!



¹ http://online.wsj.com/public/resources/documents/st_BESTJOB050104_20110105.html

² <http://www.nytimes.com/2009/08/06/technology/06stats.html>

³ <http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>

<http://lgm2012.r-inla.org/>
<http://www.geostats2012.no/>
<http://www.nordstat2012.se/>

Fra redaksjonen

Anne Marie Fenstad | Bergen

Sommer og sommerferie står for døren, og er sikkert etterlengtet for mange, enten man «bare» skal være hjemme eller ut på tur. De fleste skal vel en lett blanding – hit og dit, innland og utland, nord og syd, men alle med håp om sol og varme, og med et rykende fersk TG i hånden!

I dette sommernummeret presenterer vi en aktuars hverdag på jobb. Vi vil igjen oppfordre medlemmene til å skrive inn og fortelle om sine forskjellige oppgaver i sitt arbeide. Det er jo kjent at en statistiker er som poteten, og kan brukes til alt. Dette vil vi gjerne høre mer om!

Som den observante statistiker sikkert vet, legges alle numrene av TG ut på nettsidene til Norsk statistisk forening. Bladet legges ut i sin helhet, i PDF-format med klikkbare linker for nettadresser, e-postadresser, artikkelreferanser og alt annet som er skrevet med farget tekst. Stilig!

<http://norskstatistiskforening.no/>

I miljøene rundt de mange nasjonale registrene snakkes det nå om å dra i gang et statistikernettverk. Dette med tanke på at man stadig vil ha de samme problemstillingene og behov for de samme statistiske metodene om man registrerer pasienter med hofteproteser, hjertesykdom eller ryggglidelser. En delegasjon fra leddproteseregisteret i Bergen var akkurat på besøk hos SKDE (Senter for klinisk dokumentasjon og evaluering) i Tromsø for å utveksle erfaringer og bli bedre kjent. Jobber du ved et register og tenker at dette er noe for deg, kan du ta kontakt med oss.

Da gjenstår det bare å ønske alle en strålende sommer. Håper også at riktig mange velger å avslutte sin ferie med konferansen *33rd Annual Conference of the International Society for Clinical Biostatistic* i Bergen 19.–24. august. En bedre start på høstsemesteret kan vanskelig tenkes!

<http://www.iscb2012.info>

Jobben min som aktuar

Jorunn Slagstad | Bergen

Eg heiter Jorunn Slagstad, og er aktuar i Norwegian Insurance Partner AS (NIP). Dette er eit leiande rådgjeving- og meklarforetak med spesialkompetanse på person-, ansvars- og skadeforsikring, og med base i Bergen. Eg har studert industriell matematikk ved NTNU i Trondheim og statistikk ved Universitetet i Bergen. I 2007 leverte eg masteroppgåva mi i finasteori og forsikringsmatematikk, og blei med det godkjent aktuar.

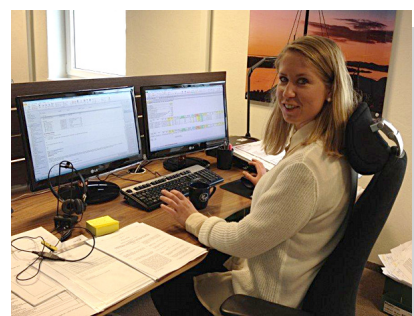
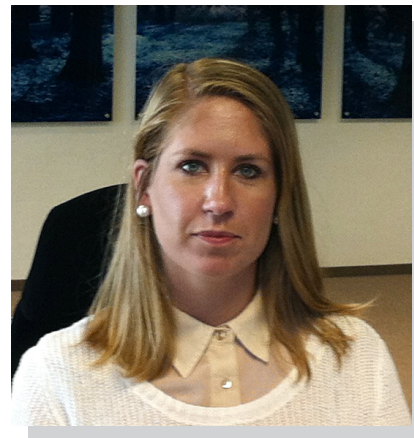
Det er tre hovudkategoriar av aktuarar – skadeaktuaren, livaktuaren og finansaktuaren. Eg er livaktuar på konsulentavdelinga i NIP, og jobbar med pensjonsforsikring. Berekning av pensjonskostandar er noko vi tilbyr kundane våre i tillegg til dei ordinære tenestene. Vi har ei fast liste over kva for selskap som ynskjer desse, men vi tek også i mot oppdrag fortløpande ved behov.

Juni månad er ikkje den mest hektiske perioden for ein livaktuar, og heller ikkje for ein livaktuar i NIP, men det er likevel nok å gjere. I dag har eg teke imot bestillingar på halvårsberekningar av pensjonskostnader, vore i dialog med livselskapet om nødvendige data, svara kundar på spørsmål om tilleggspensjonsordningar og rekna ut pensjonar for tilsette med særskilde førtidsavtalar.

Dette høyrer kanskje ikkje ut til å involvere statistikk i stor grad, så kva er samanhengen mellom pensjon og statistikk – eller rettare sagt, kva er samanhengen mellom pensjon og forsikring? I motsetning til skadeforsikring og livsforsikring, som blir utbetalt når ei uventa hending skjer, er pensjon det motsette. Den forventale levealderen i Norge er 79 år for menn og 83,5 år for kvinner⁴, så dei fleste av oss vil oppleve pensjonstilværet.

Pensjon er det vi alle skal leve av når vi ein gong trer ut av arbeidslivet, og den blir sett av i løpet av eit yrkesaktivt liv. I tillegg til folketrygda, som alle som er busette i Noreg har krav på, kjem tenestepensjon, som blir dekt av arbeidsgjeveren⁵. Det er vanleg at bedrifta kjøper ei pensjonsforsikring i eit livselskap som tek på seg ansvaret for dei framtidige utbetalingane. Selskapet betaler ein premie for medlemmene for at dei skal vere dekte av denne ordninga.

Utbetalingane frå ei ytingsordning har fleire usikre faktorar, mellom anna om medlemmene vil stå i ordninga til pensjonsalder, kva sluttløna deira vil vere og kor lenge utbetalingane kjem til å vare. Her kjem aktuaren inn i bildet. Basert på økonomiske og demografiske føresetnader, reknar aktuaren ut pensjonspliktene til



Aktuar på jobb.

⁴ Dette gjeld spedbarn fødde i 2011.

⁵ Dei to vanlegaste typene tenestepensjon i Noreg er innskots- og ytingspensjon. Mens innskotspensjon ikkje gjev nokon lovnad om framtidige utbetalingar, garanterer ytingspensjon deg ein viss prosentdel av løna di ved pensjonsalder (med frådrag for folketrygda). For føretak av ein viss storleik som har ytingspensjon for sine tilsette, er det krav til rekneskapsføring av noverdien av dei framtidige utbetalingane.

bedrifta ved slutten av året, samt den årlege pensjonskostnaden⁶.

I studietida jobba eg mykje med analyse av ulike datasett, gjennom hypotesetesting og tilpassing av regresjonsmodellar, tidsrekker og anna. Som aktuar brukar eg ferdig utarbeidde modellar, kalla *tariffar*, til å rekne ut kostnaden til eit selskap gitt opplysningar om medlemmene i pensjonsordninga. Desse tariffane blir rekna ut av aktuarar tilsette ved dei store livselskapa. I utrekning av pensjonskostnader og pensjonsplikter må ein halde seg til ferdig utarbeidde og godkjende rekneskapsstandardar. Desse er også med på å bestemme utsjånaden til rekneark vi leverer frå oss i Excel. Eg nytta programmet R i studietida, og eg tykte det var synd å ikkje kunne vidareføre denne kunnskapen i jobben min. Men det viser seg at grunnleggande kunnskap om programmering likevel ikkje er så dumt å ha med seg til Excel.

Dei aktuarberekna kostnadane skil seg frå andre kostnader i rekneskapen til bedriftene, og difor må ein som aktuar ofte forklare storleiken på desse rekneskapspostane. Gjerne er ein på telefon med finansdirektørar og revisorarar som er godt skodde i økonomiske termar, men som fortviler når dei skal forstå pensjonskostnader. Eit vanleg spørsmål er kvifor premien som blir betalt til livselskapet ikkje er den same som kostnaden i aktuarberekninga⁷, og kvifor det er endringar i pensjonskostnaden og pliktene frå året før. Det er greitt å ha repetert litt økonomi før ein skal svare på desse spørsmåla. Det viser seg at kredit og debet er komplisert når ein skal forklare til dømes estimatavvik⁸.

Mange meiner at det viktigaste ein tek med seg frå studietida er evna til å lære seg nye ting. Og dette er ei god evne å ha med seg ut i yrkeslivet. Dei færraste kjem til å jobbe med dei same problemstillingane som i studiet. Overgangen frå studiet til jobben som livaktuar krev ikkje berre læring av nye datasystem og ny programvare, men også ei omfattande opplæring i det breie fagfeltet pensjon er. Det er også eit fagfelt i endring, mellom anna grunna rørsler i det økonomiske biletet i Europa. Desse endringane påverkar også norske pensjonsordningar. Stikkord er lågrentescenarier, auka levealder og strengare krav til eigenkapital. Dei to førstnemnde er med på å auke pensjonspliktene i ytingsordningar der arbeidsgjevar ber risikoen. Dette har ført til at mange bedrifter avviklar ytingsordningane sine, og går over til innskotspensjon, der arbeidstakar ber risikoen. Då treng ikkje selskapa lenger balanseføre noverdi av framtidige utbetalningar, sidan desse er fullt ut fastsette. Dette gjer at dei heller ikkje har behov for aktuarberekningar.

Men dette vil ikkje seia at det ikkje lenger er behov for aktuaren! Aktuarutdanninga gjev deg ei unik stilling i jobbmarknaden. I tillegg til ein sterk teoretisk bakgrunn, har du ei yrkesretta utdanning. Bankar og forsikringsselskap vil alltid ha behov for prissetjing av produkta sine, og i ein konkurranseutsett marknad ynskjer ein å tilby eit levedyktig produkt til ein pris som kundane vil kjøpe for.

⁶ Brutto pensjonskostnad er den periodiserte delen av pensjonspliktene, og er pensjonsopptening det året, samt rentekostnad på pensjonspliktene.



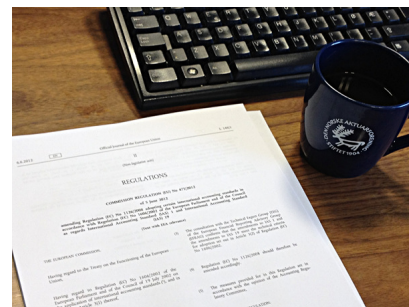
Aktuar i samtale med leiaren sin, Roy-Arild Hansen Dybvik, i Norwegian Insurance Partner sine lokale i Christian Michelsens gate 6B.



Aktuar i telefon med kundar.

⁷ Premien selskapet må betale til livselskapet vert rekna ut på gjeldande medlemmer og lønningar, medan selskapet har gitt eit framtidig løfte om utbetaling på sluttløn. Pliktene selskapet har er dermed større enn den premien dei betaler inn.

⁸ Estimatavvik er differansen mellom inngåande og utgåande balanse – balanse som ikkje kan forklarast av dei andre postane i rekneskapen, og kjem av at det har vore endringar i medlemstal og føresetnader som ein ikkje hadde kjennskap til når ein gjorde berekninga.



Forordningar og aktuarkopp – sistnemnte frå Den Norske Aktuarforening.

Jostein Lillestøl | Bergen

The future of this country depends on the preacher
and the statistician.

Etter praksis som veiingeniør skjønnte han at det ikke var noe for han, og ville finne på noe annet. Faren mente at han burde gå inn i en bransje der hans kunnskaper kunne ha gyldighet livet ut. Sønnen fulgte rådet, og begynte i et investeringsfirma i Boston, der han fort lærte det meste om de finansielle markedene. I kundens beste interesse stilte han spørsmål ved selskapets metoder (og ublue honorarer), og fikk sparken etter kort tid. Han startet umiddelbart sin egen virksomhet med salg av obligasjoner, og få år senere, i 1904, en statistisk organisasjon med formål å samle inn og formidle tilstandsinformasjon om næringsliv, markeder og investeringer. Hans analyser ble solgt som nyhetsbrevet *Babson's Report* til investeringsselskaper og private investorer. Dette var noe helt nytt, og ble en stor suksess, slik at han raskt la seg opp en betydelig formue. Etter et dramatisk børskrakk i 1907 utvidet han virksomheten til rådgivning om kjøp og salg av verdipapirer, hva man burde kjøpe eller selge, og når.

A black and white portrait of a middle-aged man with a full beard and mustache. He is wearing a dark suit jacket over a light-colored shirt and a dark tie. He is looking slightly to the right of the camera with a serious expression. The background is a plain, light-colored wall.

This Supplement is based upon an analysis (worked out by location and industry) of the fundamental statistics of Dunbar's Bureau of Retail Supplement: check transactions, bank deposits, bank deposits, money rates, commodity prices, building permits, contracts awarded, crops, exports and imports, retail and wholesale trade, employment, production, oil loadings, newspaper advertising, patent receipts, services, electric power production, and other statistical data. In selling, advertising

BABSON'S REPORTS
BABSON PARK, MASS. OCTOBER 1924

BUSINESS SUPPLEMENT

Sales Opportunities in Localities and Industries

Sales Still Improve Moderately

Activity More Evenly Distributed Thruout the Country

Sales Continue to Pick Up Gradually. The tentacles need a few months ago are developing about as expected. No wild upswing in sales is reported but conditions have definitely improved in a conservative way. No reason has thus far appeared for changing the forecast that sales should be fairly satisfactory for at least the fall and early winter. Unless you are in a line which you know always lags behind business as a whole, you can fairly ask your sales department to turn in a relatively large amount of orders than in the earlier part of the year. In other words, when allowance is made for seasonal movements, your sales curve should show an upward bend — tho the slope may not be extreme; store.

Conditions Coming Into Better Balance. In a broad way what is happening is a leveling up of conditions throughout the country. In terms of the sales map there is comparatively little of the gold color denoting expansion, but the area covered by blue color, which denotes depression, is lessening. The Map is beginning to become mainly green, the color which is used to signify fair tendencies. Hereafter the regions least favorable, situated were the farm sections. With the advance in prices of farm products the farmers' status has risen until in many cases it is now approaching a level with that of the manufacturer.

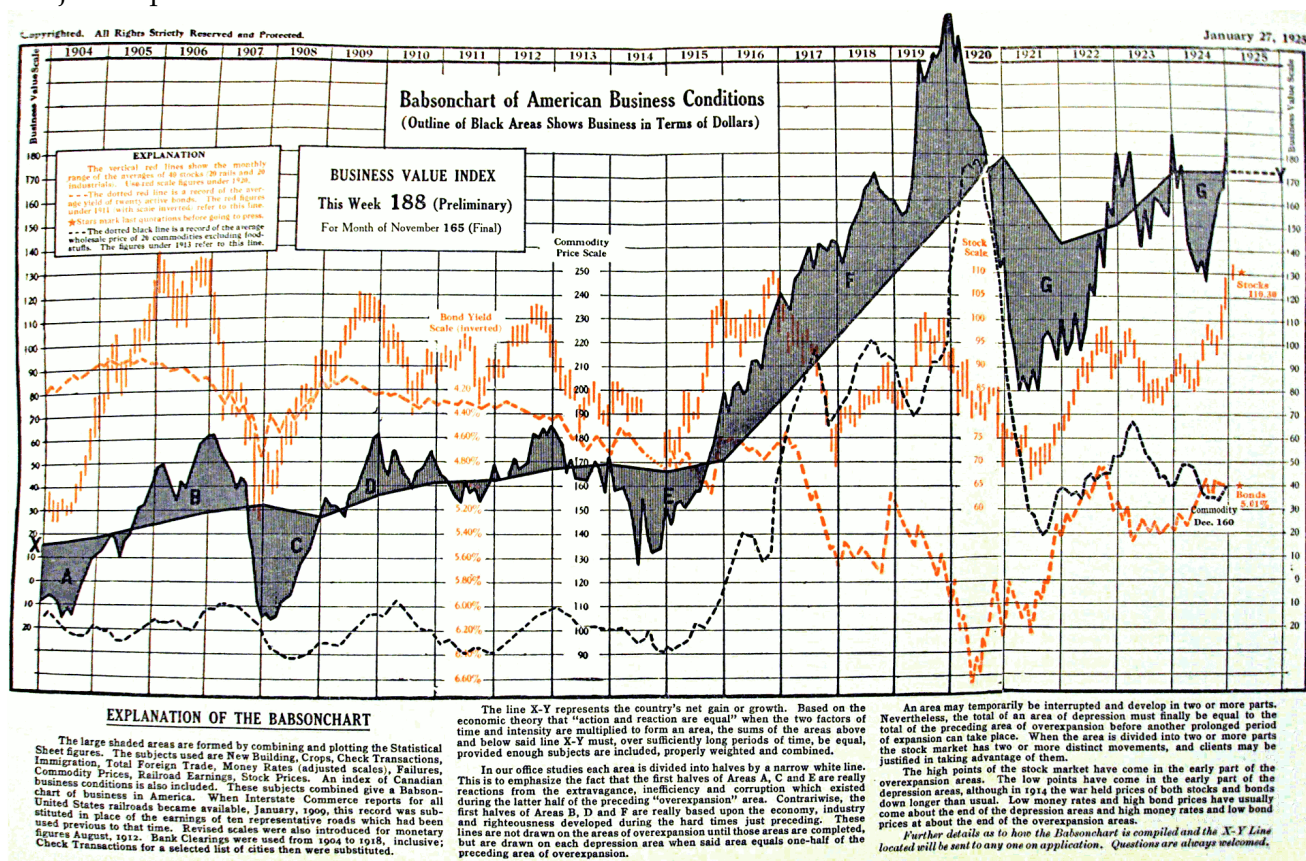
When dire agricultural areas reach the same economic plane as the industrial centers of the world, the cities will be first to take advantage of the new advance into higher ground. There is some precedent for anticipating that the cities will take the lead. This is not a foregone conclusion, however, and the relative standing of the two types of territory should be constantly watched. As to the immediate future, we favor special campaigns (advertising, selling, collecting) in selected agricultural areas on the principle that territories which have swung upward from depression offers opportunities for new capitalists.

FEATURES IN THIS SUPPLEMENT		Page
Crop Forecasts by States		1
How Should Farm Collections be Handled		2
Best Industries for Selling		3
California Somewhat Inactive		4-5
Analysis of Trends in Each State		

Care Needed in Handling Farm Collections. It has been suggested that local bankers in the farming regions, who have been very heavily burdened the last few years, should continue to handle the situation from the broad standpoint of community interests. For example, suppose that the bank presses the farmer for payment in full; in some cases this may mean that after paying the bank, the farmer will have difficulty in paying the local merchant, who likewise owes the bank. Might it not be better policy for the bank to ask the farmer to

Territories Department Confidential Report, No. M32
Copyright 1944 by BARNHART STATISTICAL ORGANIZATION, Incorporated

kvinner. Over tid ble en rekke statistiske publikasjoner lansert, med registreringer av likt og ulikt, som han tjente store penger på. Publikasjonene hadde også iøynefallende grafikk. Et eksempel er et kart over statene i USA, spekket med informasjon om hvor det kunne være lurt å satse eller ikke, og hva. Det fantes i to varianter, ett for de optimistiske og ett for de pessimistiske. Sammen med ingeniørprofessor Swain fra MIT utviklet han det såkalte *Babsonchart* for økonomiske indikatorer, som tok sikte på å evaluere nåværende og fremtidige konjunkturer, der Newtons tredje lov spilte en viss rolle.



Babsonchart var basert på en sammenfattende indeks konstruert på grunnlag av en rekke tidsserier av ulik karakter – kommersielle, finansielle og monetære. Hans modell tok sikte på å beskrive en 20-årsperiode med sykluser i fire stadier: overekspansjon, nedgang, depresjon og forbedring. En linje gikk gjennom diagrammet, og representerte den normale vekst, og det var her Babson brakte inn Newtons tredje lov om kraft og motkraft, uttrykt ved at arealer over og under linjen burde være like. Diagrammet hadde også flere andre egenskaper som vi ikke skal komme inn på her. Disse diagrammene oppnådde stor popularitet i praksis. På sett og vis er dette et tidlig eksempel på det som senere er kalt «teknisk analyse». Dette ble selvsagt sett på med skjeve øyne fra den tids toneangivende og respekterte økonomer på Harvard og Yale, herunder spesielt Irving Fisher.

Hans virksomhet fikk et ekstra løft da han i 1929 var en av de få som spådde at et nytt børskrakk var nært forestående, og

økonomene tilsynelatende var tatt med buksene nede. Senere har den kjente økonomen J.K. Galbraith uttalt seg om dette:

The methods by which he reached his conclusions were a problem. They involved a hocus pocus of lines and areas on a chart. Intuition and even mysticism played a part. Those who employed rational, objective and scientific methods failed to foretell the crash. In these matters, as so often in our culture, it is far, far better to be wrong in a respectable way than to be right for the wrong reasons. Wall Street was not at a loss as what to do about Babson. It promptly and soundly denounced him.

Selv om han hadde noen spredte artikler i seriøse vitenskapelige journaler, ble hans forsøk på faglighet oftest karakterisert som kvasivitenskap på linje med astrologi. På den annen side kan man si at han bidro til større bevissthet om verdien av statistiske opplysninger i praksis, i «forecasting» som målsetting, og bidro kanskje også til at vitenskapelige økonomer gikk dypere inn i modellering av konjunkturer.

Roger Babson gjorde seg også til uvenn med politikere ved å anklage dem for ikke å ha gjort noe for å forhindre krisen. Han sier i sin selvbiografi at de store tapene til selv nøysomme investorer ved krasj hadde ledet han til å arbeide for at slikt ikke skulle skje igjen. Han så det som sin plikt å være oppdrager og dele sine kunnskaper med offentligheten, og han var fast spaltist i flere av de store avisene, blant annet *The New York Times*. Han var også ofte å finne i populære magasiner, herunder *Popular Science*, der han hyppig ble brukt som kilde med merkelappen «the eminent statistician». I løpet av 33 år skrev han 47 bøker, som dekket et bredt spekter av temaer – næringsliv, industri, politikk, helse, sosiale forhold, reiser og religion. En rød tråd i det meste var å stimulere individer og samfunnet til å gjøre endringer til det bedre.

Babson sluttet seg ikke til de nyrike, men investerte sin formue i mange tiltak til beste for samfunnet. Han støttet blant annet industri som kunne øke folks velferd, for eksempel ved å utvikle produkter til beste for folks helse og sikkerhet, som hygieneprodukter, alarmsystemer, brannslukkere og trafikkregulering. I mange av sine prosjekt greide han (på mystisk vis) å integrere Newtons lover, og etter hvert med søkelyset på forhold knyttet til gravitasjon. Han ble så opptatt av dette at han opprettet en egen stiftelse, *Gravity Research Foundation*, med formål å støtte forskning på gravitasjon og dens, etter hans oppfatning, negative konsekvenser. I flere år hadde han tre personer fast stasjonert ved *U.S. Patent Office* for tidlig å snappe opp patenter på området. Han ga også penger til flere forskningsinstitusjoner, som knapt visste hva de skulle gjøre med pengene før etter hans død. *Gravity Research Foundation* eksisterer den dag i dag, men er

redusert til å drifte en årlig essaykonkurranse om gravitasjon, der blant annet Stephen Hawking er blitt belønnet flere ganger.

Roger Babson grunnla tre colleges, og la det finansielle grunnlaget for disse. Det meste kjente er *Babson College*, i dag en respektert akademisk institusjon, som han grunnla allerede i 1919. Hans hovedtanke var å kunne tilby et intensivt program på ett år med fire komponenter: praktisk økonomi, finansiell styring, forretningspsykologi og personlig effektivitet. Studiet var en kombinasjon av undervisning og utplassering i bedrifter, og de fleste ansatte lærere hadde erfaring fra næringslivet. Studentene arbeidet med gruppeprosjekter og presentasjoner, observerte produksjonsprosesser ved bedrifter i nærheten (altså statistikk i praksis), og hadde møter med ledelsen. Dette sto i sterk kontrast til de eksisterende 4-årige collegeprogrammene. Dette innebar ikke at vitenskapene, kunst og humaniora ble nedvurdert, bare at dette fikk man bedre andre steder.

Ikke alle hans forretningsprosjekter og sosiale prosjekter var like vellykkede, men han tok tapene med fatning, og gikk ukuelig i gang med nye prosjekter. Han var sterkt religiøs, og forsøkte gjennom organisasjonsarbeid på nasjonalt nivå å høyne moralen innen kirken. I 1940 stilte han som presidentkandidat for *Prohibition Party*. Han fikk ikke mange stemmene, noe han heller ikke ventet, men han fikk mange mennesker i tale, og ble likevel nummer tre blant ni kandidater, så det var ikke verst.

Under depresjonen etablerte han en stor naturpark, og engasjerte arbeidsløse steinhoggere i sin hjemstat til å hogge ut leveregler på steinstøtter som han plasserte ut hist og pist i parken, til glede for de vandrende: «Keep out of debt», «Help mother», og så videre. Disse kan visstnok oppleves den dag i dag, og må vel sies å være blant de vellykkede prosjekter.

En fargerik velmenende statistisk predikant og spåmann, som neppe spådde rett i 1912!



Foto: Corey Seeman – © ⓘ ⓘ ⓘ

Friedman W.A. (2007). «The Seer of Wellesley Hills: Roger Babson and the Babson Statistical Organization». *Harvard Business School Working Paper* 08-036.



Foto: David Bruce – © ⓘ ⓘ ⓘ

More Than Gold

Lisa Miller | Bergen

Working with More Than Gold, a Christian charity, I have been providing statistical support for their *Athlete Family Homestay* programme, which has consisted of a weighting system for ranking host applicants and then matching hosts with the athletes' guests.

More Than Gold has been involved in the Summer Olympic Games since Atlanta in 1996, as well as the Winter Olympic Games, including Salt Lake City and Vancouver. The *Athlete Family Homestay* programme, which has been awarded the Inspire mark by London Organising Committee of the Olympic and Paralympic Games (LOCOG), offers free bed and breakfast style accommodation to the athletes' 'support family' during the 2012 Olympic and Paralympic Games. The support family may consist of family members, team officials, personal coach or friends critical to the athlete's support.

Host applicants are ranked based on a weighting system that consists of ease of travel on public transport, time to travel to the nearest Olympic venue, cost of travel, number of rooms available, dates available to host and extra benefits, such as ability to provide Internet access, towels or additional meals. Travel information requires a mapping system of each area code to the closest venue, as well as to the main Olympic stadium.

Guests are matched to a host applicant based on travel conditions to the athletes' Olympic venue and the hosts availability during the Olympic Games. Guests are also matched to hosts using lifestyle preferences, such as smoking, personality, religion and sporting interests. This is performed on a one-to-many match, with each guest being matched to multiple hosts and the host with the highest match being selected.

Guest arrival and departure information also needs to be mapped to each airport and international train station, and monitored closely during the Olympic Games to ensure the host is at the correct venue and at the correct time to collect their guests on their arrival into London.

More detailed information on More Than Gold and the other programmes that they are involved in can be found at www.morethangold.org.uk.

Pusleri nummer 35

Jostein Lillestøl | Bergen

La X_1 og X_2 være uavhengige $\mathcal{N}(0,1)$, og la

$$Y_1 = \frac{X_1}{\sqrt{X_1^2 + X_2^2}} \quad \text{og} \quad Y_2 = \frac{X_2}{\sqrt{X_1^2 + X_2^2}}.$$

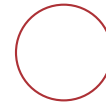
Da er opplagt (Y_1, Y_2) uniformt fordelt på sirkelen $Y_1^2 + Y_2^2 = 1$.

La isteden

$$Y_1 = \frac{X_1}{\sqrt{X_1^2 + X_2^2 + Z}} \quad \text{og} \quad Y_2 = \frac{X_2}{\sqrt{X_1^2 + X_2^2 + Z}},$$

der Z er ikke-negativ stokastisk variabel, slik at (Y_1, Y_2) er punkt på sirkelskiven $Y_1^2 + Y_2^2 \leq 1$. Hvilken sannsynlighetsfordeling må Z ha for at (Y_1, Y_2) blir uniformt fordelt på sirkelskiven? Hint: Med Arkimedes på laget trenger du ikke papir og blyant.

Send svar til Jostein.Lillestol@nhh.no.



Løsning på pusleri nummer 34

Jostein Lillestøl | Bergen

Oppgaven

I USA gjennomførte man i perioden 1999–2008 *The 50 States Quarters Program*. Kort fortalt ble store mengder quartermynter sendt ut på markedet med motiver fra hver av de 50 ulike statene. Per hadde forskningsopphold i USA på denne tiden, og skrev blant annet følgende i sin rapport hjem:

Min datter Jenny prøver å få en fullstendig *States Quarters*-samling, og hun har 47 ulike nå. Når vi får en slik quarter ser vi først etter om den passer i hennes samling, hvis ikke så legger jeg den til min egen samling, som har 34 forskjellige akkurat nå. Før vi begynte å samle, spurte jeg meg selv:

Hva er forventet antall slike mynter som trengs før Jenny har en fullstendig samling? Hva er forventet antall slike mynter som trengs før jeg også har en fullstendig samling?

I det siste har det gått lenge mellom hver gang vi finner en ny mynt, og jeg spør meg: Hva er forventet antall flere mynter som trengs for Jenny og meg akkurat nå, når Jenny har 47 ulike mynter og jeg 34?

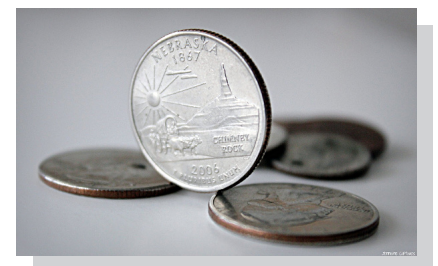


Foto: Jennuine Captures

Løsning

Ventetiden til Jenny får en fullstendig samling må være lik summen en rekke geometrisk fordelte ventetider med avtagende suksessannsynligheter når en mynt man ikke hadde før mottas, og forventningen blir

$$\frac{50}{50} + \frac{50}{49} + \dots + \frac{50}{1} = 50\left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{50}\right) = 224,96.$$

Forventet ventetid til Per også har sin samling komplett kan løses slik: La m være antall ulike mynter Jenny har alene og n være antall ulike mynter Per har (og derfor begge har). Vi har da følgende overgangssannsynligheter.

$$(m, n) \rightarrow \begin{cases} (m, n), & n/50, \\ (m+1, n), & 1-m/50, \\ (m, n+1), & (m-n)/50. \end{cases}$$

La $f(m, n)$ være forventet antall flere mynter som trengs for full samling ved start i (m, n) . Da har vi

$$f(m, n) = 1 + \frac{n}{50}f(m, n) + \left(1 - \frac{m}{50}\right)f(m+1, n) + \frac{m-n}{50}f(m, n+1),$$

som gir

$$f(m, n) = \left(1 - \frac{n}{50}\right)^{-1} \left[1 + \left(1 - \frac{m}{50}\right) \times \right. \\ \left. \times f(m+1, n) + \frac{m-n}{50}f(m, n+1) \right].$$

Vi ønsker $f(0,0)$, og har randkravene $f(m, 50) = 0$ og $f(50, n) = 50\left(1 + \frac{1}{2} + \dots + \frac{1}{50-n}\right)$. Beregninger basert på formelen ovenfor er enkel ved bruk av regneark og relativ adressering: La linjer og kolonner være henholdsvis (m, n) i omvendt rekkefølge fra 50 til 0. Etabler først marginalene i første kolonne og første linje i tabellen i henhold til randkravene ovenfor. Skriv formelen med relativ adressering i neste sør-østlige rute. Kopier formelen til hele trekanten. Da leser vi tallet $f(0,0) = 324,80$ og $f(47, 34) = 197,08$, som besvarer de to siste spørsmålene.

Kommentar: I «gamle dager» ville en trolig forsøkt å løse den bivariate differensligningen ved å innføre genererende funksjoner, og endt opp med en partiell differensialligning som trolig krever en numerisk løsning. Regneark som Excel muliggjør løsninger på mange problemstillinger uten å gå veien om en sluttet formel med tilhørende teknikaliteter. Dette bør ha en viss konsekvens for våre undervisningsopplegg, og reiser samtidig spørsmålene: Hva er en løsning? Når er et problem løst?

	A	B	C	D	E	F	G	H	I	J
1										
2	50	0	50	75	91,66667	104,1667	114,1667	122,5	129,6429	135,8929
3	49	0	100	112,5	122,2222	130,2083	137	142,9167	148,1633	152,8795
4	48	0		137,5	143,3615	149,5949	154,5649	159,0102	163,0539	166,7603
5	47	0			160,6481	164,8582	168,6775	172,1772	175,4101	178,4164
6	46	0				177,3582	180,4136	183,256	185,9154	188,4159
7	45	0					190,4136	192,7822	195,0202	197,1425
8	44	0						201,1156	203,0339	204,8654
9	43	0							210,1767	211,7793
10	42	0								218,0293

318,7373	318,9154	319,0927	319,2693	319,4451	319,6202
	320,0023	320,1759	320,3487	320,5208	320,6922
		321,2397	321,4089	321,5775	321,7453
			322,4506	322,6157	322,7801
				323,6361	323,7973
					324,7973

møter og konferanser

Rapport fra Second Workshop on Bayesian Inference for Latent Gaussian Models with Applications

Organisasjonskomiteen

Den andre workshop om latente gaussiske modeller ble avholdt på NTNU i Trondheim denne våren. Forrige gang var i Zürich i 2011. Litt over 100 personer deltok, fra cirka 20 land.

<http://lgm2012.r-inla.org/>

Programmet bestod av innsendte foredrag og en posterdel. I tillegg hadde vi seks inviterte foredrag: David Dunson (Duke University, North Carolina), Gregor Gorjanc (University of Ljubljana, Solvenia), Janine Illian (University of St Andrews, Skotland), Doug Nychka (National Center for Atmospheric Research, Colorado), Chris Williams (University of Edinburgh, Skotland), Yu Ryan Yue (Barach University, New York). Før det offisielle programmet startet arrangerte vi et kurs i *INLA for beginners* og et kurs i *Markovian models in spatial statistics*.

I tillegg til alt programmet bød på av modeller, metoder og datasett, hadde vi en trivelig byvandring i Trondheim og en festlig konferansemiddag på Rica Nidelven hotell.

Neste år arrangeres den tredje workshop om latente gaussiske modeller på Island!



Årsmøte NSF Bergen

Ågot Irgens | Referent

Tidspunkt 28.3.2012 kl. 17.00

Sted Nasjonalt register for leddproteser

Oppmøte 9 personer

- 1 Innkallingen ble godkjent.
- 2 Årsberetningen for 2010–2011 ble lagt frem, og godkjent med små justeringer. I 2011 ble det avholdt et møte der Bo Henry Lindqvist snakket om *Competing risk – fascinerende teori og viktige anvendelser*.
- 3 Regnskapet for 2011 ble godkjent med minimale endringer. Siden aktiviteten i 2011 var liten, var det et lite overskudd etter overføring fra moderforeningen.
- 4 Valg. De som sitter i styret har sittet der i flere perioder og enkelte ønsket avløsning. Nye kandidater ble foreslått og valgt med akklamasjon:

Styreleder	Anne Marie Fenstad
Sekretær	Bård Støve
Kasserer	Øystein Haaland
Styremedlem	Karl Ove Hufthammer
Styremedlem	Ågot Irgens
Styremedlem	Astrid Lunde (om hun ønsker å fortsette)
- 5 Eventuelt: Ingenting til eventuelt, men det ble fra flere medlemmer påpekt at TG, innhold og utseende, var svært bra.

Stein Atle Lie, forskningsleder ved Uni Helse, holdt et interessant foredrag med tittelen *Multi-state models in the analysis of sickness absence data*. Vi avsluttet som vanlig møtet med pizza.

29th European Meeting of Statisticians

Budapest, Hungary, 20-25 July 2013
at the Eötvös Loránd University

The Bernoulli Society, the community of Hungarian statisticians and probabilists and the organisers cordially invite their colleagues from all over the world to participate in the broadest and most prestigious regular meeting of the profession in Europe.

CALL FOR PAPERS

The Programme Committee is appointed by the European Regional Committee of the Bernoulli Society. The worldwide character of the conference is emphasised by the inclusion of members from overseas.

CHAIR: Michael Sørensen, Univ. of Copenhagen, Denmark

MEMBERS:

Victor Perez Abreu	CIMAT, Mexico
Felix Abramovich	Tel Aviv Univ., Tel Aviv, Israel
Alexander Bulinskii	Lomonosov Moscow State Univ., Moscow, Russia
Krzysztof Burdzy	Univ. of Washington, Seattle, USA
Probal Chaudhuri	Indian Statistical Institute, Kolkata, India
Rama Cont	Centre National de Recherche Scientifique, Paris, France
Kai Tai Fang	Hong Kong Baptist and Beijing Normal Universities, Hong Kong
Enno Mammen	Univ. of Mannheim, Mannheim, Germany
László Márkus	Eötvös Loránd Univ., Budapest, Hungary
Axel Munk	Univ. of Goettingen, Goettingen, Germany
Benedikt Pötscher	Univ. of Vienna, Vienna, Austria
Nancy Reid	Univ. of Toronto, Canada
Marco Reis	Univ. of Coimbra, Coimbra, Portugal
Gareth O. Roberts	Univ. of Warwick, Coventry, U.K.
Fabrizio Ruggeri	CNR-IMATI, Milan, Italy
Mark J. van der Laan	Univ. of California, Berkeley, USA
Rasmus Waagepetersen	Aalborg Univ., Aalborg, Denmark
Stephen Walker	Univ. of Kent, Canterbury, U.K.

The Local Organising Committee

CHAIR: László Márkus, Eötvös Loránd Univ., Budapest, Hungary

MEMBERS:

Miklós Arató	Eötvös Loránd Univ., Budapest,
Ágnes Backhausz	Eötvös Loránd Univ., Budapest
Sándor Baran	Univ. of Debrecen,
Villő Csizsár	Eötvös Loránd Univ., Budapest
István Fazekas	Univ. of Debrecen,
Tamás Móri	Eötvös Loránd Univ., Budapest
Gyula Pap	Univ. of Szeged
Vilmos Prokaj	Eötvös Loránd Univ., Budapest
Tamás Rudas	Eötvös Loránd Univ., Budapest
András Zempléni	Eötvös Loránd Univ., Budapest

Invited topics in statistics:

- Shape and image analysis
- Statistical Methods in Genetics, Bioinformatics and Systems Biology
- Machine Learning
- Financial time series/Extreme events
- High-Dimensional Statistical Inference
- Function Estimation/Nonparametric Statistics
- Spatial and space-time statistics
- Bayesian nonparametrics
- Model Selection
- Forensic statistics
- Statistics for SDEs, including jump diffusions (Levy-driven SDEs) and high-frequency data
- Bayesian computing
- Environmental statistics/Environmental epidemiology
- Statistical Models and Analysis of Neuronal data
- Functional Time Series
- Empirical processes
- New developments in causal inference
- Statistics for single molecule experiments

Invited topics in probability:

- Limit theorems and applications
- Random fields and geometry
- Stochastics in Biology
- Random matrices
- Stochastic Partial Differential Equations including fractional stochastic partial differential equations
- Spatial point processes
- Stochastics in Finance
- Stochastic modelling and risk analysis
- Branching processes and applications
- Percolation and random graphs
- Rough paths

Organised Contributed Paper Sessions

To promote the presentation of new ideas and innovative topics in the conference, proposals are welcomed for Organised Contributed Paper Sessions (Streams), with a designated organiser, consisting of 4 papers/stream (of 20 minutes each).

If you wish to organise a stream, please, send your proposals to ems2013@congress.hu till 28 February 2013 the latest.

On acceptance, the Contributed Paper Session organiser will be in charge for the session and will chair it. The speakers of the group will need to upload their abstracts complying with the general rules but clearly indicating the title of their Organised Contributed Paper Session when submitting.

The abstracts can be submitted through the conference website. www.ems2013.eu

Authors have to make an abstract registration before uploading their paper in LATEX format.

The abstract submission starts on 1 September and the final deadline for submission is 15 March 2013.

The Programme Committee evaluates the abstracts and will notify the authors about the acceptance of their paper for presentation until 15 April 2013 the latest.

For detailed instructions for authors, please contact the website in August 2012.

REGISTRATION

	Early before 1st May 2013	Late after 1st May 2013
Normal registration fee	300 €	390 €
Bernoulli Society / EMS members registration fee*	240 €	350 €
Student registration fee	160 €	240 €

*In accordance with the agreement between BS and EMS

According to tax regulations in force, participants possessing European VAT number or a certificate that they are subject to tax in their own country – are not charged with VAT. For all other participants, the gross fee including 27% VAT will be charged on top of the amounts given above.

The registration fee includes:

- Admission to the scientific programmes
- Lunch and Coffee Breaks during the conference
- Conference kit
- Admission to the Welcome Reception and Organ Concert

For the conference banquet tickets can be purchased in advance.



All professional activities of the Conference will take place in Eötvös Loránd University (ELTE), Budapest.

ELTE Campus is situated in a picturesque environment on the right bank of the Danube, in a walking distance from the city center.

MEETING AND SOCIAL PROGRAMMES

The scientific programme consists of the traditional Forum, Opening and Closing plenary lectures, 6 special invited lectures, 28 invited and several contributed papers and poster sessions. To stimulate the interaction between probability and statistics a balanced representation of intertwined topics of both disciplines is intended. Special commemorative lectures celebrating the anniversaries of the profession in 2013 will be held. The Future of Statistics will be discussed by a designated panel in interaction with the audience. Ample time will be given for discussion forums on issues of publications and European research funding.

Social programmes included in the registration fee:

Welcome Reception, 21 July Sunday ELTE Congress Venue	Organ Concert, 22 July Monday St. Stephen Basilique
The Welcome Reception will be held at the ELTE University Congress Center, the meeting venue.	The organ concert organized exclusively for participants of the EMS 2013 will surely remain an unforgettable experience for all.

Conference Banquet on Wednesday 24 July
Dining out on River Danube, the participants will enjoy magnificent views of historical Budapest from the boat.

USEFUL INFORMATION

For travelling and accommodation advises and also for general information about Hungary and Budapest, please visit our website: www.ems2013.org

EMS 2013

www.ems2013.eu

FURTHER INFORMATION

Eszter Zsigmond
Congress Ltd.,
EMS 2013 Secretariat
H-1026 Budapest
Szilágyi Erzsébet Fásor 79.
Hungary

Tel: +36 1 212 00 56
Fax: +36 1 356 65 81
Ems2013@congress.hu
Further Information:
www.ems2013.eu



Visualizing Data Patterns with Micromaps

Karl Ove Hufthammer | Bergen

Geografiske data

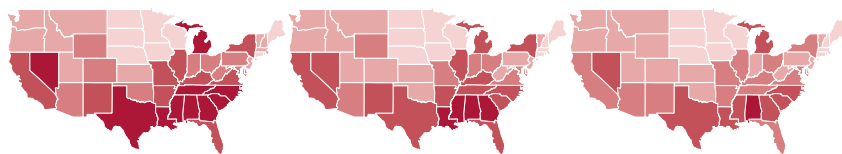
Den geografiske komponenten i statistiske data er ofte interessant; tenk berre på Snows berømte analyse av kolerautbrotet i London på 1800-talet⁹. Og har me datapunkt knytte til geografiske posisjonar, kan me lett teikna desse direkte opp på eit kart¹⁰.

Men ofte har me data aggregerte på administrative einingar, som fylke, kommune eller land. Og sjølv om dataa i utgangspunktet *ikkje* er aggregerte, kan me i ein del tilfelle med fordel aggregere dei, eller bruka ei anna form for geografisk glatting, for klarare å få fram eventuelle mønster¹¹. Anten det er snakk om krefttratar, sjølvmondsratar eller arbeidsløyseprosent kan geografiske mønster vera utgangspunkt for både forskning og debatt.

Den vanlige måten å presentera slike data grafisk på har vore såkalla *koropletkart*, der kvart område (for eksempel fylke) vert fargelagt med ein farge som indikerer verdien på den aktuelle variabelen (for eksempel brystkrefttraten), gjerne mørk farge for høge verdier og lys farge for låge verdier. Men det er egentlig litt pussig at denne typen kart har blitt så populær, då det er ein svært *dårlig* måte å framstilla slike geografiske data på.

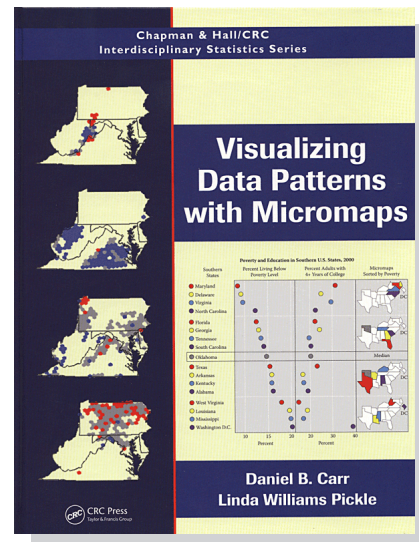
Det er fleire grunnar til at vanlige koropletkart er dårlige / lite formålstenlige. For eksempel kan eitt og same datagrunnlag gje opphav til kart som gjev svært forskjellige inntrykk av verdiane og mønsteret i dataa, avhengig av talet på fargeklassar/intervall, på grensene mellom intervalla og på val av fargar. I tillegg vil geografiske store område gjerne dominera bildet, og nærast «skjula» interessante verdier og variasjon i andre område.

Her er eit eksempel på kor forskjellige kart basert på same data – her mordrata i amerikanske delstatar – kan sjå ut. Eg har brukt populære metodar for val av intervallgrenser. Det finst for øvrig langt meir ekstreme eksempel, spesielt viss me begynner å endra talet på intervall/fargeklassar¹².



Linked micromaps

Det har vore overraskande få forslag til andre og betre måtar å framstilla data som dei ovanfor på. Eitt eksempel, som eg har stor sans for, er **Cleveland og McGill (1984)** sitt om «termometer»-symbol direkte på karta. Artikkelen var for øvrig nybrottsarbeid når det kom til faktisk eksperiment for å avgjera



⁹ Prøv eit søk etter «Snow cholera» i din favorittsøkemotor om du ikkje kjenner historia.

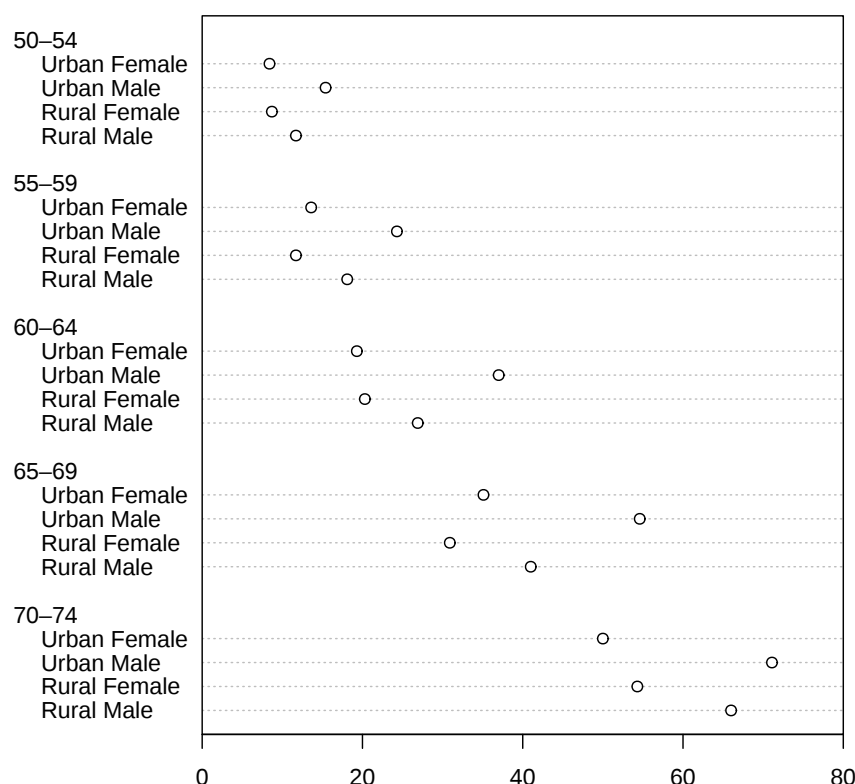
¹⁰ Og kartet kan gjerne vera interaktivt. Sjå for eksempel den flotte framstillinga av resultatane av dei nasjonale prøvane dei siste åra på <http://skoleporten.bengler.no/>.

¹¹ Det er i praksis det same me gjer når me for eksempel teiknar inn regresjonslinjer i eit spreingsplott; me fjernar støyen for å konsentrera oss om mønsteret/signalet/trenden.

¹² For nokre fine eksempel på korleis val av fargar og intervallgrenser kan påverka koropletkart, sjå artikkelen «Take Care of your Choropleth Maps» på <http://vis4.net/blog/posts/choropleth-maps/>. Eit anna godt eksempel er det interaktive kartet på <http://kartograph.org/showcase/choropleth/>. (Vel «Population», og prøv så dei tre måtane å velja intervallgrenser på. Legg merke til kor svært forskjellige inntrykk dei gjev!)

kva graftyplar som fungerer og kva graftyplar som ikkje gjer det¹³, og artikkelen er anbefalt lesnad for alle som brukar grafiske framstillingar av data i arbeidet sitt. Eit anna forslag kjem Carr og Pickle med i denne boka – *linked micromaps* (kanskje *kopla mikrokart* på norsk?). Desse grafiske framstillingane koplar vanlige radorienterte plott med små kart. Eksempel på radorienterte plott er Cleveland's prikkplott/punktplott (*dotplots*)¹⁴, som ofte er ein svært god måte å framstilla data på:

Dødsratar (per tusen) i Virginia 1940



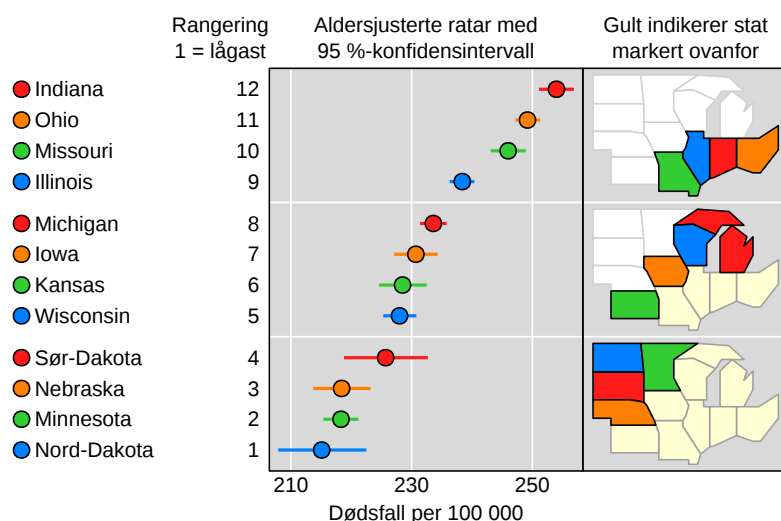
Cleveland W.S. og McGill R. (1984). «Graphical perception: Theory, experimentation, and application to the development of graphical methods». *Journal of the American Statistical Association*, **volum 79**, nummer 387, side 531-554. ISSN 01621459. [10.2307/2288400. http://www.jstor.org/stable/2288400](http://www.jstor.org/stable/2288400)

¹³ I artikkelen ser me for øvrig den første empiriske grunngevinga for at kakediagram er ein ekstremt dårlig framstillingsmåte for statistiske data. For ein nyare og grundig gjennomgang av problema med kakediagram, med gode alternativ, sjå Stephen Few's artikkel «Save the Pies for Dessert» på <http://www.perceptualedge.com/articles/08-21-07.pdf>.

¹⁴ R-brukarar kan bruka `dotchart()`-funksjonen for slike plott (figuren til venstre er basert på eksempelkode til funksjonen), eller kanskje `dotplot()`-funksjonen i `lattice`-pakken for ein litt meir avansert variant.

I kopla mikrokart vert slike plott kopla til små kart:

Mortalitetsrate kreft for menn 2001–2005



Dei ulike observasjonane (her mortalitetsraten for kvar delstat) vert sorterte, og så delte inn i små grupper. Kvar gruppe vert kopla til eit kart, og fargar vert brukte til å knyta grafobjekta til område på karta. Figuren viser eit kopla mikrokart for eit lite utval statar i USA, og me ser ein klar geografisk tendens¹⁵.

Prikkplott/punktplott, med eller utan tilhøyrande konfidensintervall, er berre éin av fleire mulige graftypar me kan bruka i eit kopla mikrokart. Me kan bruka søyle- eller stolpediagram for å vektlegga relative samanlikningar, piler til å visa endringar mellom to år, eller linjegrafar for å visa lengre tidsrekkjer. Me kan bruka boksplott – eller kanskje heller fiolinplott eller tettleiksestimat – for å visa «heile» fordelingar, og ikkje minst kan me kombinera fleire plott-typar bortover i kolonnar (der den siste kolonnen inneheld sjølve mikrokarta), for å samanlikna ulike variablar.

Men kopla mikrokart har enkelte avgrensingar. Me kan for eksempel ikkje visa *alle* kommunane i Norge samtidig (med mindre me har eit veldig langt ark). Men svært mange av eksempla i boka er basert på statane i USA, så iallfall 51 område (50 statar pluss Washington, D.C.) går bra¹⁶.

Teori

Boka har eit eige kapittel om teorien brukt til å utvikla mikrokarta, og eitt kapittel om generelle designprinsipp for datavisualisering. Forfattarane byggjer på omfattande forskning innan ei rekkje fagfelt, blant anna i persepsjon og kognisjon, datavisualisering, grafisk design og kartografi. Informasjon om korleis me reint fysisk ser, og korleis hjernen oppfattar og behandlar synsinntrykk og annan informasjon for å forstå, sjå mønster i, og å læra av, data er viktig å kjenna til om ein ønskjer å utvikla ei god grafisk framstilling av statistiske data.

Ei slik framstilling omfattar gjerne blant anna bruk av farge, form og plassering for å hjelpa hjernen å skilja, gruppera eller måla data. Det er dette som gjer at me oppfattar  som fire kolonnar med tre prikkar, og ikkje tre rader med fire prikkar (eller berre tolv prikkar i eit rutenett, som det «eigentlich» er), og som gjer at det er vanskelig å sjå om det er  eller  som er lengst, men det er svært lett å sjå det her  eller viss me samanliknar  og  (kryss-streken er plassert like langt frå venstre på dei to strekane).

Verken hovudtrekka eller detaljane i utforming av mikrokarta i boka er tilfeldige. Forfattarane har brukt tidligare forskning, men har òg utført mykje *ny* forskning direkte på mikrokart. Denne boka er resultatet (og ei samanfating) av over femten år med forskning, ikkje minst brukartesting.

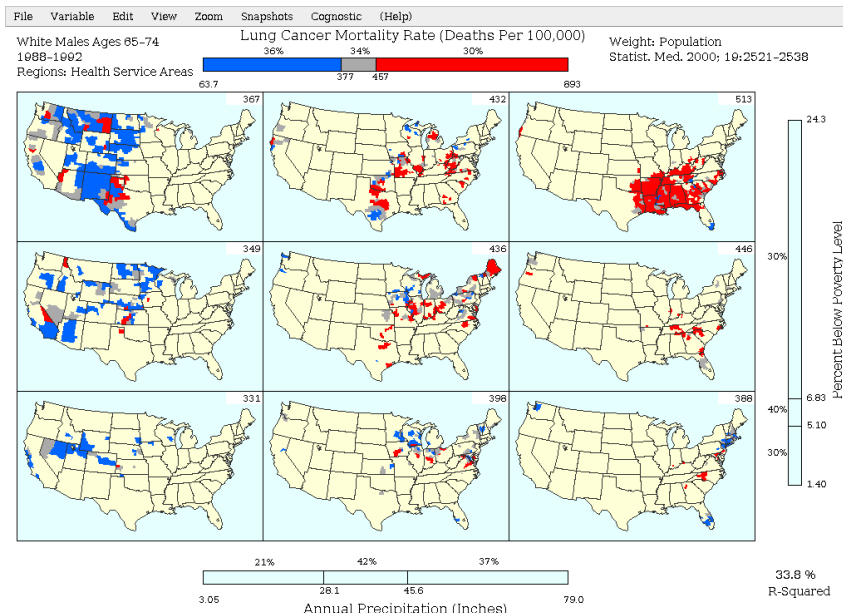
Conditioned micromaps

Boka inneheld òg annan bruk av mikrokart, og eitt kapittel handlar om det forfattarane kallar *conditioned micromaps*. Her ser me

¹⁵ Utan at me derfor automatisk skal trekka store og sikre slutningar av dette! Men me kan stilla oss spørsmål: Er tendensen like klar for andre tidsperiodar? Og for andre statar? Finst det ei relevant årsakssamanheng me kan utforska? Uansett er mikrokarta nyttige for å 1) generera hypotesar, og 2) presentera resultatata av grundigare geografiske analysar.

¹⁶ Arka i boka er noko kortare enn A4-ark, så me kan nok få plass til endå meir på «norske» ark. Og ved vising på skjerm kan me sjølv sagt ha uendelige lange «ark»!

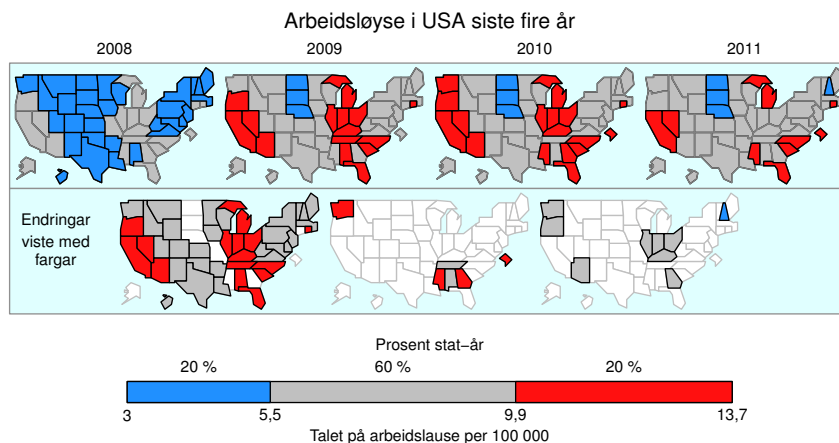
geografisk på éin responsvariabel og to forklaringsvariablar¹⁷, og me er interesserte både i samanhengen mellom responsvariabelen og forklaringsvariablane, og i eventuelle geografiske mønster. Verdiane til alle tre variablane vert delte inn i tre intervall, responsvariabelen vert koda med fargar på ni små koropletkart, og dei to forklaringsvariablane avgjer kva område som skal visast på kvart kart¹⁸:



Denne forma for mikrokart er veleigna til utforsking av (geografisk og annan) samanheng mellom variablane, for eventuell seinare formell modellering. Men merk at i motsetning til med kopla mikrokart, er det her nødvendig å bruka eit dataprogram, der ein sjølv kan flytta på intervallgrensene til dei tre variablane¹⁹.

Comparative miromaps

Den siste typen mikrokart er *comparative micromaps*, som er utvikla for samanlikning av geografiske data over tid. Dette er ikkje enkelt²⁰! Forfattarane si løysing er å visa endringane direkte²¹, på kart plasserte mellom kvart hovudkart:



¹⁷ Det kan med andre ord minna litt om ei geografisk form for tovegs ANOVA.

¹⁸ Den observante lesar ser kanskje her ein parallell i Cleveland's koplott (*conditioning plots*), der me har punktobservasjonar og vanligvis overlappende intervall. Prøv for eksempel `coplot()` i R (eller sjå tilsvarende funksjonalitet i `lattice`- og `ggplot2`-pakkane).

¹⁹ Kapittelet inneheld for øvrig forslag til automatisk val av potensielt «interessante» intervallgrenser.

²⁰ Om ein berre viser karta for kvar tidsperiode, og overlèt til brukaren å samanlikna dei, vert det lett som ei øving i «finn fem feil». For interaktiv bruk er animasjonar eit anna alternativ, men heller ikkje dette fungerer bra.

²¹ Å visa dei relevante størrelsane direkte er ein nyttig generell teknikk. Samanlikn for eksempel figur 6 og figur 28 i artikkelen til Cleveland og McGill. Figur 6 viser handelsbalansen til England på 1700-talet som to linjegravar – éin for import og éin for eksport. I figur 28 ser me direkte på differansen, og oppdagar noko interessant som var nesten umulig å sjå på den opphavlege grafen.

Eit problem her er at me ikkje alltid kan sjå direkte om endringa er positiv²² (frå blå til grå) eller negative (frå raud til grå). Dette løyser forfattarane ved å bruka fleire farge-tonar (lyseraud og lyseblå), eller ved å leggja til endå ei rad med endringskart, slik at éi rad viser positive endringar og den andre negative endringar. Dette fungerer nokolunde, men vert lett forvirrande.

Og når forfattarane seinare legg til rader med andre ordens differensar, eller prøver å samanlikna ratar for to grupper (for eksempel menn og kvinner) på same kart, vert resultatet vanskelig å forstå. Generelt er slike endringskart mindre umiddelbart forståelige enn dei to andre karttypeane (og sjølv *conditioned micromaps* er ikkje heilt enkle å forstå). Men endringskart er ein svært ny grafisk framstillingsmåte, og såleis framleis noko umoden. I boka vert han presentert som ein uferdig men mulig framstillingsmåte, som kan og bør betrast etter kvart som ein får meir erfaring i bruken. Og det er iallfall klart at *enkle* endringskart, som det ovanfor, fungerer rimelig bra til sitt formål.

Programvare

Mikrokart er interessant som konsept, men er lite verdt utan gode elektroniske verktøy for å laga dei. Heldigvis nemner boka fleire slike. *National Cancer Institute* har brukt kopla mikrokart i mange år, og har tilgjengelig gratis programvare for dette²³. Dessverre har eg ikkje fått prøvd dette, grunna tekniske problem²⁴. Men er ein berre interessert i å kjapt testa mikrokart i praksis, kan ein prøva nettsida *Cancer State Profiles*²⁵ til NCI. Her kan ein berre sjå på kreftstatistikk i USA, og ikkje alle funksjonane som er omtala i boka er støtta, men tenesta fungerer bra.

Hovudforfattaren har òg lagt ut nokre datasett og R-skript for dei ulike karttypeane på heimesida si²⁶. Men desse ber meir preg av å vera noko usystematisert, udokumentert kode, som rett nok fungerer (på dei medfølgjande eksempla), men som er langt frå så fleksible og enkle å bruka som dei burde ha vore. Nettsidene eller skripta ser heller ikkje ut til å vera aktivt vedlikehaldne. Her hadde det vore på sin plass med utvikling av ein ekte R-pakke, med skikkelig dokumentasjon!

På heimesida finn me for øvrig òg eit Java-program for *conditioned micromaps* – CCmaps²⁷. Sidan dette er ein karttype som hovudsaklig er utvikla for *interaktiv* bruk, er det nødvendig med eit slikt program, og programmet fungerer bra. Men det manglar både ein utskriftsfunksjon og eksportmuligheiter til PDF eller liknande publiseringsformat. Det finst heller ikkje R-kode for denne typen kart²⁸.

Mikrokart i Noreg

Er dei tre mikrokarttypeane noko som er aktuelt for presentasjon og analyse av geografisk relatert statistikk i Noreg? Ja, absolutt! Me har mykje talmateriale på områdenivå, anten fylke, kommune

²² Men sidan det er snakk om arbeidsløysetal, er det kanskje rimeligare å kalle slike endringar *negative*.

²³ <http://gis.cancer.gov/tools/micromaps/>

²⁴ Programmet er eit Java-program, og burde derfor automatisk fungera på alle system, men på grunn av dårlig programmering (med noko så trivielt som skråstrekar i filadresser) fungerer det ikkje på Linux. At programmet vert distribuert som ei *.exe*-fil gjer det òg noko vanskelig å komma til sjølv. Java-programmet dersom ein ikkje brukar Microsoft Windows.

²⁵ <http://statecancerprofiles.cancer.gov/micromaps/>

²⁶ <http://mason.gmu.edu/~dcarr/Micromaps/>

²⁷ <http://mason.gmu.edu/~dcarr/CCmaps/>

²⁸ Men det burde ikkje vera spesielt vanskelig å mekka til slike kart manuelt, for eksempel med `mapplot()`-funksjonen i `latticeExtra`-pakken.

eller grunnkrets, og mikrokart, spesielt kopla mikrokart, passar utmerkt for slike data. Men samtidig dukkar det opp nokre utfordringar som gjeld spesielt for vårt land.

Noreg er nemlig eit relativt langt (i motsetning til breitt) land. Dette fører til problem når det skal presenterast som mikrokart. Enten vert kartet så lite at det er vanskelig å sjå kvart fylke skikkelig, eller så må me ha stor loddrett avstand mellom kvart grafobjekt eller kvar observasjonsgruppe²⁹. Eitt alternativt vil vera å dreia kartet 90 gradar, slik at det vert liggande, men då vert det gjerne for lite gjenkjenneleg. Mitt forslag er å dela kartet i to (nord og sør), og plassera delane ved sidan av kvarandre, slik det av og til vert gjort i andre samanhengar. Dette er uheldig, men det er vanskelig å sjå betre løysingar.

Eit anna problem er at me har nokre svært små fylke, som Vestfold, som vert nærast usynlige på små kart. Her er løysinga rett og slett å forstørre desse. Dei nøyaktige grensene er ikkje viktige³⁰, og i mikrokarta brukte i boka er grensene svært forenkla, og nokre statar er forstørre, nettopp for å fungera på små kart. Samanlikn for eksempel det nøyaktige USA-kartet på side 16 med mikrokartet på side 19. Oslo fylke kan for øvrig vera aktuelt å flytta utanfor hovudkartet, slik det er gjort med *Washington, D.C.* på mikrokartet.

Diskusjonar som den ovanfor er noko eg saknar i boka. Nesten alle eksempla gjeld statane i USA³¹, og eksempel og forslag på korleis karta kan tilpassast land med annan geografi manglar. Sidan boka er meint å vera ei referansebok for mikrokart, burde det klart vore plass til slike detaljar.

Oppsummering

Boka er ei fin samanfating av over femten års forskning på korleis enkelte typar geografiske statistiske data kan presenterast og analyserast. Eit alternativ til å kjøpa/låna boka er sjølvstelt å lesa dei fleire titals forskingsartiklane forfattarane har skrive om emnet, men i boka vert alt sett saman og presentert på ein pedagogisk og oversiktlig måte, med lærdom trekt av gamle feil. Kapitla om designprinsipp i datavisualisering er òg generell nyttig lesnad, sjølv om det finst meir omfattande og betre bøker å lesa om ein er interessert i emnet.

Eitt problem med boka er at eksempla er USA-orienterte, at dei er noko lite varierte innhaldsmessig, og at ein del av dei rett og slett er noko *kjedelige*. Analysane forfattarane gjer er ikkje alltid særleg interessante, og heller ikkje alltid like overbevisande. Dette gjeld spesielt kapitla om *conditioned micromaps* og *comparative micromaps*. Men forklaringa på karta, teorien bak dei og detaljane er grundig forklart, og klart nyttig for ein som ønskjer å ta i bruk mikrokart.

Entusiasmen til forfattarane er òg synlig; dei *ønskjer* at mikrokart skal takast i bruk, og såleis forbetra grafisk framstilling av geografisk statistikk. Og her er eg einig. Mikrokart er noko me bør sjå meir av!

²⁹ Me kunne sjølvstelt valt å ha fleire observasjonar i kvar gruppe, men dette fungerer dårlig. Forfattarane tilrår maks fem i kvar gruppe.

³⁰ Noko som for Noreg kanskje kan vera like greitt, då Kartverket berre har svært *grove* grenser til «fri» bruk. Les for eksempel den triste saga om éin manns årelange (og enno ikkje avslutta) kamp for å få veta nøyaktig kor kommunegrensene går her i landet: <http://www.frigeonorge.net/>

³¹ Men eit litt artig eksempel i boka er eit kart over spelarplasseringar på ei baseballbane. Dette illustrerer at det ikkje berre er vanlige geografiske område som mikrokart kan brukast på.

Carr D.B. og Pickle L.W. (2010). *Visualizing Data Patterns with Micromaps*. Chapman & Hall/CRC. ISBN 978-1-4200-7573-1. [10.1201/EBK1420075731](https://doi.org/10.1201/EBK1420075731)

Nytt fra Norges teknisk-naturvitenskapelige universitet

Nye ph.d.-grader

Rupali Akerkar (12. juni)

Approximate Bayesian Inference for Survival Analysis

Veileder: Håkon Tjelmeland

Komit : Professor Emmanuel Lesaffre, Catholic University of Leuven, Belgium; professor Jan Terje Kval y, Universitetet i Stavanger; f rsteamanuensis Ingelin Steinsland, Institutt for matematiske fag, NTNU.

 yvind Salvesen (1. desember)

Statistical Models in Ice Hockey

Veileder: H vard Rue.

Komit : Professor Paul Fearnhead, Department of Mathematics and Statistics, Lancaster University, Storbritannia; Professor Jostein Lillest l, Institutt for Foretaks konomi, Norges Handelsh yskole, Bergen; F rsteamanuensis Mette Langaas, Institutt for matematiske fag, NTNU, Trondheim.

Nye mastergrader

Joseph Glory

Experimental Design for reliability improvement

Veileder: John Tyssedal

Kari Halle

Statistical methods for multiple testing of GWAS data

Veileder: Mette Langaas

Marthe L. Hestnes

Modelling the Conditional Correlation Structure Between Real Rate of Returns

Veileder: John Tyssedal

Kahle G. Ismail

The use of interest rate models to analyse macroeconomical changes

Veileder: Jacob Laading

Marie Lilleborge

Bruk av skranker for Markovfelt for   effektivisere utregning av optimale beslutninger i grafiske modeller

Veileder: H kon Tjelmeland

Bj rn E. Loeng

The probability of collision between petroleum wells

Veileder: Jo Eidsvik

Johanne Norstein

Parameter in a Markov mesh model by reversible jump MCMC simulation

Veileder: H kon Tjelmeland

Kjetil Ones

On the use of a quasi-distribution as the distribution of the first observation of a population sizes time series

Veileder: Jarle Tufto

Sindre Solibakke

Realized GARCH: evidence in oil contracts

Veileder: Arvid Næss

Elisabeth Straume

Prediction of petrophysical variables from seismic data

Veileder: Jo Eidsvik

Paul V. Vo

Ensemble Kalman filter on the Brugge field

Veileder: Henning Omre

Nytt fra Universitetet for miljø- og biovitenskap

Nye mastergrader

Kristine Navestad

Estimation of mixture-distributions using the Direct Look-up method

Veiledere: Solve Sæbø (hovedveileder) og Trygve Almøy (medveileder)

Ólafur H. Kristjánsson

Development and evaluation of prediction equations for NIR instrument, measuring fat in Salmon fillets

Veiledere: Trygve Almøy (hovedveileder) og Solve Sæbø (medveileder)

Nytt fra Universitetet i Bergen

Nye mastergrader

Kristine B. Solbjørg

Lee-Carter-modellens tilpasning til norske dødelighetsdata for perioden 1846 til 2009

Veileder: Hans A. Karlsen

Håkon Otneim

Local Likelihood

Veiledere: Dag Tjøstheim og Hans A. Karlsen

Hallvard Røyrane-Løtvedt

GAMLSS-modeller i bilforsikring

Veilder: Hans J. Skaug

Lisbet Lien Harjo

Statistiske modeller basert på skjulte markovkjeder i kontinuerlig tid

Veileder: Ivar Heuch

Tommy Nordbø

Copulas and local Gaussian correlation

Veilder: Dag Tjøstheim

Nytt fra Avdeling for biostatistikk, UiO

Nye ph.d.-grader

Yiting Xue (8. mars 2012)

Mathematical and statistical models of infectious diseases and their applications in health economic evaluation (prøveforelesning: *Health economic perspectives on interventions to manage infectious diseases exemplified with school closure during an influenza pandemic*)

Bedømmelseskomité: Associate professor Christian Kronborg, 1ST Forskningsenheden for Sundhedsøkonomi, Syddansk Universitet, Odense, Danmark; associate professor Lisa Brouwers, Kungliga Tekniska högskolan, Skolan för informations- och kommunikationsteknik, Kista, Sverige; førsteamanuensis Knut Reidar Wangen, Avdeling for helseledelse og helseøkonomi, Institutt for helse og samfunn, Universitetet i Oslo. Veileder: Birgitte Freiesleben de Blasio, Avdeling for Biostatistikk.

Hayat Mohammed (12. mars 2012)

Statistical tools for high-dimensional genomic data (prøveforelesning: *Nonparametric methods for density comparison*)

Bedømmelseskomité: Associate Professor Clelia Di Serio, University Centre of Statistics in the Biomedical Sciences, Vita-Salute San Raffaele University, Milano, Italia; førsteamanuensis Solve Sæbø, Institutt for kjemi, bioteknologi og matvitenskap, Universitetet for miljø- og biovitenskap, Ås; førsteamanuensis Anja Bråthen Kristoffersen, Institutt for informatikk, Det matematisk-naturvitenskapelige fakultet, Universitetet i Oslo. Veileder: Arnoldo Frigessi, Avdeling for biostatistikk.

Ingvild Dalen (15. juni 2012)

Correction for measurement error and misclassification in epidemiologic exposure data (prøveforelesning: *Betydningen av målefeil for estimering av tilskrivbare andeler*)

Bedømmelseskomité: Professor Håkon K. Gjessing, Nasjonalt folkehelseinstitutt, Oslo; professor Geir Egil Eide, Institutt for samfunnsmedisinske fag, Det medisinsk-odontologiske fakultet, Universitetet i Bergen; professor Lene Frost Andersen, Institutt for medisinske basalfag, Universitetet i Oslo. Veileder: Magne Thoresen, Avdeling for biostatistikk.

Nytt fra Avdeling C ved Matematisk institutt, UiO

Nye mastergrader

Torbjørn D. Jakobsen

Variance Swaps as a Hedge for Insurers

Veileder: Erik Bølviken

<http://urn.nb.no/URN:NBN:no-30708>

Charlotte Lima

Marginal Structural Models applied to treatment for HIV infection

Veiledere: Odd Aalen og Sven Ove Samuelsen

Kristina Rognlien Dahl
Convex duality and mathematical finance
Veileder: Bernt Øksendal

Iben Cathrine Simonsen
Pricing of spread options in energy markets
Veileder: Fred Espen Benth

Vu Duy Pham
Trading in the energy market by multicointegration and stochastic optimal control theory
Veileder: Fred Espen Benth

Alexander Dyrseth
Optimal Bond Portfolios in a Heath-Jarrow-Morton Framework
Veileder: Fred Espen Benth

Sindre Froyen
Computation of Greeks in Financial Markets Driven by Levy Processes
Veileder: Fred Espen Benth

Olav Nikolai Breivik
Optimale belastningsvalg i terskelsystemer
Veileder: Arne Bang Huseby

Rune Hoff
Properties of Estimators for Cumulative Baseline Hazard from Case-control Studies in the Relative Risk Model
Veileder: Sven Ove Samuelsen

Fratredelser

Professor Inge Helland går av med pensjon fra 1. august. Han fortsetter som professor emeritus.

Pål Lillevold slutter som amanuensis II i forsikringsmatematikk.

Nytt fra Universitetet i Tromsø

Ny mastergrad

Berner Larsen
A threshold cointegration analysis of Norwegian interest rates
Veiledere: Edvin Bredrup og Sigrunn Holbek Sørbye

Ny ph.d.-grad

Kristian Hindberg (23. mars)
Scale-Space Methodology Applied to Spectral Feature Detection, Multinormality Testing and the k-sample Problem, and Wavelet Variance Analysis
Veiledere: Fred Godtliebsen og Håvard Rue
Komité: Professor Lasse Holmström, Department of Mathematical Sciences, University of Oulu, Finland; professor Harald Krogstad, Department of Mathematical Sciences, NTNU; associate professor Robert Jenssen, Department of Physics and Technology, Universitetet i Tromsø.

SCANDINAVIAN JOURNAL OF STATISTICS

Volume 39 • Number 2 • June 2012

Contents:	Page
SJS Lecture	
The Current Duration Approach to Estimating Time to Pregnancy <i>Niels Keiding, Oluf K. Højbjerg Hansen, Ditte Nørbo Sørensen & Rémy Slama</i>	185
Discussion of 'The Current Duration Approach to Estimating Time to Pregnancy' <i>Odd O. Aalen</i>	205
Discussion of 'The Current Duration Approach to Estimating Time to Pregnancy' <i>Philip Hougaard</i>	207
Reply to the Comments by Drs Aalen and Hougaard on 'The Current Duration Approach to Estimating Time to Pregnancy' by Niels Keiding et al. <i>Niels Keiding, Oluf K. Højbjerg Hansen, Ditte Nørbo Sørensen & Rémy Slama</i>	210
Original Articles	
Regularized Posteriors in Linear Ill-Posed Inverse Problems <i>Jean-Pierre Florens & Anna Simoni</i>	214
A Matching Prior for the Shape Parameter of the Skew-Normal Distribution <i>Stefano Cabras, Walter Racugno, María Eugenia Castellanos & Laura Ventura</i>	236
Regression Analysis of Longitudinal Data with Time-Dependent Covariates and Informative Observation Times <i>Xinyuan Song, Xiaoyun Mu & Liuquan Sun</i>	248
Estimating the Conditional Error Distribution in Non-parametric Regression <i>Sebastian Kiwitt & Natalie Neumeyer</i>	259
A Convolution Estimator for the Density of Nonlinear Regression Observations <i>Bård Støve & Dag Tjøstheim</i>	282
Model Checks in Inverse Regression Models with Convolution-Type Operators <i>Nicolai Bissantz, Holger Dette & Katharina Proksch</i>	305
Scale Checks in Censored Regression <i>Holger Dette & Cedric Heuchenne</i>	323
Confidence Regions for Means of Random Sets Using Oriented Distance Functions <i>Hanna Jankowski & Larissa Stanberry</i>	340

Two-Sample Test Against One-Sided Alternatives 358
Teresa Ledwina & Grzegorz Wyłupek

False Discovery Rate Control of Step-Up-Down Tests with Special
Emphasis on the Asymptotically Optimal Rejection Curve 382
Helmut Finner, Veronika Gontscharuk & Thorsten Dickhaus

Published on behalf of:
Danish Society for Theoretical Statistics
Finnish Statistical Society
Norwegian Statistical Society
Swedish Statistical Society

Har du husket å betale kontingenten i år?

Medlemskontingenten i NSF er kr 200,- per år for ordinære medlemmer, mens mastergradsstudenter får gratis medlemskap. På adresselappen bakpå bladet står det anført siste år vi har mottatt kontingent fra deg. Sjekk adresselappen! De som ikke har betalt for tidligere år gjør dette sammen med kontingenten for 2012. Dersom du mener betalingsinformasjonen ikke stemmer, gi beskjed til clara-cecilie.gunther@nr.no. Dersom arbeidsgiveren din betaler kontingenten, sørg for at navnet ditt kommer med på innbetalingen.

De største utgiftspostene for NSF er overføringer til lokallagene i Oslo, Bergen og Trondheim, utdeling av Sverdrup-prisen, NSFs medlemskap i The European Mathematical Society (EMS) og the International Statistical Institute (ISI), trykking og utsending av TG og utgifter i forbindelse med diverse arrangementer – blant annet støtte til studenter ved deltakelse på det norske statistikermøtet.

For at medlemsregisteret skal være best mulig oppdatert, er det viktig at du oppgir eventuell ny post- eller e-postadresse, ved å sende meg en e-post. Vi har også opprettet en egen e-postliste for NSFs medlemmer, nsf-medlemmer@math.uio.no. Dersom du ikke mottar e-post fra denne lista, send meg din oppdaterte e-postadresse.

Betalingsfrist for innbetaling av kontingent er snarest. Vennligst oppgi hvem kontingenten gjelder for ved innbetaling.

Clara-Cecilie Günther
Kasserer NSF
Telefon: 22 85 25 63
Konto: 0530 27 73299

<http://www.euro-math-soc.eu/>

<http://isi-web.org/>

Betalingsinformasjon

Kontingent NSF 2012 for [Navnet ditt]

GIRO

Betalings-
frist

snarest

Underskrift ved girering

Betalt av

Betalt til

NSF ved Clara-Cecilie Günther
Norsk Regnesentral
Postboks 114 Blindern
0314 Oslo

Belast
konto

Kvittering
tilbake

Kundeidentifikasjon (KID)

Kroner

Øre

Til konto

Blankettnummer

H

200

00

<

>

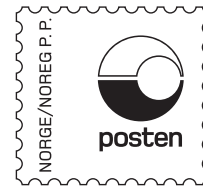
0530 27 73299

<

>

Avsender

Norsk statistisk forening
v/Clara-Cecilie Günther
Norsk Regnesentral
Postboks 114 Blindern
0314 Oslo

**Redaksjon**

Anne Marie Fenstad
Karl Ove Hufthammer
Bård Støve