

Qvintensen

MEDLEMSTIDNING FÖR SVENSKA STATISTIKFRÄMJANDET NR 3 2011

Klimatets statistik:

Klimatet är fördelningen av väder

SIDAN 16

Ett liv med
statistik
– möt Ove
Frank

SIDAN 6

Är det
farligt att
titta
på TV?

SIDAN 9



**Det blir allt viktigare att
snabbt upptäcka utbrott
av epidemiska sjukdomar.»**

Linus Schiöler

SIDAN 19

Statistik,
sanning och
kunskap

SIDAN 12

PRICKSÄKRA KUNSKAPER OM IMORGON. BARA SÅ DU VET.

DET FINNS TRE SLAGS LÖGNER SA MARK TWAIN EN GÅNG: LÖGN, FÖRBANNAD LÖGN OCH STATISTIK.

Vi håller inte med. I vår värld finns det tre slags sanningar: fakta, avancerad analys och statistik.

Det är med de grunderna som SAS Institute skapar kunskap och beredskap inför framtidens scenarion. Just lösningar för beslutsstöd med avancerad analys är vi världsledande på.

På sas.com/sweden hittar du alltid aktuella seminarier, webb-seminarier och utbildningar, vare sig du är befintlig eller blivande kund.

Välkommen in.



SAS Institute är ledande på lösningar för beslutsstöd med avancerad analys. SAS Institute, även världens största privatägda mjukvaruföretag, omsatte 2009 2,31 miljarder dollar i 118 länder och har sedan 1976 erfarenhet av att utveckla verktyg och metoder som låter stora organisationer lära av sin historia, mäta och kommunicera pågående aktiviteter och inte minst att skapa insikt om framtiden. Världen runt har SAS Institute totalt gjort 45 000 kundinstallationer, bland annat i 92 procent av Fortune 500-företagen. I Sverige startade SAS Institute AB år 1986 och har idag drygt 100 anställda på kontoret i Stockholm. Bland de svenska kunderna finns landets mest betydande företag och organisationer. För mer information besök: www.sas.com/sweden



Medicinutveckling.



Pernilla
Andersson
– ny i statistik-
främjandets
styrelse.



Seans.



Londons tunnelbanesystem.

Innehåll

ETT LIV MED STATISTIK

– möt Ove Frank 6

Prolonged TV viewing and risk of chronic
disease and premature death..... 9

Causal inference – a book review 11

Några funderingar om statistik
och kunskapsteori..... 12

MINA ÄVENTYR I PARAPSYKOLOGIS FÖRTROLLADE VÄRLD

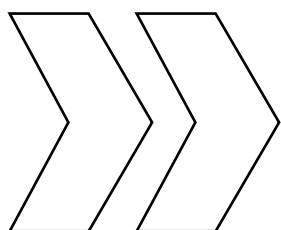
Del 1. Vad är parapsykologi och
vilka är forskarna? 14

KLIMATET/

Seminarium om klimatmodeller..... 16

Upp- och nedskalning av klimatmodeller..... 17

ÖVERVAKNING AV UTBROTT 19



**Att som student, under handledning,
få praktisera som konsulterande
statistiker är otroligt värdefullt och
lärorikt.»**

Anna Torrång, sidan 24

Fasta spalter

I HUVUDET PÅ EN REDAKTÖR 4

ORDFÖRANDE HAR ORDET 5

HÄNT & HÖRT 8

KALENDARIUM 28



Våra föreningar

SURVEYFÖRENINGEN

Öppna data vinner 21

Surveymetoder på underbar plats..... 22

FÖRENINGEN FÖR MEDICINSK STATISTIK

Minnen, MBDD och stipendiater..... 24

Model based drug development..... 25

CRAMÉRSÄLLSKAPET

Sommarskola på Asa herrgård 27

Ansvarig utgivare

Hans Nyquist

Redaktör

Dan Hedlin, 070-578 4334

Ingegerd Jansson (biträdande)

Redaktion

Hans Ekholm, Främjandet

Alf Fyhrlund, Främjandet

Marie Linder, FMS

Lennart Nordberg, Surveyföreningen

Lars Söderström, Industriell statistik

Ann-Marie Flygare, Cramérssällskapet

Adress

Qvintensen, c/o Dan Hedlin

Runbyvägen 31, 194 44 Upplands Väsby

E-post qvintensen@gmail.com

Produktion

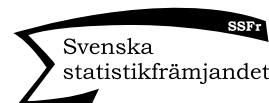
Form och redigering: Mezzo Media AB

Tryckeri: Trydells Tryckeri AB

Annonser

Annonser i Qvintensen bokas med redaktören.

Annonssutskick per e-post bokas med Svenska statistikfrämjandets sekreterare. Priset är 5 000 kronor för institutionsmedlem eller motsvarande, och 7 500 kronor för övriga annonsörer. På alla priser tillkommer 25 % moms.



SVENSKA STATISTIKFRÄMJANDETS STYRELSE

Ordförande

Hans Nyquist, 08-16 11 32,

hans.nyquist@statstat.se

Vice ordförande

Jun Yu, 090-78 68 468

Sekreterare

Chandra Adolfsen

Pernilla Andersson

Skattmästare Jonas Selling

Klubbmästare Therese Andersson

Surveyföreningen

Joakim Malmidin, 08-506 949 45

FMS

Mats Rudholm, 031-703 73 72

Industriell statistik

Göran Lande, 070-526 26 18

Cramérssällskapet

Jan-Eric Englund 040-41 52 36

Övriga ledamöter

Henrik Hult, Mathias Lindholm

Adress

Svenska statistikfrämjandet

c/o Chandra Adolfsen

Statistiska centralbyrån

701 89 Örebro

E-post sekrfram@gmail.com

Webbplats www.statistikframjandet.se

Allt som påverkar oss att välja



Långt efteråt inser man vilken enorm betydelse ens gamla lärare och handledare hade. Min mattelärare i gymnasiet hade jag ett komplicerat förhållande till men hon påverkade det yrkesval som jag gjorde långt senare.

I det här numret minns Anna Torráng Erik Leander, hans hårsvall och hans stora betydelse för hennes utveckling i yrket. Vi gör också ett avsteg från att välja intervjupersoner slumpmässigt, vilket är naturligt för statistiker men för inga andra. Ingegerd Jansson intervjuar Ove Frank, som betydde mycket för hennes fortsatta yrkesliv.

Något det nu skrivs mycket om är öppna data, alltså data som är fritt tillgängliga. Det finns numera en förväntan att allt ska vara åtkomligt när som helst. Att en del data är

»Vad är egentligen förhållandet mellan data och kunskaper?»

konfidentiella glöms ofta bort. I sin översikt av den officiella statistikens historia, med start i Qvintensen nr 4/2009, berättade Ulf Jorner att dåvarande regeringen ville hålla befolkningsstatistiken hemlig medan vetenskapsmännen ville sprida den. Här skriver Joakim Malmidin om olika initiativ världen över mot ökad öppenhet. Steget är långt från hemlighetsmakeriet på 1700-talet till helt öppna data på internet.

Vad vet vi egentligen om vår värld? Vi kan samla in data, analysera dem, lägga ut dem på ett öppet sätt, men vad är egentligen förhållandet mellan data och kunskaper? I Qvintensen nr 3/2010 skrev Olof Ljungström om skallmätningar som trots överraskande stringent insamling av data aldrig ledde till kunskaper, eftersom frågeställningarna var djupt förfelade. Joakim Westerlund startar nu en artikel om parapsykologi, detta evigt kunskapsfattiga ämne trots delvis rigorös forskning.

Våra numera goda kunskaper om klimatet har ett komplext förhållande till data. Georg Lindgren och Peter Guttorp skriver på sidan 16 om aktuella statistiska problem i klimatforskningen. Tommy Johnsson går ännu djupare i sin artikel. Vad har statistik för roll i vårt byggande av kunskap? Det är inte alls unikt för statistik att utövarna har dålig koll på sina mest grundläggande begrepp. Tänk efter själv, hur skulle du förklara begreppet sannolikhet för dig själv och för andra?

DAN HEDLIN



FOTO: MONICA HOLMBERG

ORDFÖRANDEN HAR ORDET

VI HAR EN SPÄNNANDE HÖST ATT SE FRAM EMOT

Välkomna tillbaka efter sommarledigheterna. Nu när hösten kommit hoppas jag att alla har haft en skön och vilsam sommar och att solen har visat sin allra vänligaste sida. Efter en hektisk vår med alla bestyr och göromål som skall vara klara innan semesterna bryter ut behöver vi verkligen en skön och avkopplande ledighet. Och efter en skön avkoppling och vila kommer entusiasmen att ta tag i nya arbetsuppgifter tillbaka. Därför kändes det extra roligt att kavla upp ärmarna och starta upp hösten för Svenska statistikfrämjandet med ett styrelsemöte.

År 2010 hade vi dryga räkningar att ta hand om, vilket allvarligt ansträngde Statistikfrämjandets ekonomi. Samtidigt var medlemsantalet vikande. På vårt första möte för hösten kunde vi glädjande nog konstatera att ekonomin är på uppåtstående igen. Vad gäller medlemsantalet är det också på väg uppåt, men här behöver vi jobba ytterligare lite för att nå upp till de mål vi har satt upp. Vi behöver både fler enskilda personer och fler företag som medlemmar. Ett mycket bra sätt att värva nya medlemmar är att var och en av oss berättar för våra kollegor och vänner om fördelarna med att vara medlem i Svenska statistikfrämjandet. Ta det som en utmaning att värva nya medlemmar, både personer och företag är välkomna!

Styrelsen har blivit uppmärksam på att Statistiska föreningen hade en lägre årsavgift för pensionärer, något som inte har tagits upp av Svenska statistikfrämjandet. Jag tror att det finns en klar majoritet bland medlemmarna som tycker att pensionärer skall ha reducerad årsavgift. Styrelsen kan dock inte besluta om årsavgifternas storlek; detta kan bara beslutas på medlemsmöten. Därför måste vi ha kvar de beslutade avgifterna i år och vid nästa årsmöte föreslå differentierade avgifter för kommande år.



**Hans Nyquist,
ordförande.**

**»Vi behöver jobba
ytterligare lite för
att nå upp
till de mål vi har
satt upp»**

Ett av de större problemen vi har att brottas med nu är medlemsregistret. Registret är tungarbetat och förmår inte uppfylla de krav som Främjandet och dess medlemsföreningar ställer på ett medlemsregister. Styrelsen har därför tillsatt en arbetsgrupp med uppgift att utforma ett nytt och mer ändamålsenligt medlemsregister. Arbetsgruppen består av Pernilla Andersson, Annika Lindblom, Mikael Möller och Joakim Malmelin. Målsättningen är att skapa ett gemensamt register för medlemmar i både Svenska statistikfrämjandet och medlemsföreningarna. Registret skall vara lätt att administrera för alla som behöver arbeta med registret.

Det skall t.ex. vara lätt att ändra adressuppgifter, det skall vara lätt för föreningarna att hitta sina medlemmar för utskick av information om föreningsaktiviteter och det skall vara lätt att ta reda på exakt hur många som är medlemmar i respektive förening. Som en följd av ett gemensamt medlemsregister ser jag också framför mig en förenklad procedur vid betalning av årsavgifterna, nämligen att det räcker med ett gemensamt formulär och en gemensam räkning.

Ett intressant kommande datum är den 28 oktober. Då ordnar nämligen Cramérsällskapet en ämneskonferens som är gemensam för de båda ämnena statistik och matematisk statistik. Konferensen kommer att äga rum i Uppsala. Låt oss hoppas att detta är början till en lång tradition där båda våra statistikämnen kan mötas och utbyta erfarenheter.

Avslutningsvis vill jag nämna att planeringsarbetet för Nordiska konferensen i matematisk statistik pågår för fullt, bl.a. är nu listan på huvudtalare klar. Titta gärna på den information som arrangörerna ger på konferensens webbplats. Se kalendariet på sidan 28.

Som ni förstår har vi en spännande höst i Svenska statistikfrämjandet att se fram emot med många intressanta arbetsuppgifter och utmaningar. Då är det bra att sommaren verkligen har varit skön och vilsam.

Ett liv med statistik

– NEDRÖKTA RUM, DECHIFFRERING, TVÄRVETENSKAP OCH BRETT SAMARBETE

■ Ove, hur kom det sig att du började läsa matematisk statistik?

– Jag tänkte bli kemist men fick veta att man skulle börja med att läsa matematik. Det gjorde jag, och det var roligt, så jag letade efter andra ämnen som var matematiska och hittade matematisk statistik, mekanik och nationalekonomi. Mekanik var knepigt, nationalekonomi lätt och matematisk statistik roligast.

■ Du skrev din licentiatavhandling i matematisk statistik?

– Ja, den handlade om signalstörning, en sänd signal störs och ska sedan rekonstrueras. Jag använde entropimått, det var första gången jag arbetade med det (Ove skrev en diskussionsartikel i Qvintensen nr 2/2011 om nyttan med entropimått). Inspirationen kom från min militärtjänst vid FRA (Försvarets radioanstalt). Det var en bra skola. Vi lyssnade av signaler men det fick jag inte skriva, det var ju hemligt, så jag gjorde om problemet till stokastiska processer med diskreta tillstånd. Någon tillämpning fick jag inte nämna. På FRA fanns Arne Beurling, han som knäckte tyskarnas krypteringsmaskin G-skrivaren under andra världskriget.

Många ryssar har skrivit om sådant här men maskerat som ren



Ove Frank.

Vi inleder här en serie intervjuer med personer som har lång erfarenhet av statistiskt arbete.

Vi börjar med Ove Frank, professor emeritus vid Stockholms universitet.

matematik, de fick ju inte skriva om militär verksamhet som var hemlig. De här gamla metoderna har dykt upp igen i ny form, till exempel i bankvärlden.

■ Vad gjorde du när du var klar med din lic?

– Då började jag på FOA, Försvarets forskningsanstalt. Där kom jag i kontakt med nätverksproblem för första gången. Det är tillämpligt även i det militära, till exempel på järnvägsnätet. Jag skrev om militära transporter på Inlandsbanan, det gällde att optimera användningen så att inte alla tåg samlades i Norrland. De militära nätverkstillämpningarna ledde vidare till mitt intresse för grafer. Det här var på sextiotalet, på sammanträdena på FOA rökte de så mycket att man måste tvätta håret varje kväll. Ulf Grenander som varit min handledare för licen ville att jag skulle följa med till USA när han flyttade till Brown University. Jag kände mig inte redo för det, men fick senare tillfälle att besöka honom och Tore Dalenius vid Brown flera gånger.

Så småningom fick jag medel från Riksbanken och ett lektorat. Tore Dalenius var min handledare när jag fortsatte med avhandlings-

arbetet och jag disputerade 1971 på avhandlingen Statistical inference in graphs.

Innan dess var jag även på KTH där jag var assistent tillsammans med Olav Kallenberg. En enormt intelligent och kreativ man som senare skrev en mycket lovordad bok i sannolikhets-teori. Vi gjorde tentor på KTH, men Olavs tal var ibland alldeles för svåra för teknologerna.

■ Efter avhandlingen har du fortsatt verka inom universitetsvärlden.

– Ja, jag blev biträdande professor i Uppsala där jag var i tre år. Därefter blev jag professor i Lund, där stannade jag i tio år innan jag kom till Stockholms universitet 1984. Där tyckte jag att forskningsmiljön var den klart bästa.

■ Berätta lite om ditt internationella samarbete.

– Jag har samarbetat mycket internationellt genom åren i min forskning, det är ofta väldigt trevligt. På åttiotalet tog jag ett sabbatsår och var i Kalifornien, vid Stanford University och University of California, Riverside, och jag har varit inbjuden till USA många gånger, det har varit väldigt berikande. Där fanns det många stora namn inom grafteori och tillämpningen sociala nätverk som jag kommit i kontakt med och samarbetat med, till exempel Frank Harary och Linton Freeman. Jag träffade också Paul Erdős, den oerhört produktive ungerske matematikern och grafteoretikern. Han reste runt på konferenser med sin mamma som tog hand om honom och han var alltid barfota i sandaler.

Sociala nätverk är ett tvärvetenskapligt om-



Optimering av tågtrafik är en tillämpning av nätverksteori.

råde där folk har olika bakgrund, de kan komma från psykologi, sociologi, datavetenskap eller något annat. Det är naturligt med samarbete inom ett sådant område och att som statistiker ta tag i deras problem. Jag har samarbetat tvärvetenskapligt både inom och utanför Sverige. Speciellt har arbete vid Statens institutionsstyrelse och inom ett EU-projekt om ungdomar med missbruksproblem varit givande. Mitt samarbete med statistiskt intresserade kriminologer och psykologer har också varit mycket givande.

Det är viktigt att forskningen känns rolig och betydelsefull. Man ska ut och träffa folk, åka på konferenser, initiera samarbeten, inte bara sitta på sin kammare.

Det är också ett bra sätt att få ut doktorander, det är viktigt som handledare att ha kontakter och träffa folk med olika sätt att se på forskningsproblemen.

■ Vad är enligt din erfarenhet viktigt i undervisningen?

– När jag själv hade läst en del matematisk statistik vid Stockholms högskola fick jag börja i den statistiska forskningsgruppen på institutionen. Det var en fristående grupp där man arbetade praktiskt som konsult och där lärde jag mig mer, utöver läroböckerna. Den typen av verksamhet

borde in i undervisningen. Eleverna måste förstås ha läst några kurser först, men sedan ska de få ansvaret för att diskutera problem med uppdragsgivare och självständigt eller tillsammans med handledare finna lösningar. Det ger inblick i många olika sorters tillämpningar och metodfrågor, och om det är på konsultbasis så kan det ju dessutom göra att eleverna får in en fot på arbetsmarknaden.

Statistiker har god nytta av matematik. Men matematik har kommit i stryckklass i gymnasiet. Problemlösning är inte viktigt längre. Man kan inte bara räkna samma uppgifter om och om igen, det blir tråkigt. Man måste få upp ögonen för att matematiken är det naturliga språket för att exakt och koncist beskriva frågeställningar och

representera företeelser man vill undersöka. Matematiken berikar ens liv. Det är mycket mer än att räkna, det är ett sätt att se möjligheter, ett sätt att tänka och strukturera. Man borde betona modelltänkande mycket mer i statistiken och inte vara främmande för att pröva flera alternativa modeller och variabelval innan man bestämmer sig för en lämplig analys.

Dessutom borde man modernisera läroböckerna i statistik. De borde i högre grad ta upp statistiska kunskaper som är relevanta i andra ämnen. Inom sociologi till exempel behöver man speciella statistiska kunskaper som inte tas upp i de vanliga statistikböckerna. Det är inte så konstigt att sociologer tycker att de klarar sig utan statistiker. Där borde statistikerna ta sitt ansvar och skriva om till exempel skaltyper och variabeldefinitioner på ett användarvänligt sätt. Man borde generellt i statistiska läroböcker ge systematiska översikter av modelltyper i olika situationer och inte nödvändigtvis gå in på beräkningsmetoder för olika analyser innan man lärt sig vilken rikedom av möjligheter som finns. Risken är stor annars att man tror att den modell man lärt sig bäst alltid går att använda.

INGEGERD JANSSON

NYA I SVENSKA STATISTIK- FRÄMJANDETS STYRELSE

Jun Yu, vice ordförande

– Jag är professor i matematisk statistik och verksam vid Biostatistikum vid Sveriges lantbruksuniversitet (SLU) i Umeå. Jag är född i Kina och tog en magisterexamen i matematik med inriktning mot matematisk statistik vid Hangzhou University 1982. Så småningom blev jag universitetslektor där. 1990 flyttade jag till Umeå för forskarutbildning i matematisk statistik. Efter en tid vid institutionen flyttade jag till SLU där jag anställdes som forskare i matematisk statistik. Arbetet har bestått i undervisning, handledning, forskning och konsultverksamhet.

– Jag har varit studierektor för forskarutbildningen vid Institutionen för skoglig resurshushållning och geomatik och för forskarskolan RESE (fjärranalys för miljön) och är också verksam vid Centrum för medicinsk teknik och fysik. Min nuvarande forskning handlar främst om stokastisk modellering och statistisk analys av tidsrumsliga data. Jag har arbetat med ett flertal forskningsprojekt med tillämpningar som sträcker sig över ett brett spektrum, från skogsbruk, vattenbruk, miljöövervakning, ekologi och populationsbiologi till fjärranalys, ekonomi och medicinsk teknik.



Mathias Lindholm, ledamot

– Mitt nuvarande arbete handlar om att värdera AFA Försäkrings åtaganden gentemot de försäkrade samt att analysera de risker som finns i dessa åtaganden. Primärt syftar detta arbete till att säkerställa att AFA Försäkring sätter av tillräckliga reserver för att kunna klara sina försäkringsåtaganden, ett arbete som görs i nära samarbete med bolagets kapitalförvaltning. Inom ramen för detta gör jag även diverse utvecklingsarbete.



HÄNT OCH HÖRT

■ Den 10 juni disputerade Susanna Björkwall, Matematisk statistik vid Stockholms universitet, på en avhandling med titeln Stochastic claims reserving in non-life insurance: bootstrap and smoothing models. Fakultetsopponent var Dr Peter England, Towers Watson, New York.

■ Även Bertil Wegmann disputerade den 10 juni i Stockholm. Bertil är på Statistiska institutionen och hans avhandling har titeln Bayesian inference in structural second-price auctions. Fakultetsopponent var Samita Sareen, Associate Professor, på Charles River Associates i New York.

■ Den 26 augusti disputerade Patrik Andersson, Matematisk statistik vid Stockholms universitet, på en avhandling med titeln Four applications of stochastic processes: contagious disease, credit risk, gambling and bond portfolios. Fakultetsopponent var professor Stewart Ethier, University of Utah, USA.

Qvintensen rättar

■ I Magnus Arnér och Kerstin Vännmans artiklar "Att mäta duglighet hos en tillverkningsprocess" och "Vad är kapabilitetsindex?" publicerades olyckligt nog utkastversionerna av figurerna. Rätt version finns på Qvintensens sida hos Svenska statistikfrämjandets webbplats. Vi ber Magnus och Kerstin om ursäkt för misstaget.



Mer på Qvintensens webbplats

■ Magnus Arnér och Kerstin Vännmans artikel "Att mäta duglighet hos en tillverkningsprocess" i Qvintensen nr 2/2011 har lästs av många. På Svenska statistikfrämjandets webbplats under fliken Qvintensen finns en diskussion av Lars Ängquist.

Bli medlem

■ Svenska statistikfrämjandets syfte är bland annat att främja sund användning av statistik som beslutsunderlag och att väcka och sprida intresse för statistik i samhället. Om du önskar gå med som medlem i Svenska statistikfrämjandet skicka ett e-brev till Chandra Adolfs-



Har du flyttat?

■ Meddela din nya adress genom att gå in på Svenska statistikfrämjandets webbplats www.statistikframjandet.se, välj Medlemsregister och följ instruktionerna. Ditt användar-id är tryckt ovanför din adress på tidningens baksida.

son och Pernilla Andersson på sekrfram@gmail.com. Du får Qvintensen i brevlådan och platsannonser via e-post. Det ställs inga krav för att bli medlem; alla som är intresserade av statistik och vill stödja statistikens roll i samhället är välkomna.

Prolonged TV viewing and risk of chronic disease and premature death:

IS THE RELATIONSHIP CAUSAL?



Är det farligt att titta på tv? Eller är det så att alltför många dagliga timmar framför tv:n tränger ut mer hälsosamma aktiviteter och istället inför ett övermått av chipsätande? En vetenskaplig studie fick stor uppmärksamhet i media där vinklingen ofta gav intrycket att tv-tittande är farligt i sig. Dagens Nyheters rubriksättare valde rubriken "Tv-tittande kan orsaka för tidig död" (15 juni). I Science Daily (också 15 juni) säger en av författarna till studien, Frank Hu, att "The message is simple. Cutting back on TV watching can significantly reduce risk of type 2 diabetes, heart disease, and premature mortality." Men hur är det egentligen? Vi frågade medförfattaren Anders Grøntved. »»»»

»»» PROLONGED TV VIEWING AND RISK OF CHRONIC DISEASE AND PREMATURE DEATH: IS THE RELATIONSHIP CAUSAL?

We have recently published an article quantitatively summarizing all available cohort studies on the association of television

viewing with the risk of type 2 diabetes, cardiovascular disease and all-cause mortality [1]. The results from these analyses showed that prolonged TV viewing was associated with increased risk of these health outcomes in a dose-response manner. Now that evidence from large cohort studies has confirmed to be robust across different populations from different parts of the world, can it be accepted as causal? To answer this in more detail, a number of important questions must be answered related to what drives the association with TV viewing: i) Is the association due to confounding or other biases? ii) Is it the daily time spent sitting that is harmful? iii) Is it other risk behaviors performed while viewing TV or risk behaviors acquired from viewing TV? Answering these questions is crucial for designing public health messages and interventions tailored to reduce risk associated with prolonged TV viewing time.

All associations in individual studies were statistically adjusted for measures of physical activity and associations persisted. Thus, it is unlikely that the associations are completely confounded by physical activity. However, physical activity information is obtained from questionnaires in large scale epidemiological studies and these are often not designed to capture light intensity activity such as activities of daily living. Therefore, it is likely that the effects of prolonged TV viewing are mediated by light activity as TV viewing may displace time spent on activities at home such as cooking, cleaning, and other standing/walking activities rather than time spent on exercise such as running and bicycling. Other putative confounding factors such as smoking, diet, and socioeconomic position were already controlled for in most of the available studies; nevertheless, unknown or residual confounding may still exist.

Because previous studies on the associations of sedentary behaviors similar to TV viewing or non-specific sedentary time with chronic disease and mortality have been reported in cohort studies (see references [2]-[6]), this gives some confidence that sedentariness per se is an



important part of the harmful effect of prolonged TV viewing. Prolonged daily sitting while e.g. reading may therefore also be an independent risk factor as it could displace time spent on daily activities with light intensity at home or other places. A previous cohort study among US nurses supports that other sitting time at home beyond TV viewing is positively associated with type 2 diabetes risk, whereas time spent walking and standing at home is protective [3].

Other unhealthy lifestyle behaviors could be picked up or modified by daily exposure to prolonged TV viewing. These include eating more unhealthy food, increasing or initiating alcohol drinking and smoking during and beyond TV viewing time. Evidence from experimental studies which, strictly speaking, is the only study design that can refute or accept hypotheses regarding lifestyle factors to be causal and not a result of bias such as confounding, has provided some support of these inferences. Two randomized trials among children have shown that reducing TV viewing time can lead to favorable changes in adiposity status (i.e. being obese) at least by altering dietary patterns [7,8]. Another randomized trial among overweight or obese adults did not find a statistically significant change in adiposity status from restricting TV viewing time, but did see an increase in objectively measured

energy expenditure [9]. The limited available evidence from these trials suggests that reducing TV viewing can lead to improvement in energy balance which is a product of energy intake and expenditure. Non-experimental evidence also suggests that prolonged TV viewing is associated with initiation of smoking [10]. Together this indicates that TV viewing to some degree may influence dietary intake and smoking during and beyond TV viewing time.

Prolonged TV viewing is a very common daily activity among both children and adults in western countries. It may be so pervasive that it impairs health for all of the above mentioned reasons. Important future studies to conclusively test the causal role of TV viewing in influencing health should come from larger experimental studies randomly allocating both healthy participants and individuals at high risk of chronic disease to restricted TV viewing time or control conditions. Outcomes in such studies could be biomarkers closely associated with chronic disease risk or if possible hard endpoints. Adherence to restricting TV viewing time among participants is an important feature in such a study and devices electronically monitoring TV viewing behavior have previously been used with success among children and adults [8,9]. Thus, such studies should be feasible to carry out. Furthermore, future studies should attempt to disentangle the importance of TV viewing in altering total daily sitting time, dietary intake, and smoking with greater detail to gain insight into what elements of prolonged TV viewing that cause disease and premature mortality. Finally, more investigation in other daily sitting time apart from TV viewing and in engagement in light intensity activity is warranted as many individuals are being sedentary for a vast amount of time every day regardless of the amount of TV viewing.



ANDERS GRØNTVED,
SYDDANSK UNIVERSITET
AND HARVARD SCHOOL
OF PUBLIC HEALTH, USA

Referenser:

- [1] Grøntved A, Hu FB. Television Viewing and Risk of Type 2 Diabetes, Cardiovascular Disease, and All-Cause Mortality. *Journal of the American Medical Association*. 2011;305:2448-2455.
- [2] Manson JE, Greenland P, LaCroix AZ et al. Walking Compared with Vigorous Exercise for the Prevention of Cardiovascular Events in Women. *New England Journal of Medicine*. 2002;347:716-725.
- [3] Hu FB, Li TY, Colditz GA, Willett WC, Manson JE. Television Watching and Other Sedentary Behaviors in Relation to Risk of Obesity and Type 2 Diabetes Mellitus in Women. *Journal of the American Medical Association*. 2003;289:1785-1791.
- [4] Katzmarzyk PT, Church TS, Craig CL, Bouchard C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Medicine and Science in Sports Exercise*. 2009;41:998-1005.
- [5] Weller I, Corey P. The Impact of Excluding Non-Leisure Energy Expenditure on the Relation between Physical Activity and Mortality in Women. *Epidemiology*. 1998;9:632-635.
- [6] Patel AV, Bernstein L, DeKa A et al. Leisure Time Spent Sitting in Relation to Total Mortality in a Prospective Cohort of US Adults. *American Journal of Epidemiology*. 2010;172:419-429.
- [7] Robinson TN. Reducing Children's Television Viewing to Prevent Obesity. *Journal of the American Medical Association*. 1999;282:1561-1567.
- [8] Epstein LH, Roemmich JN, Robinson JL et al. A Randomized Trial of the Effects of Reducing Television Viewing and Computer Use on Body Mass Index in Young Children. *Archives of Pediatrics and Adolescent Medicine*. 2008;162:239-245.
- [9] Otten JJ, Jones KE, Littenberg B, Harvey-Berino J. Effects of Television Viewing Reduction on Energy Intake and Expenditure in Overweight and Obese Adults: A Randomized Controlled Trial. *Archives of Internal Medicine*. 2009;169:2109-2115.
- [10] Gidwani PP, Sobol A, DeJong W, Perrin JM, Gortmaker SL. Television Viewing and Initiation of Smoking Among Youth. *Pediatrics*. 2002;110:505-508.

CAUSAL INFERENCE – A BOOK REVIEW

■ **Statistical Models and Causal Inference: A Dialogue with the Social Sciences (Cambridge University Press), by David A Freedman**

The term causal inference refers usually to the estimation of effects of treatments or the finding of the causal relationships between (random) events. Causal inference is a hot topic in the statistical literature nowadays mainly due to, among other reasons, negligence of the topic by the research community in the past. The most solid basis of causal inference is randomized control experiments. However, we often find situations where we cannot perform them, perhaps due to ethical reasons or inability to go to the past; therefore we need to perform the task with observed data perhaps along with some subject domain knowledge. There are approaches for this: the Markov graphical modeling framework that is based on probabilistic relationships between events, thus variables, using (conditional) (in)dependencies, the Neyman–Rubin potential outcome framework that is based on counterfactual outcomes, structural equation modeling, and so on.

Definitions of probability

This book, a timely publication in this form, is a collection of some works particularly in social science by David Freedman and co-workers, spanning over 20 years starting in the early 90s. However, many of the applied issues presented are related to the medical sciences despite the title of the book. After reading the book, one may be convinced that causal questions cannot be solved

completely by statistical methods alone, but they can be supportive to the logical reasoning and knowledge in the subject domain, which implies that statisticians need to work with subject domain experts in the enterprise of causal inference.

Freedman starts by talking about various definitions of probability and shows how they can lead to trouble when trying to interpret probability in different contexts; the objectivist idea of existence of the probability of an event separate from data, the simplicity of empirical relative frequency, incompleteness or partial rationality of the Bayesian way of thinking about probability, unavailability of stable prior probabilities, and so on. All are of importance to applied scientists. He goes on to argue that traditional statistical models such as regression models are difficult to apply due to non-testable assumptions and thus they become an integral part of modeling when used and modest departure from assumptions may cause substantial changes in results due to bias.

He also points out how problematic it is to translate applied research questions into statistical hypotheses and related distributional assumptions. Freedman's opinion is that regression models that are applied for causal inferences through the assessment of statistical significance of parameters are less helpful in doing the job, and further they provide less help for controlling confounding. He shows that inferring causal relationships from observed data by fitting regression models is not successful unless we have considerable prior knowledge on the data generating mechanism. Furthermore, adjustments do not often fix the problems with the regression models. His stance is that other popular causal models such as Markov graphical models

also have limited successful practical applications where their causal interpretation can be defended. In one of the articles he criticizes some of the graphical modeling researchers who are regarded as successful in formulating and applying their theory in social sciences. Though Freedman is also very critical about, for example, assumptions on unobserved errors he is saying almost nothing against probability causal models such as the Neyman–Rubin model that contains probability statements on counterfactual events that are also not testable.

Questions the statistical developments

This book provides not only applied researchers with a lot to consider when they apply traditional statistical tools to their research questions but also statistical methodologists to rethink their tools for applied questions; how model predictions can go wrong when required conditions are not met and how mathematical convenience often leads to failures in applications. Thus, applied researchers are left wondering about their need of clear statistical guidance in order to proceed with their work. However the book does not provide any such guidance. One of the papers starts by saying “*This chapter gives an informal account ...*”. The paper is mathematically discouraging. Most of the contents in the book question the statistical methodological developments we see today, but, of course, favouring some models such as the Neyman–Rubin model of causation. This book should ideally be read along with the author's earlier book *Statistical Models: Theory and Practice* (Cambridge University Press).

PRIYANTHA WIJAYATUNGA,
UMEÅ UNIVERSITET

Några funderingar om kunskapsteori

Som statistiker vill man gärna tro att statistik och statistiska metoder spelar en roll i den kunskapsgenererande processen.

Därför kan det vara på sin plats att fundera lite över vilka kunskapsteoretiska uppfattningar som kan förknippas med denna vår vetenskap.

Sanning

Sanning kan definieras på flera sätt. Aristoteles föreslog korrespondensteorin som säger att det är sant som överensstämmer med de verkliga förhållandena. En intuitivt tilltalande teori som dock råkar i svårigheter när även själva överensstämmelsen skärskådas med just detta sanningsbegrepp. Ett annat synsätt är koherensteorin, vars krav på motsägelsefrihet ofta uppfattas som ett nödvändigt men sällan tillräckligt villkor för epitetet sanning. Men notera att inte ens den gamla idealvetenskapen fysik är helt koherent. Pragmatismen laborerade med praktisk användbarhet eller rent av majoritetsbeslut som sanningskriterier. Det förra bekvämt för en diktator, det senare ett bra sätt att göra sanningar av värderingar. Evidensteorin, som kräver belägg för att något ska få hållas för sant, är kanske den till synes mest vetenskapliga definitionen. Men med positivismens död och en huvudsakligen modellbaserad vetenskap, är begrepp som bevis, belägg och verifierbarhet inte längre helt absoluta.

Så långt alltså inget sanningsbegrepp som kan köpas rakt av. I praktiken får man välja och vraka lite. Intuitiv överensstämmelse med verkligheten och motsägelsefrihet är inte så dumt. Inte heller litet evidens, vad det nu betyder. En smula inte alltför relativistisk pragmatism kanske inte heller skadar.

Sedan finns ju alltid den ontologiska utvägen att förhålla sig skeptisk till sanningens blotta existens. Men intar man den positionen ska man nog inte syssla med vare sig statistik eller kunskapsteori.

Kunskap

Kunskap brukar traditionellt definieras som sann, berättigad tro. För att kunna säga att



Charles Sanders Peirce, 1839-1914, vände på deduktionen och formulerade en princip för abduktion.

man vet något ska man alltså tro på det och ha goda skäl för denna tro. Dessutom ska det vara sant. Det senare gör att kunskapen hänger intimt samman med sanningen och den teori man har för denna. Sanningsledet är också den problematiska delen i dessa kunskapens krav. Det går visserligen att konstruera situationer där sanningskravet är uppfyllt men där tron eller berättigandet sviktar. Som exempelvis när Gettier på 1960-talet visade att inte alla former av berättiganden duger. Men det handlar som sagt om konstruktioner och i de flesta praktiska fall fungerar sann, berättigad tro som tillräckliga och nödvändiga kunskapskriterier.

Även här finns möjligheter till olika grader av skepticism. Oberoende av om man bejakar existensen av en sanning eller inte, så kan man förneka möjligheten att nå denna eventuella sanning. Men med en sådan kunskapsteoretisk inställning är det föga meningsfullt att över huvud taget bedriva vetenskaplig verksamhet.

Deduktion, induktion och abduktion

Ibland sägs det att kunskapsteoriens främsta mål är att vederlägga den skeptiska inställningen till möjligheten att erhålla kunskap om världen. Utmärkta argument i den diskussionen är givetvis att påtagligt anvisa vägar till just kunskap.

Historiskt sett brukar man skilja på deduktion och induktion som framkomliga vägar till den hägrande kunskapen. Ser man till hur det verkligen går till kan man kanske också prata om abduktion.

Deduktionen är logiskt tämligen oantastbar. Från de aristoteliska syllogismerna och framåt har man tillåtit sig att från allmänna premisser dra slutsatser om det speciella fallet; om A medför B och A föreligger så följer att även B är vid handen. Inga observationer av verkligheten krävs: kunskapen om B kan resoneras fram om man accepterar A. En förespråkare för denna rationalism var Descartes. Eftersom han inte litade på sina sinnesintryck, sökte han sig tillbaka till den ovedersägliga sanningen *cogito ergo sum*, jag tänker därför finns jag, för att därifrån med deduktionens hjälp vinna kunskap om världen. Att han förmodligen uttryckte det som *je pense donc je suis*, Descartes gillade inte latin, och att hans sanning egentligen inte är speciellt ovedersäglig, är en annan historia.

En helt annan position intog de engelska empiristerna. Här var metoden att göra faktiska observationer av verkligheten och sedan, när man tyckte att man observerat tillräckligt mycket, dra slutsatser om de allmänna förhållandena. Detta sätt att resonera, induktion, är givetvis inte logiskt vattentätt. Så länge man inte observerat alla förekomster av ett visst fenomen, så kan det ju alltid dyka upp ett fall som avviker från alla andra. Men detta var inget man hymlade med och en av empiristerna, Hume, tillskrivs formulerandet av detta induktionsproblem. Ett känt exempel på hur fel det kan bli var färgen på svanar. Eftersom samtliga hittills observerade

statistik och

svanar varit vita, så drog man slutsatsen att alla svanar i världen var vita. När man sedermera upptäckte Australien visade sig denna ö, eller var det möjligen en kontinent, bebos av bland annat svarta svanar.

Ytterligare en väg till kunskap anvisades i början på förra seklet av Charles Sanders Peirce. Det han benämnde abduktion vänder på deduktionen och drar slutsatser om premissen. Om det är känt att A medför B och man observerar B, så drar man slutsatsen att även A föreligger. Detta resonemang är inte logiskt invändningsfritt, men Peirce menade att eftersom det är så man söker efter förklaringar i det dagliga livet, så kunde abduktionen även få ha en roll i det vetenskapliga arbetet. Problemet är givetvis att fenomenet B kan ha andra orsaker än A; även om man anser att just A är den troligaste. Ett nutida exempel på abduktionens risker är den svenska marinens jakt på främmande ubåtar. Man visste att ryska ubåtar (A) gav upphov till ett visst ljud (B), A medför B alltså. När man sedan observerade detta ljud drog man enligt abduktionens princip slutsatsen att även A, dvs att ryska ubåtar fanns i närheten. Tyvärr visade det sig senare att B kunde ha även andra förklaringar. I detta fall minskar som var på rymmen. Och det var inte ens ryska utan troligen svenska minskar.

Sannolikhetsteori och statistisk inferens

Om man, efter Kolmogorov, betraktar sannolikhetsteorin som blott en konstruktion med axiom och teorem blir den dels analytisk, d v s oberoende av verkligheten, dels deduktiv. Med sannolikhetsteoretiska resonemang kan man utifrån allmänna förhållanden eller premisser uttala sig om vad som kan tänkas hända i det enskilda fallet. Dock bara i termer av sannolikheter för olika utfall. De fördelar och nackdelar som vidlader rationalismen och deduktionen i allmänhet drabbar också sannolikhetsteorin. Slutsatserna blir logiskt hållbara men resonemanget tillför inget nytt eftersom all kunskap redan finns i premisserna. Dock, att dra ut konsekvenserna, i stokastiska eller deterministiska termer, av givna antaganden eller enligt föreslagna modeller är en väsentlig del av det vetenskapliga arbetet.

Den statistiska inferensen är tvärtom induktiv till sin natur. Här går man ju från det speciella

till det allmänna. Baserat på ett urval eller en försöksserie drar man slutsatser om en bakomliggande population eller process. Dock inte på det naiva sätt som möjliggjorde utsagan om att alla svanar är vita. Det statistiska påståendet skulle istället, givet ett slumpmässigt urval av observerade svanar, säga att andelen svanar med vit fjäderdräkt ligger mellan exempelvis 92 och 100 procent.

Statistisk inferens skulle också kunna betraktas som ett fall av abduktion. Denna form av slutledning kallas ibland för inferens till den troligaste förklaringen och exempelvis en *maximum-likelihood*-skattning är just en sådan optimering i överensstämmelsen mellan estimat och parameter. Med de till minskar förvandlade ubåtarna i färskt minne låter detta måhända litet skakigt. Men oavsett om man ser den statistiska inferensen som induktion eller abduktion, så vilar den på en solid sannolikhetsteoretisk grund och dess utsagor beledsagas alltid av sådant som medelfel, konfidensintervall eller p-värden. Eller hur man nu väljer att uttrycka den statistiska kvalitets säkringen.

Statistik och kunskap

Den statistiska vetenskapen har alltså båda deduktiva och induktiva inslag. Och detta gäller även för vetenskapen i stort. Om man bortser från de mer tolkande humanistiska disciplinerna, så bedrivs mycket forskning enligt den hypotetiskt-deduktiva metod som bl.a. Karl Popper använde som vetenskaplig modell. Litet schematiskt uttryckt tänker man sig framkastandet av teorier och modeller, av mer eller mindre induktivt ursprung, vars konsekvenser och förklaringskraft testas med rent deduktiva metoder. För dessa olika faser i det vetenskapliga arbetet använde Popper uttrycken *context of discovery* respektive *context of justification*.

I upptäcktsfasen är alla medel tillåtna. Här kan fantasin flöda när det gäller att hitta förklaringsmodeller till olika observerade fenomen. Induktion, abduktion och mer eller mindre galna infall i en salig blandning alltså. Det finns exempelvis en anekdot om hur Kekulé i ett lätt omtöcknat tillstånd fick idén till strukturformeln för bensen genom att titta på rökslingor som likt ormar bet sig i svansen och bildade ringar. Även om just denna idé visade sig hålla, är man i denna fas

långt ifrån den vederhäftiga vetenskapliga teori som kan delges omvärlden.

För att idéerna eller de tentativa teorierna ovan skall få namn av vetenskap, krävs det en berättigandefas med tester och utvärderingar utformade på ett sådant sätt att eventuella felaktigheter avslöjas. Den teori som klarar sig utan att förkastas får fortsatt förtroende som förklaringsmodell för de aktuella fenomenen. Men, nota bene, bara så länge som inga anomalier upptäcks. Den förhandenvarande teorin utgör nämligen inte nödvändigtvis den slutliga sanningen; bara en vetenskapens nuvarande ståndpunkt. Som mycket väl kan komma att utvecklas och revideras när nya rön framläggs.

I förstone kan det förefalla som om statistiska metoder vore förbehållet den senare fasen. Här känner man ju igen formulerandet av strikta hypoteser och tester med kontrollerade felrisker. Men även i den tidigare fasen finns det utrymme för statistik i olika former. Även om kraven på stringens där är betydligt lägre och man kanske rent av sysslar med multipel inferens som närmar sig så kallat signifikansfiske, så är det inga större problem så länge man har klart för sig var i den vetenskapliga processen man befinner sig. Och tolkar resultaten, även de statistiska, utifrån denna insikt.

Statistisk kunskapsteori

Det är förvisso förmätet att kalla de spretiga funderingarna ovan för statistisk kunskapsteori. Men möjligen borde man fundera mer över vilken roll som statistisk teori och metod spelar i det som ibland kallas den kunskapsgenererande processen. Och om man dessutom inkluderar ett inomvetenskapligt perspektiv kanske man kan skapa litet balans i den vetenskapens värld där statistik ömsom övertolkas och betraktas som evig sanning, ömsom misstros och inte alls kommer till användning.



TOMMY JOHNSSON,
GÖTEBORGS UNIVERSITET

Joakim Westerlund skriver här första delen om sin forskning i ett kontroversiellt område. I början var det spännande, och lovande, men hur blev det sen? Och varför är det så att de första resultaten är lovande, sen dalar det utför – tills nästa omgång resultat kommer, som är lovande...

Mina äventyr i parapsykologins förtrollade värld

DEL 1. VAD ÄR PARAPSYKOLOGI OCH VILKA ÄR FORSKARNA?

Under lite drygt tio år bedrev jag parapsykologisk forskning. Det var en spännande och på många sätt lärorik tid. För några år sedan tröttnade jag dock på området.

Parapsykologin har sina rötter i spiritismen som frodades kring förra sekelskiftet och än i dag finns några få parapsykologer som studerar fenomen som på ett eller annat sätt implicerar att "själen" kan överleva döden: medier som påstår sig kunna kommunicera med avlidna, spöken, reinkarnation, poltergeistfenomen och liknande. Men de allra flesta som bedriver parapsykologisk forskning nuförtiden undersöker det de kallar psi-fenomen. Dessa är indelade i två grupper; extrasensorisk perception (ESP) och psykokenesi, se figuren på nästa sida.

Extrasensorisk perception är tanken att vi har ett "sjätte sinne" med vilket vi kan få information på (än så länge) okänd väg och innefattar de tänkta fenomenen telepati (tankeöverföring), klärvoajans (fjärrsyn) och prekognition (att se in i framtiden). Vad dessa fenomen egentligen innebär kan bäst förstås med exempel.

Experiment om det "sjätte sinnet"

Telepati. I ett rum finns en dator och en person som vi kan kalla sändare. I datorn finns ett antal bilder som bedömts vara positiva respektive negativa till sin emotionella karaktär. En slumpmekanism avgör var tjugonde sekund med lika sannolikhet om en positiv eller negativ bild skall visas för sändaren. Samtidigt som bilden börjar visas för sändaren blinkar en lampa till i ett annat rum där det finns en person som vi kan kalla mottagare. Mottagaren gissar då om bilden som visas för sändaren är positiv eller

negativ. Om mottagaren gissar signifikant bättre än 50% rätt anses försöket vara lyckat (Sven Ove Hansson diskuterade denna ansats i Qvintensen nr 3/2010).

Klärvoajans. Ett klärvoajansexperiment skulle kunna gå till på precis samma sätt som telepatiexperimentet som skisserats ovan med den enda skillnaden att datorn visar bilderna i ett tomt rum utan mänsklig sändare.

Prekognition. Experiment kan göras både med och utan sändare och sändaren kan dessutom vara samma person som mottagaren. Ett experiment om prekognition kan använda samma dator och samma bilder som i telepatiexperimentet ovan. Experimentet börjar 20 sekunder innan första bilden visas med att försökspersonen gissar om första bilden kommer att vara positiv eller negativ. Experimentet fortsätter på samma sätt för resten av bilderna.

Experiment om psykokenesi...

Mikropsykokenesi. Psykokinesi är idén att vi enbart med "tankekraft" kan påverka vår omgivning. Vår gamla dator med bilder från telepatiexperimentet kan fortfarande användas. Försökspersonen sätts framför datorn och uppmanas att med tankekraft försöka få datorn att visa till exempel så många positiva bilder som möjligt. Efter ett tag växlar man och försökspersonen ska i stället försöka få fram så många negativa bilder som möjligt.

Makropsykokenesi. Försökspersonen uppmanas att med tankekraft försöka få datorn att lyfta från bordet! Denna typ av försök förekommer inte längre.

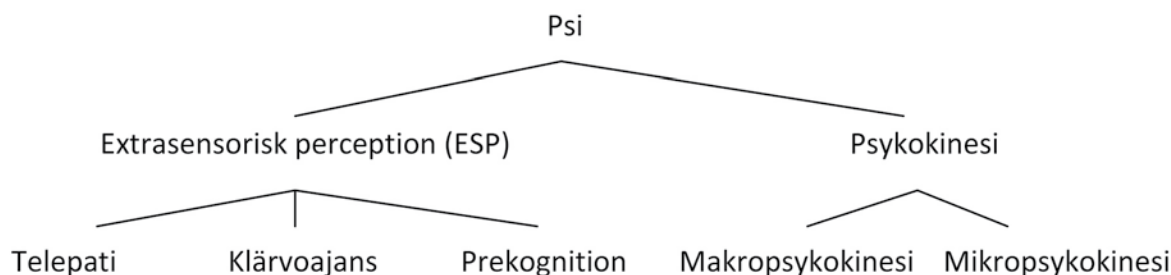
Är fenomenen lika konstiga?

Om vi bortser från det empiriska bevisläget inom parapsykologin och från de anekdoter som vi alla hört, men som ju enklare kan förklaras utifrån psykologiska mekanismer som till exempel oförmåga att förstå hur slumpen kan slå, hur tokiga är de här sakerna egentligen? Jag vill påstå att alla sakerna inte är lika tokiga. Framförallt behöver telepati inte nödvändigtvis innebära någon radikalt annorlunda naturvetenskaplig syn på världen. Positiva resultat i många telepatiexperiment skulle till exempel kunna förklaras av att vi, liksom många andra djurarter, har något slags elsinne och att vi med detta elsinne skulle kunna reagera olika beroende på om mänskliga hjärtan i vår närhet slår snabbt eller långsamt. De andra fenomenen, i synnerhet makropsykokenesi och prekognition, är betydligt svårare att tänka sig utan att samtidigt behöva göra drastiska förändringar i den naturvetenskapliga världsbilden.

Vilka håller på med parapsykologi?

För det första måste sägas att det inte är särskilt många som bedriver parapsykologisk forskning. Konferenser som anordnas av the Parapsychological Association hålls varje år, omväxlande i Europa och Amerika, vanligen USA. På dessa konferenser kommer ungefär femtio personer som själva bedriver parapsykologisk forskning vilket jag skulle tro är ungefär hälften av alla parapsykologer som finns.

I Europa kommer de flesta parapsykologer från Storbritannien, mest från University of Northampton och University of Edinburgh, och från Tyskland där det finns ett privat institut i



Psi-fenomen är uppdelade i två grupper; extrasensorisk perception (ESP) och psykikinesi.



Parapsykologin har sina rötter i spiritismen från förra sekelskiftet.

Freiburg med ett tjugotal anställda. I Storbritannien har drygt ett dussin personer doktorerat i parapsykologi på olika psykologiska institutioner, de flesta på University of Edinburgh under nyligen avlidne Robert Morris ledning. Nästan alla dessa har efter disputationen fått anställning på brittiska universitet, ofta som lärare i forskningsmetod och statistik. Robert Morris innehade 1985-2004 den då enda professuren i parapsykologi, möjliggjord tack vare testamenterade pengar från författaren Arthur Koestler. Robert Morris var mycket uppskattad, inte bara bland parapsykologins tillskyndare utan även bland dess kritiker. Fortfarande finns endast en professur i parapsykologi i hela världen. Den finns i Lund och innehas av Etzel Cardena. Även denna tjänst har finansierats genom privat donation. I samband med tillträddandet uttalade sig Cardena mycket onyanserat i Aftonbladet om bevisläget inom parapsykologin och jag och min kollega Jan Dalkvist kritiserade honom för dessa uttalanden i Sydsvenska Dag-

bladet (Aftonbladet 7 juni 2005 och en serie replikskiften i Sydsvenskan med start 17 juni 2005).

De som forskar inom parapsykologi brukar vara medlemmar i the Parapsychological Association som sorterar under the American Psychological Association som i sin tur sorterar under the American Association for the Advancement of Science. Detta gör att tidskriften *Journal of Parapsychology* finns på de flesta stora universitetsbibliotek. De parapsykologiska tidskrifterna använder sig (förstås) av peer-review. Personer med en mycket kritisk inställning till uppfattningen att det skulle finnas några parapsykologiska fenomen finns ibland representerade både som artikelförfattare och granskare av inskickade manuskript. Artiklarna redovisar oftast utförda enskilda experiment men ibland görs även metaanalyser varav några publicerats i vanliga (icke-parapsykologiska) psykologiska tidskrifter. Så på många sätt bedrivs parapsykologisk forskning på precis samma sätt som forskning inom mer etablerade och accepterade discipliner.

Många parapsykologer som undersöker psi-fenomen har nog en substansdualistisk syn på kropp-och-själproblemet och verkar tro att om man bara kan visa att psi-fenomen är reella så har man därmed bevisat att människan verkligen har en slags mystisk själ – en ande i maskinen – som under vissa omständigheter kan lämna kroppen och, utan öron och utan ögon, höra och se vad som sker i världen och som så klart kan överleva döden. De drivs alltså av ett slags religiöst motiv i sin forskning. Detta behöver förstås inte hindra att de ändå kan producera bra forskning inom ämnet. Men för somliga parapsykologer tror jag att de "religiösa" motiven hänger ihop med stark psykologisk uppbundenhet

kring föreställningen att parapsykologiska fenomen verkligen finns. Den parapsykologiska forskningen som de bedriver blir då en slags plikt, något som man måste göra – av sociala och rituella skäl – men vars resultat, hur de än blir, inte på något sätt kommer att påverka deras

uppfattning om något som de ändå redan "vet" är verkligt. Dessa parapsykologer tror jag ligger i riskzonen att lura sig själva och att glömma bort obekväma resultat och argument.

Men en andel av parapsykologerna drivs verkligen av ett behov av att få reda på sanningen. De har ofta hamnat inom området av en tillfällighet.

Vad har man kommit fram till?

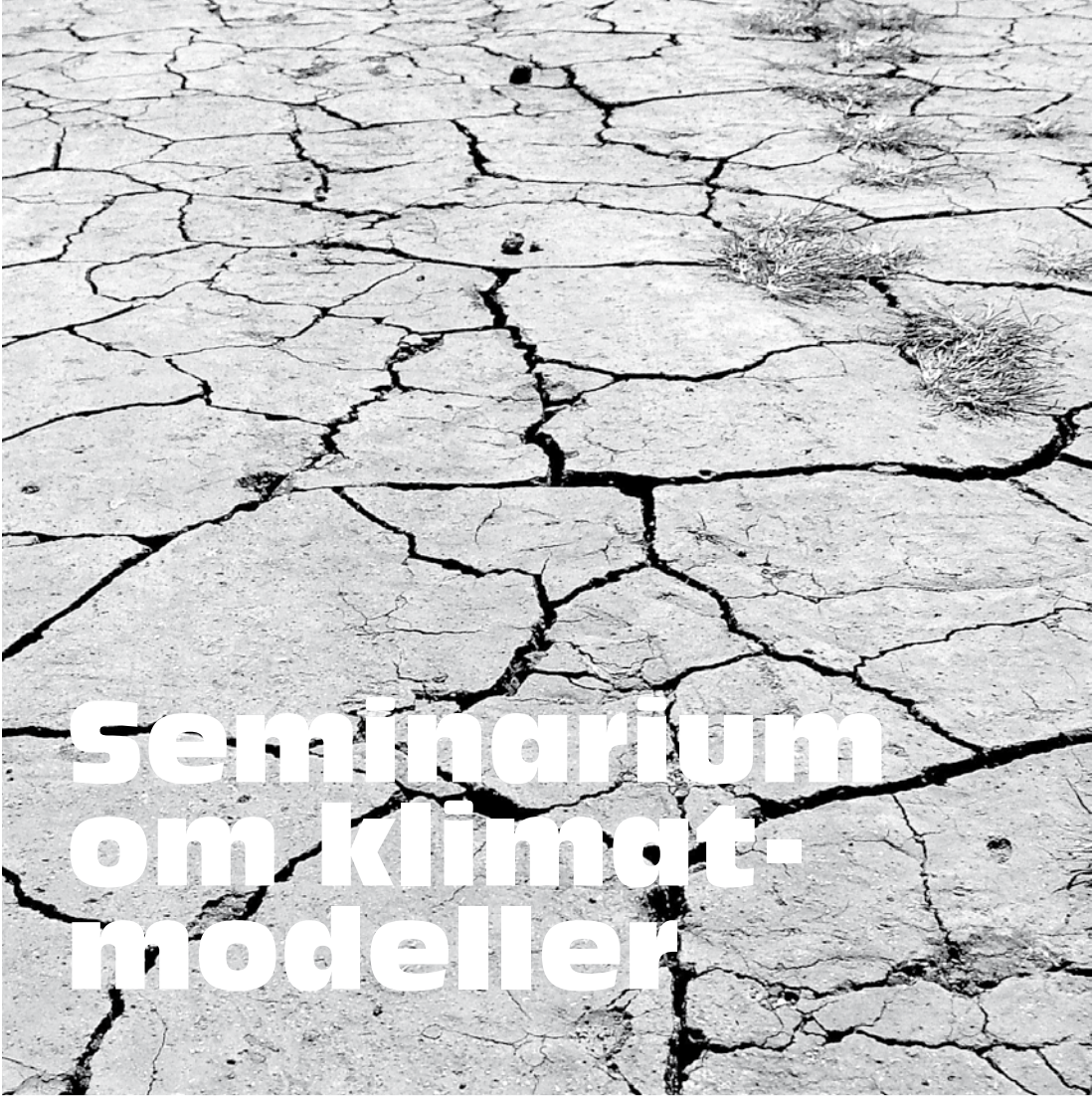
Experimentell forskning om psi-fenomen har bedrivits i ungefär 80 år. Visserligen har inte särskilt många personer bidragit till denna forskning, men totalt sett är det ändå en ganska stor mängd experiment som gjorts och som sammantagna ändå borde kunna säga oss något om huvudfrågan: har vi ett sjätte sinne?

Jag tror att många forskare utanför parapsykologins värld skulle bli förvånade om de tog sig tid att läsa igenom något nummer av *Journal of Parapsychology* från den senaste tioårsperioden eller ta del av föredragen på Parapsychological Associations årliga konferenser. De skulle då, åtminstone vid en första anblick, se att de flesta av de undersökningar som rapporteras inte verkar innehålla några felkällor. I stället verkar författarna vara minst lika medvetna om metodologiska och statistiska problem som forskare inom till exempel psykologi. Dessutom skulle de se att de allra flesta artiklarna presenterar signifikanta resultat med hyfsade effekttstyrkor som är väldigt svåra att förstå hur de har uppkommit givet att det inte finns några paranormala fenomen. Det var i alla fall så jag själv reagerade när jag började sätta mig in i den parapsykologiska forskningslitteraturen.

Men efter hand upptäckte jag att en viss typ av händelseförlopp i den kollektiva parapsykologiska forskningsprocessen återkom och upprepade sig genom parapsykologins 80-åriga historia. I Del 2 fortsätter jag min diskussion av parapsykologisk forskning.

JOAKIM WESTERLUND,
STOCKHOLMS UNIVERSITET

**I nästa nummer
fortsätter Joakim
med slutet
på berättelsen...**



Seminarium om klimat- modeller

På vilka sätt kan statistiska metoder utvecklas och användas i klimatforskningen? Med exempel från just klimatforskning gav Olle Häggström i Qvintensen nr 1/2011 goda argument för att statistikämnenas behöver flytta fram sina positioner inom den högre utbildningen i allmänhet. Vi ger här en inblick i ett viktigt och svårt forskningsområde där stora framsteg i våra kunskaper görs.

■ **Ett område inom klimatforskningen** där det är nödvändigt att använda statistiska metoder och tänkesätt är skalningsproblem, dvs hur man skall koppla samman globala klimatmodeller som beskriver det riktigt storskaliga klimatet med modeller som beskriver klimatet på regional eller lokal nivå. Denna sammankoppling innebär både att man utnyttjar data för att finna statistiska samband och att man kombinerar fysikaliskt grundade deterministiska modeller med stokastiska modeller. Ett internationellt möte, som hölls i Lund strax före "siste april", gav en översikt över en del av problematiken. Mer än femtio personer hade samlats för att diskutera "statistical approaches to down- and upscaling in climate models", som var mötetets titel.

På seminariet togs nuvarande praxis för användning av modeller för framtida klimat och klimateffekter och för rekonstruktion av gångna tiders klimat upp, liksom presentationer av nya statistiska metoder för att lösa kända och

okända (!) problem. De många deltagande doktoranderna från Institutionen för geo- och ekosystemvetenskaper och från Ekologiska institutionen i Lund vittnar om intresset för mer samarbete med statistikämnenas.

Peter Guttorp delar sin tid mellan att vara statistikprofessor vid University of Washington i Seattle och att från Norsk Regnesentral i Oslo leda det nordiska toppforskningsnätverket SARMA. Tillsammans med Matematisk statistik i Lund och de strategiska forskningsinitiativen MERGE och BECC stod SARMA för programmet. Markku Rummukainen, professor i klimatmodellering vid Lunds universitet, inledde med en översikt över de väsentligaste problemen. Därefter följde föredrag av Shinya Sugita, Carl XVI Gustavs gästprofessor i miljövetenskap, och ytterligare sexton föredrag och åtta posterpresentationer.

GEORG LINDGREN,
LUNDS UNIVERSITET

Mer om SARMA-nätverket kan man läsa på: www.nrcse.washington.edu/NordicNetwork/

Upp- och nedskalning av klimatmodeller

Alla föredrag finns tillgängliga på www.maths.lth.se/matstat/climate/Presentations.html, och vi refererar till dem i det följande med namnet på talaren inom parentes.

Behovet av nedskalning

Regionala effekter av ändrat klimat måste beräknas med större upplösning än den som används i de globala modellerna. Detta har lett till utveckling av regionala modeller, dvs modeller som löser samma ekvationer som de globala modellerna men på finare rutnät (areell indelning). Då kan man använda en mer detaljerad beskrivning av jordytan, t ex bergighet och landanvändning.

Många viktiga klimatprocesser har rumslig utsträckning som är mindre än rutnätet på vilket en klimatmodell löser de partiella differentialekvationer som beskriver atmosfären och havet. Globala klimatmodellers rutnät kan handla om en upplösning på någonstans mellan 100 och 300 km, medan regionala klimatmodellers rutor är av storleken 10-50 km. De olika dimensionerna påverkar bl a hur man kan hantera de effekter moln har på klimatet. Moln reflekterar inkommande solstrålning och ger en kyleffekt vid jordytan men de hindrar också en del av värmestrålningen från jordytan att transporteras uppåt i atmosfären och ger därför en värmande effekt vid jordytan. Men moln och även molnsystem är ofta små jämfört med rutorna i klimatmodellerna. Dessutom har modellerna en vertikal upplösning (rutnät på höjden) som varierar från ett hundratal meter närmast marken till någon kilometer högre upp i troposfären. Hur mycket molnlagren på olika höjder täcker varandra avgör både uppvärmnings- och avkylningseffekter men kan inte direkt beskrivas i de flesta globala klimatmodellerna (Guttorp). Istället används olika typer av approximationer och medelvärden.

Beskrivningen av hydrologiska processer behöver detaljerade nederbördssimuleringar på olika skalor i rum och tid, från enstaka kvadratkilometer till hela flodområden, från

minuter till decennier. Klimatmodellerna är oanvändbara för hydrologisk planering om man inte tar hänsyn till detaljerade nederbördsdata (Uvo). Resultat från klimatmodeller behöver först bearbetas och kombineras med empiriska modeller.

Processer i vegetation och mark varierar stort över jorden, t ex ljusabsorbering och fotosyntes i vegetationen eller syreanvändning och kvävebindning i marken. Jordsystemmodeller är den senaste generationen av klimatmodeller. De kräver statistiska metoder som tillåter flexibel modellering av sådana processer (Anderson, Shurgers).

Nedskalningsmetoder

Regionala klimatmodeller kallas även *dynamisk nedskalning* eftersom de använder randvärden från en global modell i ett finare rutnät än det som finns i randvärdesmodellen. Så fort värdena i den globala modellen ändras, justeras också den regionala modellen. Men också i regionala modeller måste viktiga processer parametreras på grund av att deras rutnät fortfarande är grövre än det som karakteriserar många klimatprocesser (Rummukainen).

Statistisk nedskalning är måhända den metod som känns mest hemtam för en statistiker. Metoden är ett verktyg för regionalisering av globala eller andra storskaliga klimatdata. Den baseras på någon typ av statistisk modell för hur processer som försiggår på liten skala förhåller sig till, exempelvis, storskalig cirkulation. Vanliga metoder är empiriska ortogonala funktioner som är en sorts regressionsmodeller där antalet dimensioner i mångdimensionella observerade storheter reduceras så att man får en enklare version av den regionala modellen (Benestad, Bolin). Metoden är en speciell typ av principalkomponentanalys.

En annan metod för nedskalning är *stokastisk nedskalning* där en fysikaliskt grundad modell kompletteras med en stokastisk modell för sådana processer som inte går att modellera i detalj med en deterministisk modell, t ex nederbörd eller moln. Man kan säga att i en deterministisk

modell beskrivs verkligheten som om den vore exakt känd men i en stokastisk modell läggs en slumpterm till som beskriver det vi inte kan fånga i den deterministiska modellen. För att vara användbar som en nedskalning måste processen bero på klimatvariabler på de skalor som modelleras i klimatmodeller, såsom lufttryck, höjd över havet, vind och temperatur (Guttorp).

Osäkerhet och modellbedömning

Många faktorer påverkar resultatet av en regional modell, t ex vilken global modell som använts, hur olika startvärden för oceanerna sätts i den drivande globala modellen eller vilket scenario som använts för klimatprojektioner. I likhet med väderprognoser används ofta en ensemble av

modeller, dels för att förbättra prognosen/projektionen, och dels för att bedöma modellosäkerheten (Kjellström, Nikulin). Man kan då se om de olika modellvalen leder till samma sorts slutsats eller ej. Från statistisk synpunkt är ensemblemetoder

för närvarande något problematiska, eftersom det inte finns något givet sannolikhetsrum av alla tänkbara modeller. Har vi ett stickprov från något universum av modeller, eller är universum valt ur bekvämlighetssynpunkt? Man kan använda sig av långa klimatsimuleringar för att studera klimatets interna variabilitet. Tillförlitligheten hos ensemblemetoder är ett centralt problem inom klimatforskningen där det finns behov av statistiska tankesätt.

Ett sätt att skatta den naturliga variationen i klimatet är att använda paleodata, såsom pollen från svunna tider eller mineralinnehåll från sjösediment (de Luna, Sugita). De utgör indirekta mätningar och kan vara svåra att direkt relatera till klimatvariabler av intresse.

För att kunna jämföra en klimatmodell med data måste man tänka på vad som är en korrekt jämförelse. Man kan till exempel jämföra en empirisk fördelning av regn över en viss period med värden som fås från en regional klimatmodell som till randvärden använder globala meteorologiska analyser. En klimatmodell producerar däremot inte en rum-tidsserie som är direkt jämförbar med väderobservationer. Ur

»»»

»»» UPP- OCH NEDSKALNING AV KLIMATMODELLER

statistisk synvinkel är klimatet fördelningen av väder över en lång tidsperiod. Det går att jämföra fördelningarna för modellgenererade respektive observerade värden över ungefär 30-40 år (Orskaug). Man måste relatera de areella modellvärdena (jfr rutnät) med de punktvisa stationsvärden som observationerna ofta är, t ex med interpolering, dvs uppskalning från stationsvärdena. Detta kan vara en utmaning, till exempel när mätvärden av biomassaproduktionen från ett begränsat antal mätstationer skall jämföras med modellvärden i flera tiotusentals modellceller.

Om klimatmodellerna inte är tillfredsställande kan man undersöka om det går att kalibrera modellvärdena så att de två fördelningarna kommer närmare varandra (Berrocal, Haug).

Statistiska verktyg

När man sysslar med statistisk analys relaterad till nedskalning av klimatmodeller har man användning av verktyg såsom Markovfält, hierarkiska modeller, icke-stationära gömda Markovkedjor och metoder för att jämföra statistiska fördelningar (Anderson, Guttorp, Lindgren, Lindström). Från teoretisk synpunkt behövs egenskaper för icke-stationära empiriska processer och icke-gaussiska spatiala interpolationsmetoder (Podgorski).

Vart går vi nu?

Seminarier gav en god belysning av flera centrala statistiska problem inom klimatforskningen. En central fråga är hur säkra slutsatser man kan dra av olika klimatmodeller. Så gott som alla modeller är byggda på kända fysikaliska principer, men de innehåller alla olika grader och typer av förenklingar. Vilka slutsatser av en modell är säkra och vilka är mindre säkra? Eftersom användning av modeller oftast är det enda sättet att utföra klimatexperiment, och varje datorkörning tar mycket lång tid, behövs (och används) ordentlig statistisk försöksplanering och utvärdering. En annan statistisk svårighet är hur man mäter överensstämmelsen mellan olika modeller, och mellan modell och data. En nederbördsmodell som är bra på att förutsäga hur mycket regn det kommer men placerar det på fel plats eller vid fel tidpunkt kan vara användbar till mycket, men helt oanvändbar för planering av kvällens grillfest. En tredje fråga är hur man värderar de mycket stora datamängder som genereras i modellerna i förhållande till en begränsad tillgång till empiriska data.

En av slutsatserna av mötet var vikten av att utveckla en gemensam terminologi. Termer som komplexitet, icke-linjära processer och parametrisering har olika betydelser beroende på vem som använder dem. Det enda sättet att undvika språkförbistring är ökat samarbete, vilket vi hoppas ska bli ett av resultaten av denna workshop. Ett sådant samarbete kommer att påverka klimatforskningen och samtidigt ge upphov till många nya utmaningar för den statistiska forskningen. Det behövs mer av skarpa statistiska verktyg!

PETER GUTTORG,
UNIVERSITY OF WASHINGTON, USA,
OCH NORSK REGNESENTRAL, OSLO

NYA I SVENSKA STATISTIKFRÄMJANDETS STYRELSE

Pernilla Andersson, sekreterare

– Jag jobbar på SCB men är inte statistiker. Hur gick det till egentligen?



FOTO: MONICA HOMBURG

Som barn ville jag helst bli cirkusartist. Eller slalomstjärna. Som 16-åring blev jag antagen till Malungs alpina skidgymnasium och flyttade från Örebro till Dalarna. Ett antal år och knäskador senare lade jag skidorna på hyllan. Under nittioalet läste jag filosofi och nussvenska vid Umeå, Örebro och Stockholms universitet. Under denna period närdes en liten dröm om att fortsätta med filosofin på forskarnivå. Efter några år som kriminalvårdare och vårdare inom den kommunala socialpsykiatrien halkade jag in på SCB. Där har jag stannat kvar i tio år och stortrivs

med både arbetsuppgifter och kollegor. Jag arbetar i en grupp som kallas Enheten för mätteknik. Vi är en samling beteende- och språkvetare som arbetar med att utforma och förbättra frågor, informationsbrev, instruktioner och annat material som används i kontakten med uppgiftslämnare. I det arbetet är det främst kvalitativa testmetoder som är aktuella (t.ex. kognitiva intervjuer), men även metoder med mer kvantitativa inslag som experiment och beteendekodning av den sociala interaktionen i intervjusituationer. Mina kollegor Andreas Persson och Fredrik Scheffer skrev om att utvärdera frågor i Qvintensen nr 1/2011. För att komplettera min utbildning läser jag statistik på deltid. Följaktligen har jag för tillfället inte mycket till fritid. Men jag försöker ändå ta en löptur då och då, spela lite onlinepoker och slappa med en bok i sommarstugan tillsammans med sambon och katten.

Göran Lande, ledamot

– Under mina 40 år inom svensk industri har jag praktiserat och populariserat tillämpad statistik.



FOTO: LINA LANDE

Mina anställningar har varit på Bjerking Ing Byrå, Nitro Consult AB, Nitro Nobel AB, Saudi Chemicals LTD, Kema Nobel, Norabel AB, Nitro Metall AB, Volvo-Nobel JV, Vibration Measurement Engineering – VME LTD, FMV, Ericsson Telecom AB, Ericsson Radio Systems AB, Ericsson Business Consulting, Ericsson Radio, Ericsson Quality Management Insti-

tute, Ericsson AB och Sony Ericsson Mobile Communications AB.

– På universitet har jag läst geovetenskap, statistik, matematik, geofysik och kvalitetsteknik och inom telekomindustrin har jag som lärare i Lean Six Sigma utbildat flera hundra linjechefer i statistiskt tänkande och lärt ut statistik till ca 550 Six Sigma Black Belts.

– Nu, efter min pensionering från Sony Ericsson, arbetar jag som egen företagare med konsultationer inom kvalitetsteknik och industriell statistik. Mitt mål är att hjälpa företagare och anställda bli bättre genom att påvisa att kvalitetsteknik och statistik stärker företagets lönsamhet och ökar deras kvalitetsförmåga.



Linus Schiöler.

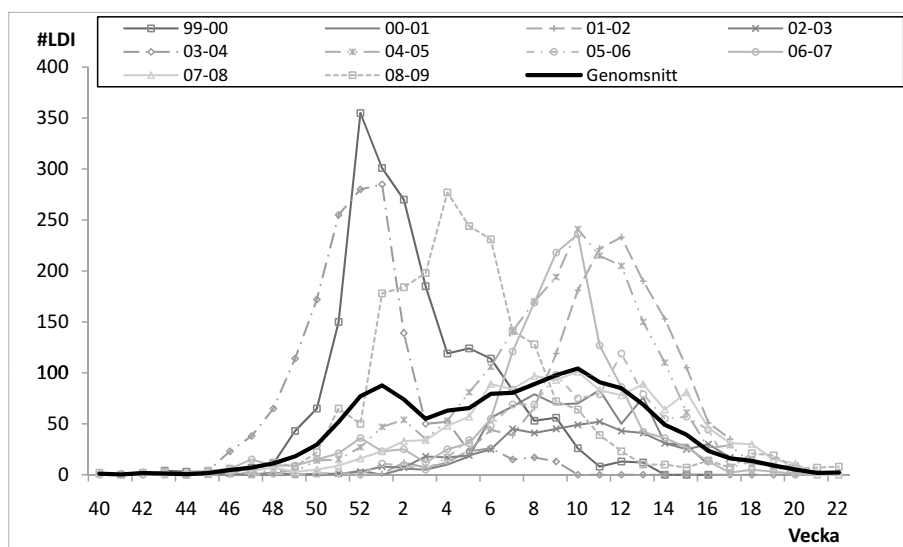
Det skrivs nu många intressanta doktorsavhandlingar i statistik. De behandlar ett brett spektrum av tillämpningar. Linus Schiöler skriver här om statistisk övervakning.

Övervakning av utbrott

I nom många områden görs fortlöpande observationer i syfte att snabbt upptäcka förändringar. Inom tillverkningsindustrin började enkla metoder användas redan 1930. Behovet av övervakning inom finans, miljö och medicin har medfört att teorin för statistisk övervakning, och därmed möjligheten att konstruera optimala metoder, har utvecklats snabbt de senaste åren.

Det blir allt viktigare att snabbt upptäcka utbrott av epidemiska sjukdomar. Utbrott av nya virus som svin- och fågelinfluensa visar att det finns behov av beredskap vid krissituationer. Genom att snabbt upptäcka nya utbrott kan sjukvård planeras och preventiva åtgärder vidtas. Detta kan avsevärt minska förlusten av människoliv samt kostnader för sjukskrivningar och sjukvård. Detsamma gäller det årliga influensautbrottet under vintern, som varje år leder till tusentals dödsfall och kostar samhället stora belopp. Intresset för statistisk övervakning är större än någonsin.

När man försöker upptäcka en förändring måste man göra en avvägning mellan att göra upptäckten så snabbt som möjligt och säkerheten i bedömningen. Genom att använda metoder för statistisk övervakning kan denna avvägning göras på ett optimalt sätt och effektiva beslutsystem som grundar sig på vedertagen statistisk teori kan utvecklas. En övervakningsmetods egenskaper kan utvärderas genom att härleda dem ur statistisk teori eller genom att göra simuleringsstudier med en statistisk modell som efterliknar verkligheten. Detta gör att det på förhand går att ha en mycket god uppfattning om hur en statistisk metod fungerar när den tillämpas i praktiken. Det går alltså att uttala sig om hur säker man är på att en förändring verkligen har skett när man får ett larm från systemet. Man får också en uppfattning om hur



Figur 1. Antal fall av laboratoriediagnostiserad influensa (LDI) i Sverige under nio säsonger samt en genomsnittskurva.

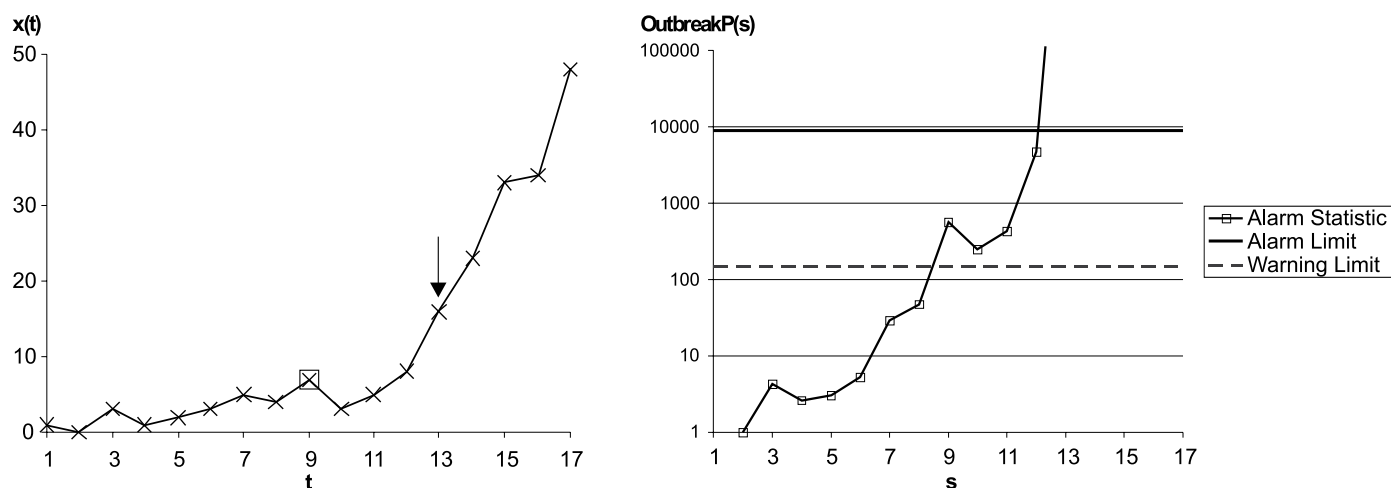
lång fördröjning man kan förvänta sig från att en förändring sker till att ett larm ges. Denna objektivitet är en stor fördel i praktiken.

Vi genomförde ett experiment där vi lät studenter i medicinsk statistik på subjektiv väg försöka avgöra när ett utbrott hade börjat. Samtliga studenter följde fyra simulerade influensautbrott (i randomiserad ordning) och avgjorde när de tyckte att utbrottet börjat. Resultatet visar att det objektiva systemet upptäcker förändringar snabbare och säkrare. Det visade sig även att de subjektiva bedömningarna var starkt beroende av vem som gjorde dem.

För att göra övervakningsmetoden enkel att använda har vi utvecklat användarvänliga programtillägg för Microsoft Excel. Dessa program kan fritt laddas ned från www.statistics.gu.se/surveillance. Metoden har även gjorts tillgänglig i Smittskyddsinstitutets övervakningssystem CASE, samt i paketet "surveillance" för R.

Liksom vid all annan statistisk analys är man vid statistisk övervakning tvungen att göra vissa antaganden. Många metoder för upptäckt av influensautbrott baseras på avvikelser från en genomsnittskurva beräknad över flera år. Detta kan bli ineffektivt eftersom det finns stor variation mellan åren (se Figur 1). När utbrott sker sent under säsongen har dessa metoder svårt att upptäcka utbrottet.

Felaktiga antaganden kan få som konsekvens att övervakningssystemet blir ineffektivt. Inget antagande görs angående en baslinjenivå, i stället utnyttjas att influensautbrott kan karakteriseras av en ändring från konstant incidens till ökande. En enkel men användbar modell för detta är att anta att observationerna är oberoende och likafördelade med medelvärde som efter utbrottet ökar monotont. En larmstatistika konstrueras för fallen där observationerna är normalfördelade med känd varians eller pois-



Figur 2. Vänster: Antal observerade laboratoriebekräftade fall av influensa i Sverige säsong 07-08. Den första observationen är för vecka 40 år 2007. Varnings- och larmsignaler för metoden är markerade med fyrkant respektive pil. Höger: Värdet på larmstatistikan vid olika tidpunkter för data till vänster.

»»» ÖVERVAKNING AV UTBROTT

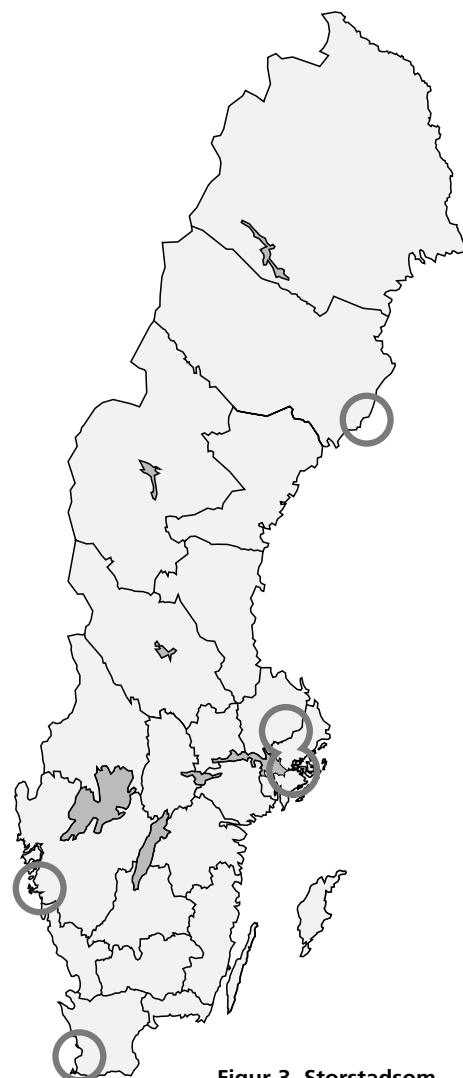
sonfördelade. Larmstatistikan bygger på den generaliserade likelihoodkvoten mellan en konstant nivå och en monotont ökande. Då maximum-likelihoodskattningen under dessa antaganden ges av icke-parametriskt isoton regression är metoden semiparametrisk. I Figur 2 visas resultat när metoden tillämpas på svenska influensadata från 2007.

Vid övervakning av sjukdomsutbrott har man ofta data tillgängliga från flera olika platser. I [4] analyseras flera års svenska influensadata från ett tjugotal laboratorier. Det finns inga tydliga tendenser till ett geografiskt spridningsmönster. Slutsatsen är i stället att influensautbrottet i Sverige börjar cirka en vecka tidigare i storstäderna än i övriga landet (se Figur 3). I [5] utnyttjar vi denna information för att konstruera ett effektivare övervakningssystem. Det rör sig om övervakning av flera variabler, så kallad multivariat övervakning.

Vid multivariata problem är det en fördel att reducera komplexiteten så mycket som möjligt utan att gå miste om information. Då man

med en uttömmande statistiska inte går miste om någon information är det ofta önskvärt att hitta sådana. I [3] härleder vi en uttömmande statistika för fallet där man följer ett antal oberoende och lika fördelade processer och är intresserad av att upptäcka en förändring från ett medelvärde till ett annat. Statistikan bygger på att man antar att det mellan de olika processerna finns en känd fördröjning av när förändringen sker. Ofta är det dock mer lämpligt att beskriva ett utbrott som ett gradvis men monotont ökande av medelvärdet. I [5] vidareutvecklar vi den uttömmande statistikan till att gälla för detta fall. Vi vidareutvecklar även övervakningsmetoderna från [1] för att utnyttja denna statistika och tillämpar systemet på de årliga influensautbrotten i Sverige. Resultatet blir ett övervakningssystem som upptäcker utbrott snabbare än ett som behandlar hela landet som en enhet.

LINUS SCHIÖLER,
GÖTEBORGS UNIVERSITET



Figur 3. Storstadsområden där influensautbrotten sker först.

Referenser

1. Frisén, M., Andersson, E. and Schiöler L. (2009) **Robust outbreak surveillance of epidemics in Sweden.** *Statistics in Medicine*, 28, 476-493.
2. Frisén, M., Andersson, E. and Schiöler L. (2010) **Evaluation of Multivariate Surveillance.** *Journal of Applied Statistics*, 37, 2089-2100.
3. Frisén, M., Andersson, E. and Schiöler L. (2011)

Sufficient Reduction in Multivariate Surveillance.

Communications in Statistics – Theory and Methods, 40, 1821-1838.

4. Schiöler, L. (2011)

Characterisation of influenza outbreaks in Sweden. *Scandinavian Journal of Public Health*, 39, 427-436.

5. Schiöler, L. and Frisén M.

Multivariate outbreak detection (i tryck).

Journal of Applied Statistics.

ORDFÖRANDEN HAR ORDET

Öppna data vinner



Under en av de sista semesterveckorna läser jag i tidningen att det ska anordnas en okonferens (ja, den kallas så på grund av sitt okonventionella upplägg) vid namn Sweden Social Web Camp. Tilldragelsen är tänkt att locka 400 internetentusiaster och anordnas i år för andra gången. Vad som får mig att höja på ögonbrynen är möjliga samtalsämnen på (o)konferensen. I fjol var tydligen "öppna data" det hetaste ämnet och det tros bli den mest diskuterade frågan även i år. Arrangören exemplifierar med att SMHI inte har sina data öppet tillgängliga, till skillnad från YR (den norska motsvarigheten).

När Hans Rosling talade på det Nordiska statistikermötet 2010 i Köpenhamn tog han upp parallellen till väderdata, nämligen samhällsdata. Han halade fram en dansk sedel ur fickan: "Den här ägs av den danska riksbanken, men det är fritt fram för mig att använda den." På samma sätt, menar Rosling, bör vi som tar fram och tillhandahåller data om samhället (alltså statistik) se på saken. Vi kvalitetsmärker data med en tänkt vattenstämpel, men låter användarna ta del av dem utan avgifter och krångligheter.

Idén vad gäller att dela med sig av information är snarlik den som tillämpas av dem som förespråkar öppen källkod. Datorprogram som skrivs med *öppen källkod* är fritt tillgängliga att granska och använda för var och en som är intresserad. I sammanhanget är det viktigt att poängtera skillnaden mellan dem som argumenterar för öppen källkod och dem som anser att allt på nätet bör vara gratis. Ett program som är skrivet med öppen källkod kan mycket väl ingå som en del i en paketerad vara som är till salu. Det finns till exempel skräddarsydda Linuxdistributioner att köpa.

Tittar vi oss omkring så ser vi att fenome-



Londons tunnelbanesystem med alla tunnelbanetågs positioner i realtid. [Http://PublicData.eu](http://PublicData.eu).

net öppna data är något som vinner större och större utrymme. Internationellt rapporterar Världsbanken (<http://data.worldbank.org/>) att Kenya tagit täten bland utvecklingsländerna vad gäller att göra data tillgängliga, i detta

fall *key government data*. För europeiska data finns webbplatsen <http://PublicData.eu> där det till exempel går att finna ett ypperligt prov på visualisering: en karta över Londons tunnelbanesystem med alla tunnelbanetågs positioner i realtid! I Sverige ligger regeringen bakom

webbplatsen www.openaid.se vars syfte är att öppna upp för insyn och idéer för att göra biståndet mer effektivt. Här handlar det alltså om *öppna offentliga data*, vilket är överlappet mellan öppna och offentliga data.

Det här med öppna data är ett område där det tydligen händer saker. Efter en stunds googlande hittar jag sidan som berättar om *The Open Data Challenge*, en tävling som gick av stapeln tidigare i år och som flera större städer och länder i Europa stod bakom. Tävlingen gick ut på att göra öppna data tillgängliga för allmänheten med hjälp

av dataprogram. Vann gjorde programmet *bePart*, som underlättar för intresserade invånare att ta del av, och kanske i förlängningen påverka, den urbana utvecklingen. Näst sista veckan i september anordnades en konferens i Brisbane i Australien: *Open Government Data Conference and Data Camp* (<http://creativecommons.org.au/ogdbristbane>). Arrangörerna understryker vikten av öppna data och menar att politiskt beslutsfattande, transparanta demokratiprocesser och forskning underlättas och främjas av denna öppenhet.

I Kanada rapporterades förra året om projektet *Open Data Ottawa* där deltagare samlades för att arbeta på appar som på något sätt skulle bidra till samhällsnyttan med hjälp av information som delas digitalt. Ett program som lyfts fram heter *fixmystreet* (<http://fixmystreet.ca/>) med vars hjälp personer kan rapportera in om gatlyktor som inte längre fungerar, sönderslagna soptunnor och så vidare. Speciellt efterlyser man statistiker för att bearbeta alla data till meningsfulla rapporter. Och det låter väl inte som en helt avlägsen tanke – att ju mer tillgängliga offentliga data blir, desto mer kommer alla vi som på något sätt arbetar med statistik att behövas.

Framför mig ser jag hur något stort kommit i rullning. Nu får vi hoppas, och bidra till, att statistiker är med från allra första början. Det vore olyckligt om scenariot blev detsamma som när *data mining*-stenen började rulla – ett område nästan helt och hållet ockuperat av systemvetare och datatekniker.

JOAKIM MALMDIN
joakim.malmdin@scb.se



FOTO: MONICA HOLMBERG

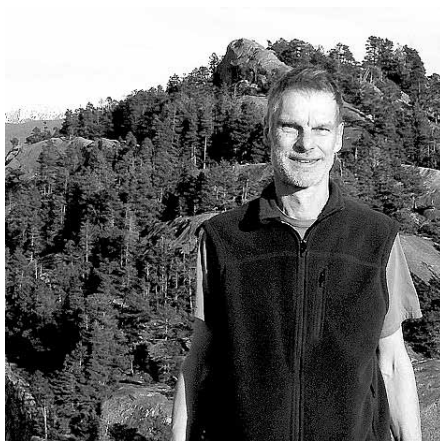
Surveymetoder på underbar plats

REFERAT FRÅN BALTISK-NORDISKA KONFERENSEN I SURVEYSTATISTIK

Med siktet inställt på Norrfällsviken vid Höga kusten för ett femtiotal statistiker från nitton länder för att konferera under en vecka i juni.

Det var den tredje konferensen inom det baltisk-nordisk-ukrainska, och numera även vitryska, nätverket för surveystatistik. Ett soligare försommarväder hade varit välkommet under veckan, även om väderguden höll hammaren kvar i bältet under sin himmelsfärd. Men några av de långväga besökarna hade lyckligtvis prickat in Europas varmaste väder ett par dagar innan konferensstart.

Veckan innehöll bland annat längre föredrag av de tre huvudtalarna Jean-Claude Deville (Centre de Recherche en Économie et Statistique, Frankrike), Giovanna Ranalli (universitetet i Perugia, Italien) och Steven Thompson (Simon Fraser University, Kanada). Steven Thompson talade om adaptive sampling. Adaptiv stickprovsteori är ett samlingsbegrepp för designer där man väljer urvals-sannolikheterna så att de beror på undersökningsvariabelns värde. (Därmed kan de inte vara kända i förväg.) Thompson diskuterade särskilt två konkreta fall där adaptiv sampling är användbar: spatiala stickprov där man vill undersöka klustrade fenomen, exempelvis djur och växter som lever samlade i grupper, och stickprov ur dolda populationer med ett socialt nätverk mellan individerna, såsom HIV-bärare, droganvändare eller prostituerade. För



Steven Thompson, som var en av tre huvudtalare, talade om adaptive sampling.

de spatiala stickproven dras i princip först ett litet område slumpmässigt. Visar sig området ha många förekomster av det man söker så ges omgivande områden sedan en högre sannolikhet att ingå i urvalet. För stickprov ur dolda populationer använder man sig av det sociala nätverket för att välja ut objekt till stickprovet och för att

bygga bra estimatorer. Sammantaget gjorde Thompsons förmåga att kombinera förståelse av praktiska problem med matematisk stringens hans presentation till en av konferensens höjdpunkter.

ESS sägs vara världens mest fel- bemängda under- sökning

Många var nog de deltagare som sett fram emot att lyssna på Jean-Claude Deville, som är en av samplingsteoriens viktigaste bidragsgivare. De flesta syntes också belåtna, om än något duvade, efter vederbörandes framställning av kalibrering, dess tillämpningar vid svarsbortfall

samt balanserade urval. Deville levererade även visdomsorden "Statistics is the study of data, not of models". Detta höll nästa huvudtalare, Giovanna Ranalli, med om men tillade vidare "I have realized that we should model more" (citerande programkommitténs ordförande Daniel Thorburn från dagen innan). Även Ranalli utgick huvudsakligen från ett designbaserat perspektiv då hon gav en exposé över modellassisterade estimatorer baserade (delvis) på icke-parametriska modeller. Hon tog även upp tillämpningar vid bortfall och mätfel, samt den aktuella frågan om hur man får meningsfulla skattningar även då stickprovet är väldigt litet (så kallad Small Area Estimation, SAE) då modellerna används på ett mer explicit sätt. Framställningen berikades av många pedagogiska exempel. Giovanna Ranallis generella rekommendation för icke-parametriska hjälpmodeller i surveyer var att använda penaliserade splines.

Förutom de tre huvudtalarnas föredrag bjöds åhörarna även på kortare presentationer från sju särskilt inbjudna gäster. Risto Lehtonen från Helsingfors universitet talade om SAE, och Natalie Shlomo från universitetet i Southampton om olika metoder för att bedöma risken att identifiera enskilda individer ur aggregerade data, såsom statistiska tabeller. Imbi Traat från universitetet i Tartu i Estland diskuterade en typ av kalibreringsestimatorer som ger konsistens även mellan olika undersökningar.

Lennart Bondesson från Umeå universitet diskuterade flera olika algoritmer för dragning av π ps-urval, bland annat så kallad Extended Sampford Sampling (ESS). Förkortningen ESS kan dock även avse den europeiska undersökningen



Vy från Höga kustenbron. Ett femtiotal statistiker från nitton länder samlades i Norrfällsviken för att konferera under en vecka i juni.

European Social Survey, vilken ”icke-statistiskern” Ineke Stoop (The Netherlands Institute for Social Research) berättade om i sitt anförande. I denna typ av geografiskt gränsöverskridande undersökningar finns ofta stora harmoniseringsproblem. Förutom rent praktiska problem med att få alla länder att genomföra undersökningen på ett likartat sätt finns mer subtila sådana där kulturskillnader mellan länderna påverkar invånarnas tolkningar av en och samma fråga. Stoop berättade om ESS höga ambition vad gäller dokumentation av statistiska fel, vilken ska ha gett undersökningen epitetet ”the survey with the most errors in the entire world”.

En annan av de inbjudna gästerna var Carl-Erik Särndal (SCB och Örebro universitet), som i ett mer filosofiskt anförande delgav sin syn på surveystatistik ”i dåtid, nutid och framtid”. Han beskrev officiell statistikproduktion som ett ”fragmenterat område” ur ett vetenskapsteoretiskt perspektiv: flera konkurrerande teorier existerar samtidigt och den utveckling som sker är inte enhetlig. Särndal citerade vidare fysi-

kerna och filosoferna D. Bohm och F. D. Peat vilka menar att en hög grad av fragmentering inom ett forskningsfält starkt försvårar långsiktig utveckling. Huruvida det finns något egentligt alternativ till denna fragmentering inom området officiell statistikproduktion lämnades dock som en öppen fråga.

Anders Norberg från SCB, även han en av de inbjudna gästerna, talade slutligen om så kallad selektiv granskning. Både ur ett ekonomiskt perspektiv och med hänsyn till uppgiftslämnarbördan är det viktigt att inte besvara uppgiftslämnare i onödan för att kontrollera rapporterade värden. Det gäller att ta hänsyn till objektens potentiella påverkan på skattningen, vilket är vad selektiv granskning syftar till. Norberg gav även en kortfattad introduktion till granskningsverktyget SELEKT vilket utvecklats på SCB och nyligen implementerats inom en rad undersökningar där.

Även om konferensen bjöd på många nyttiga möten och viktiga diskussioner såväl inom som

mellan grupperna av teoretiker och mer praktiskt inriktade statistiker, så är ett dylikt evenemang inte enbart begränsat till sakfrågorna. I halvtid återfanns en utflykt till lokala sevärdheter, där naturupplevelserna stod i centrum. Höga kustenområdet, vars landhöjning efter istiden uppmätts till världens snabbaste och därmed kvalificerat det till en plats på UNESCO:s världsarvslista, är en lisa för själen väl så god som norska fjordlandskap och Toscanas gröna dalar. Den gemytliga stämningen gjorde även att skrattet aldrig var avlägset. Här kan bland annat nämnas att en av organisatörerna (medvetet) tog orden ”the floor is yours” bokstavligt och formligen kröp fram under bordet för att nå fram till talarstolen. I avslutningstalet avslöjade nätverkets nestor Gunnar Kulldorff att de två närmaste årens större aktiviteter inom nätverket är planerade att äga rum i Lettland 2012 och Vitryssland 2013. Vi ses förhoppningsvis då!

OLIVA HELLBERG OCH NICKLAS PETERSSON,
STOCKHOLMS UNIVERSITET



Minnen, MBDD och stipendiater



Vilken förmån det var att ha en lärare som Erik Leander. Med glimten i ögat och ett oförglömligt kroppsspråk och uttryckssätt lotsade han sina elever genom statistikens grunder. Hans axellånga, gråa, rufsiga hår och busiga uppsyn fick honom att påminna om Einstein. När jag blundar kan jag se honom komma släpande med två stora plastpåsar i vilka han hade våra stencilar, en stor väckarklocka och en gigantisk sladdförsedd räknemaskin. Det som betydde allra mest för mig var ändå Eriks konsultuppdrag bl. a. på Regionsjukhuset i Linköping. Han erbjöd oss studerande att arbeta med aktuella medicinska frågeställningar och riktiga data samt ordnade små korta extraknäck åt oss som var intresserade. Jag ville verkligen jobba med medicin, det visste jag från början. Efter ett antal sommarjobb på Sandbyhovs sjukhus som vikarierande vårdbiträde var jag nyfiken på området.

Att som student, under handledning, få praktisera som konsulterande statistiker är otroligt värdefullt och lärorikt. Jag önskar att alla skulle få den chansen under sin utbildning.

Sista sommarlovet under mina universitetsår ordnade Erik ett fantastiskt sommarjobb åt mig. Han hade fått ett uppdrag från Oskarshamns sjukhus. Ett arbete som krävde några veckors intensivt dataknådande. Erik frågade den statistiksupportsökande läkaren om han kunde tänka sig att ta hjälp av en studerande om Erik handledde. Mer än gärna

sade doktor Joep Perk. Jag packade en väska och reste ner till Oskarshamn och tog in på hotell Post. Oskarshamn hade kärnkraft och en väldigt lång bänk, över 70 meter, men mer än så visste jag inte om kuststaden. Joep arbetade med rehabilitering av hjärtinfarktspatienter och hans forskning undersökte vikten av olika åtgärder och dess effekt på risken för återfall. Jag kommer ihåg hur jag kämpade de första dagarna innan jag hade klurat ut hur jag skulle hantera datamängden och förstätt frågeställningen. Jag minns att jag ringde till Erik och frågade om jag var på rätt spår. Inte så lätt att veta vilken metod som ska tillämpas.

»Vi samarbetar med medicinska experter och hjälper dem att komma ett steg närmare förståelse och kunskap om sjukdomar och behandlingsmetoder.»

Idag känns det så självklart och det är skrattretande att tänka på hur jag funderade på alternativa metoder. De första åren som statistiker är ju lite trevande. I verk-

tygslådan finns ett antal redskap men det blir inte så lyckat att använda en hammare till en skruv. Skönt att Erik fanns där och handledde, stöttade och kontrollerade. Det inledande arbetet i Oskarshamn avlöstes med fortsatt periodiskt arbete i universitetets datorsalar. Så lärorikt och stimulerande att få förmånen att jobba parallellt med sin utbildning. Jag fick en mycket bra förståelse för vad statistikyrket egentligen går ut på. Vi biostatistiker samarbetar med medicinska experter och hjälper dem att komma ett steg närmare förståelse och kunskap om sjukdomar och behandlingsmetoder.

Många år senare fick jag Joeps doktorsav-

Inte så lyckat.

FOTO: LINNEA HEDLIN

MODEL BASED DRUG DEVELOPMENT

Around 60 participants attended the conference Model Based Drug Development, jointly organized by FMS and the Statisticians in Finnish Pharmaceutical Industry (SSL) and hosted by KTH.

After some welcoming words and introduction to the program by Anna Torrång, FMS, and Tommi Nurminen, SSL, and a short overview about the Center for Industrial and Applied Mathematics (CIAM) at KTH by Anders Lindquist, the presentations took off.

The first speaker of the day was Jouni Kerman, Novartis Pharma, who gave a presentation with the title "Bayesian modeling in clinical trials: from early development to phase III". Several arguments for using Bayesian modeling were presented. One is that it provides a framework for incorporating all relevant information to make informed decisions. Incorporating historical controls is one example. Jouni described the process of incorporating information about the effect of an already known (control) drug to be used for the prior distribution for the treatment effect. In addition, Bayesian modeling in clinical trials provides a framework for quantifying and incorporating uncertainty. He argued that Bayesian modeling improves the communication around clinical trials since decision rules and results from studies can be formulated to mirror the scientific questions of interest. The need for both Bayesian and frequentist approaches was stressed, they can and should live in a peaceful coexistence. However, challenges in adopting the Bayesian approach in drug development exist: ignorance (lack of acceptance) and the need of mathematical modeling and programming experience. The latter is also an opportunity for industry statisticians to develop in their role.

Jukka Corander, University of Helsinki, was the second speaker. While the use of Bayesian methods was also a central theme

in his talk "Bayesian prediction of protein-ligand interactions", he discussed a very different problem from the biophysics area with applications in the preclinical phase of drug development. The aim is to build a statistical model for describing the distance of a ligand from the protein and the angle in relation to the protein. Jukka illustrated and visualized this problem both with graphs in his presentation and by using his laptop as a ligand and himself as protein. He successfully managed to introduce the topic of protein-ligand interactions to the audience despite it being a new topic to presumably most of the participants. An important characteristic of his modeling problem is that no purely data-driven modeling is desired since physical theory gives information about possible structures. As a consequence, he stressed the importance of supervising the parametric model and to restrict the parameters to what physical theory suggests. His talk provided a good example of knowledge-based modeling where the knowledge is based on theory (here biochemical) rather than on collected data which otherwise often is the focus.

Matts Kågedal, AstraZeneca, spoke about "Model based design of Phase IIb studies: focus on dose-finding". He described how pharmacometrics can be used in the analysis of clinical dose-finding trials to support the choice of dose(s) to be used in larger patient trials. Pharmacometrics is a newly defined area coming from pharmacokinetics. In

contrast to the traditional method of pairwise comparisons of different dose groups in a dose-finding trial, pharmacometric modeling describes the dose-exposure-response relationship. Methods for assessing the precision of

the estimated optimal dose were presented. Based on an example from neuropathic pain drug development Matts showed how predicted precision in dose selection for a larger patient study (based on simulations) can be used to guide the choice of the dose-finding study design. He pointed out that the selection of the study design depends on many factors: the assumed shape of the dose-exposure-response relationship, the uncertainty (which is often underestimated) »»»

handling med posten med tackord till mig. Resultaten av mitt första riktiga arbete tillsammans med en läkare fanns nu dokumenterade i en liten bok.

Nu sitter jag här nära tjugo år senare och skriver min spalt. Erik har lämnat oss, Joep är numera hälsoprofessor och jag är ordförande i FMS.

I maj var det äntligen dags för endagskonferensen Model Based Drug Development - MBDD på KTH i Stockholm som FMS arrangerade tillsammans med vår finska motsvarighet SSL. Förberedelsearbetet hade pågått länge, med hög intensitet i slutspurt och vi i programkommittén njöt av att få uppleva det efterlängtda resultatet. CIAM (Center for Industrial and Applied Mathematics) på KTH var proffsiga värdar och medarrangörer. Jag tycker dagen blev lyckad och vi visade att statistiker kan göra så mycket mer än att bara räkna stickprovsstorlekar och leverera p-värden (se referatet på nästa sida). När ni läser denna spalt har redan nästa tillställning, Statistikerträffen i Göteborg, ägt rum.

2011 års FMS-stipendiater utsågs i juni. Stort grattis till Anna Ekman, Hatef Darabi, Anna Lindam och Rebecca Rylance. Ett stort tack till våra partner som gör stipendiet möjligt.

Nu är det dags att ta tag i ännu en höst. En årstid jag faktiskt gillar och ser fram emot. Vi planerar härnäst att anordna ett föreningsmöte tillsammans med Surveysektionen. Hopas vi ses på kommande evenemang.

ANNA TORRÅNG



»»» MODEL BASED DRUG DEVELOPMENT

since the true model is in reality unknown), the desired precision in estimation, and the method of analysis, but also practical aspects like the availability of dose strengths.

Tobias Rydén, KTH, gave an “Introduction to hidden Markov models – with applications to longitudinal data”. For example, in a study where patients can be either in a relapse or in a remission state, it is evident that Markov models can be useful to describe dynamics over time. In this application, however, the relapse or the remission state of the patient is not directly observed, i.e. the state is “hidden”. Instead, another variable is observed. This variable is assumed to have a distribution that depends on the state of the patient. After introducing the concepts using a simple example with two hidden states and an observed variable with a normal distribution conditional on the hidden state, he described two examples from clinical drug development. The first was an example from a multiple sclerosis study where the number of lesions in the brain was observed providing information about the state of the patient (relapse or remission). The second example was a schizophrenia study where frequent side effects measured with rating scales gave information about in which (hidden) health state the patient was. In these models, the following types of inference are often of interest: prediction (estimating the state of the patient at time t given the observations until time $t-1$), filtering (estimating the state at time t given observations until and including time t), and smoothing (estimating the state at time t given all observations including those observed after time t). Different methods for estimating the parameters of interest in these models were also discussed.

Stig-Johan Wiklund, AstraZeneca, gave examples of how quantitative models can be used to provide decision basis under the title “Modeling to evaluate drug development strategies”. His key message was that modeling should be used for high-level drug development decisions, and not only for decisions regarding a specific study. One example he gave addressed the question of to what extent animal models (preclinical trials) and biomarkers should be used in early phases of drug development. Clearly, the answer to this question depends on the predictive ability of

the animal models and biomarkers. Stig-Johan presented a conceptual model for investigating where the cut-off point is in terms of the correlation needed between animal trial outcome (or biomarker) and clinical outcome, and gave examples of situations where abandoning them completely would in fact increase the return of investment. Another example of high-level decision in drug development is the choice of program for clinical trials. The current strategy of sequentially evaluating one candidate drug from the same drug family at a time was contrasted with a parallel drug development paradigm where several candidate drugs are evaluated in parallel in the same studies. The comparison of the sequential and the parallel paradigms was also based on the return on investment. The parallel paradigm was shown to be beneficial in several most probable scenarios. The main explanation to this finding was suggested to be the ability to proceed faster with the best candidate drug in a parallel program.

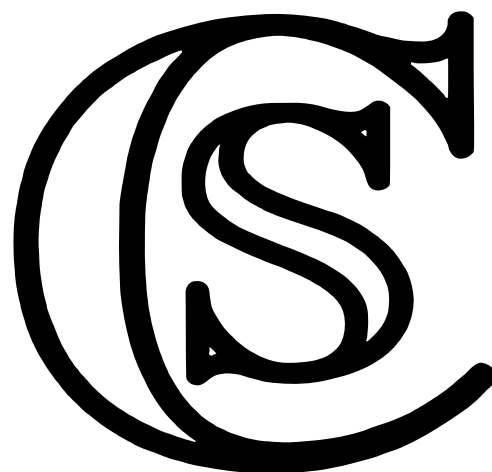
In the last presentation of the day, Alun Bedding, GlaxoSmithKline, talked about “Model based adaptive designs in dose-finding including case studies”. He motivated modeling in dose-finding by contrasting it to traditional pairwise testing. Dose-response is an integral component of drug development but the problem is that it is often not investigated in a good way. In an adaptive design doses can be dropped or added in the course of the study based on the results collected so far. Thus adaptive designs

can help to avoid investigation of suboptimal doses. He highlighted the importance of pre-planned adaptation, i.e., the rules how the doses are chosen in the course of the study should be pre-specified. As examples, he showed an early phase diabetes trial where a tolerability outcome was modeled using logistic regression. The doses of the patients in subsequent stages were determined based on the estimated (Bayesian) probabilities for this tolerability-event. A second example was a dose-finding study for a severe asthma population with the aim to determine the minimum effective dose and to gain an understanding of the dose-response. By simulation, the characteristics of an adaptive design were investigated. In this adaptive design 300 patients were enrolled in the first stage and treated with placebo, an active comparator, or one of four doses of the study drug. An interim analysis was performed and doses of the study drug could be dropped according to pre-defined criteria. These criteria depended on the model chosen for the dose-response and he investigated different models like an Emax-sigmoid model or a normal dynamic linear model. The characteristics of this adaptive design were compared with traditional non-adaptive parallel group and cross-over designs.

The day was wrapped up with a panel discussion, with all the presenters in the panel and Carl-Fredrik Burman as chair. Among other things, it was discussed what skills are needed for modeling in the pharmaceutical industry. Communication and collaboration skills and curiosity to continuously learn more were identified as important. The ability to formulate a real world problem in the right way to allow for modeling was also raised as being crucial. Different views were raised around the need for subject knowledge and experience and seem to depend on the subject area. Some areas might require substantial bioscience knowledge, while in some instances communicating and collaborating skills enable a modeler to do a good job without subject knowledge by collaborating in cross-skills teams.

ANNICA DOMINICUS AND FRANK MILLER,
ASTRAZENECA





ORDFÖRANDEN HAR ORDET

Sommarskolan på Asa herrgård

I skrivande stund, när vardagen har återkommit efter en förhoppningsvis vederkvickande semester för oss alla, kan det passa med en tillbakablick på gångna aktiviteter och en framåtblick på de nya som väntar. För Cramérsällskapets del handlar tillbakablick om sommarskolan, som ägde rum i juni på idylliska Asa herrgård utanför Växjö. Temat var statistical learning (se Cramérsällskapets webbplats, som du enklast når via Svenska statistikfrämjandets webbplats). Den största andelen deltagare kom från Uppsala, Stockholm och Jönköping, men även några andra personer med specifikt intresse för ämnesområdet hade letat sig fram till vår skola.

Vi hade tre huvudföredragshållare. Inledningsvis talade Mattias Villani (tidigare forskare vid Riksbanken, numera professor i statistik i Linköping) om bayesianska metoder i allmänhet och deras användning när man har stora datamaterial i synnerhet. Eftersom Mattias började med en fyllig genomgång av grundläggande begrepp var det lätt att hänga med även om man som jag är lite av en novis i den bayesianska världen. Och trots att Mattias spenderade mycket tid på grunderna hade han gott om tid att diskutera mer komplicerade forskningsproblem. Bra gjort!

Efter Mattias tog John Shawe-Taylor vid. John är professor i "computational statistics and machine learning" vid University College of London. Han presenterade ämnet statistical learning från ett mera teoretiskt/klassiskt perspektiv än Mattias, med speciell inriktning

mot klassificering. Även om Johns föreläsningar höll en något svår teknisk nivå tycker jag att de var mycket intressanta. Det kändes en smula fascinerande att höra hur man kan se på denna typ av problem med en delvis helt annan grundläggande begreppsapparat än vad vi statistiker är vana vid. En annan intressant poäng var att den metodik som John presenterade även är tillämpbar för små datamaterial! Alltså är det fel att naivt sätta likhetstecken mellan statistical learning och data mining.

Den sista huvudföredragshållaren var Jennifer Castle. Jennifer har disputerat i nationalekonomi/ekonometri i Oxford med den kände David Hendry som handledare och nu arbetar hon vidare i Oxford som lärare och forskare. Hon presenterade metodiken automatisk inferens. Idén är att man stoppar in sina data i en "svart låda" (alltså ett statistiskt programpaket) och ut kommer ett färdigt förslag på modell. Metodiken är speciellt anpassad för univariat modellering av tidsserier, men den är under utveckling för det multivariata fallet. Det ligger i sakens natur att detta är lämpligt för stora datamängder. Mitt intryck var att Jennifers föredrag gav en god bild av vad automatisk inferens kan åstadkomma och hur detta går till.

Vidare fick vi lyssna på deltagarbidrag av doktoranderna Mehrdad Jafari Mamaghani (Stockholm) och James Blevins (Uppsala). Dessutom ägnades en eftermiddag åt en utflykt med statistiska förtecken. Det slumpade sig nämligen så att SLU har en forskningssta-

tion just vid Asa. Vi fick möjlighet att ta del av deras skogsförsök på ett mycket konkret sätt, med besök på två olika försöksfält i skogarna runt Asa. Den experimentdesign som används är split-plot, vilket vår egen styrelsemedlem Jan-Eric Englund förtjänstfullt introducerade oss i.

Fritiden på sommarskolan ägnade vi åt social samvaro och rekreation, bland annat i form av bastubad, kanotpaddling och promenader.

När det gäller den närmaste framtiden står Cramérsällskapets höstkonferens för dörren. Självklart är den öppen för alla. Den kommer att avlöpa i Uppsala den 27:e oktober och samordnas med ämneskonferensen i statistik/matematisk statistik dagen efter på samma plats. Temat för höstkonferensen blir pedagogiskt, preliminärt "datorn i undervisningen", medan ämneskonferensen kommer att koncentrera sig på kursböcker med mera.

För att hitta några av föredragen från sommarskolan och få mer information om de aktiviteter jag har talat om här och annat som vi gör hänvisas till vår webbplats. Där finns också postgironumret till Cramérsällskapet och vår nya logotyp. Om du inte redan har gjort det, betala 50 kronor och bli medlem!

ROLF LARSSON



DETTA HÄNDER 2011..

	KONFERENS	PLATS	WEBBPLATS
27 okt	Cramérsällskapets höstmöte	Uppsala	http://statistikframjandet.se/?page_id=224
1-4 nov	Strategies for Standardization of Methods and Tools – How to get there	Ottawa, Kanada	www.statcan.gc.ca/conferences/symposium2011/
8 nov	Industriell statistik höstmöte. "Kan vi tolerera det? — toleranskedjor, felfortplantning"	Södertälje	http://indstat.se/nastaseminarium.html.php
10 nov	Measurement Error	London, England	http://membership.rss.org.uk
28-31 dec	International Conference on Advances in Probability and Statistics	Hong Kong, Kina	http://faculty.smu.edu/ngh/icaps2011.html

..OCH 2012

	KONFERENS	PLATS	WEBBPLATS
1-6 jan	22nd Annual Conference of The International Environmetrics Society	Hyderabad, Indien	www.ties2012.com/
23-25 feb	5th Annual Bayesian Biostatistics Conference	Houston, USA	biostatistics.mdanderson.org/BBC2012
16-18 feb	Statistical Practice	Orlando, USA	www.amstat.org/meetings/csp/2012/
30 maj – 1 juni	European Conference on Quality in Official Statistics	Aten, Grekland	www.q2012.gr
11 – 14 juni	Fourth International Conference on Establishment Surveys	Montreal, Kanada	www.amstat.org/meetings/ices/2012
10-14 juni	Nordic Conference on Mathematical Statistics	Umeå	www.nordstat2012.se
9-14 juli	8th World Congress in Probability and Statistics	Istanbul, Turkiet	www.worldcong2012.org
16-18 juli	Longitudinal Data Analysis Subject to Outliers, Measurement Errors, and/or Missing Values	St. John's, Kanada	www.iss-2012-stjohns.ca
28 juli – 2 aug	Joint Statistical Meetings	San Diego, USA	www.amstat.org/meetings/
31 okt – 3 nov	International Conference on Methods for Surveying and Enumerating Hard-to-Reach Populations	New Orleans, USA	www.amstat.org/meetings/h2r/2012

