

STOR SPÄNNVIDD
PÅ GÖTEBORGS-
TRÄFF

SIDAN 6

NOMINERA
ÅRETS STATISTIK-
FRÄMJARE

SIDAN 7

Qvintensen

MEDLEMSTIDNING FÖR SVENSKA STATISTIKFRÄMJANDET

NR 4 2011



En kohortstudie blir självklart inte framgångsrik om man använder frågor eller datainsamlingsmetoder som gör de tillfrågade arga och irriterade.»

Marika Wenemark

SIDAN 17

I para-
psykologins
förtrollade
värld

SIDAN 20

Om termer i allmänhet
och likelihood i synnerhet

SIDAN 16

TEMA:

Bilder av vår värld

SIDORNA 8-15

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



Bild av vår värld.



Nytt brev.



Marika
Wenemark.



Inga intryck.

Qvintensen

MEDLEMSTIDNING FÖR
SVENSKA STATISTIKFRÄMJANDET

NR 4 2011

Innehåll

STATISTIKERTRÄFF I GÖTEBORG 6

NOMINERA ÅRETS STATISTIKFRÄMJARE 7

BILDER AV VÅR VÄRLD/

Förståelse för skalan 8

Vinterkonferens i rumslig statistik 12

Poverty mapping 12

On the controversy on

poverty mapping methods 15

SPRÅKET/

Om termer i allmänhet och likelihood i synnerhet 16

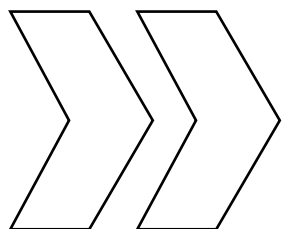
DAGS ATT TA GEMENSAMT ANSVAR

FÖR VÅRA ENKÄTSTUDIER 17

MINA ÄVENTYR I PARAPSYKOLOGINS FÖRTROLLADE VÄRLD

Del 2. 20

Kommentarer till forskningsområdet parapsykologi 23



**The poverty maps we have produced
have been the primary determinant
for allocation of over USD 1,000,000,000
of food aid alone.”**

Stephen Haslett, sidan 13

Fasta spalter

I HUVUDET PÅ EN REDAKTÖR 4

ORDFÖRANDEN HAR ORDET 5

HÄNT & HÖRT 18

KALENDARIUM 30



Våra föreningar

SURVEYFÖRENINGEN

Ordförande har ordet 25

Bortfall – seminarium med Surveyföreningen 26

Tävling om bästa uppsats 27

Europeisk forskning om urvalsundersökningar 28

FÖRENINGEN FÖR MEDICINSK STATISTIK

Ordförande har ordet 29



I HUVUDET PÅ EN REDAKTÖR

Qvintensen

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@stat.su.se

Vice ordförande

Jun Yu, 090-78 68 468

Sekreterare

Chandra Adolffsson

Pernilla Andersson

Skattmästare Jonas Selling

Klubbmästare Therese Andersson

Surveyföreningen

Joakim Malmelin, 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 Adolffsson

Statistiska centralbyrån

701 89 Örebro

E-post sekrfram@gmail.com

Webbplats www.statistikframjandet.se

Hur ser jorden egentligen ut?

I bokslukaråldern blev jag ofta enormt fascinerad över små mysterier i alla dessa science-fictionromaner jag plöjde. I en av dem betraktade någon jorden genom en värmekamera som återgav en annan bild än den vi kan se med ögat och ställde frågan "Hur ser jorden egentligen ut?"

Tittar man bredvid kameran ser den ut på ett sätt, genom kameran på ett annat sätt. Jag har läst någon annanstans att de första teleskopen mottogs med en hel del skepsis. Hur vet man att den bild man får som har transformerats genom ett intrikat system av linser är "den sanna bilden"? I det här numret har vi bilder av vår värld. Se hur många bilder man kan ha av världen, alla gjorda för något ändamål, under givna förutsättningar och begränsningar.

»Området parapsykologi tvingar fram vetenskapliga frågor.«

Kartor och andra bilder av vår värld är modeller av verkligheten. En annan modell är

den bild hjärnan formar utifrån de signaler som kommer från mätinstrumentet ögat. Stephen Haslett skriver om fattigdomsmätningar och hur modeller används i den statistiska analysen som till slut blir till kartor över fattigdom. Kartorna ligger till grund för beslut om hjälp till fattiga. En av kartorna visar att i stora områden i Bangladesh är fler än hälften av barnen extremt kortvuxna.

Parapsykologi har jag däremot blivit mindre fascinerad av genom åren. Men på ett sätt är parapsykologi fascinerande på samma sätt som värmekameran och teleskopen: hur vet vi hur världen ser ut? Området parapsykologi tvingar fram vetenskapliga frågor. Hur fungerar vetenskapen egentligen, är det inte bara ovilja från forskarna som gör att det inte kommer fram någon forskning från det hållet? Joakim Westerlund, som är forskare i psykologi på Stockholms universitet, berättar om sitt arbete och sina intryck efter en längre tids forskning om parapsykologiska fenomen.

Igår ringde en intervjuare. Jag hade inte tid att svara på några frågor. Håller vi på att köra slut på våra uppgiftslämnare? Läs Marika Wenemarks diskussion om våra attityder till dem vi är beroende av.

Ett stort tack till alla som har bidragit till detta nummer av Qvintensen!

DAN HEDLIN



FOTO: MONICA HOLMBERG

ISSN 2000-1819

Författare och andra som utan ersättning bidrar med material för publicering behåller upphovsrätten till sina verk. Svenska statistikfrämjandet äger rätten att publicera dessa verk i valfri form och upplaga samt att behålla alla eventuella intäkter från sådan publicering, om inget annat avtalas. Författare och andra som utan ersättning bidrar med material för publicering garanteras rätten att själva få sprida sitt verk under förutsättning att Qvintensen anges som originalkälla med angivande av nummer och år. Svenska statistikfrämjandet äger upphovsrätten till publicerat material då det är framställt av redaktionsmedlem, styrelsemedlem, någon till Främjandet knuten förenings styrelse eller person som erhållit arvode för sitt bidrag. © Svenska statistikfrämjandet 2011.

ORDFÖRANDEN HAR ORDET

VÅRKONFERENSEN ÄR NU PÅ GÅNG

Så här på hösten brukar planeringen av vinterns och vårens aktiviteter ligga högt upp på att göra-listan. Sol- eller skidseminstrar ska bokas, konferensresor planeras, projekt följas upp osv. När det gäller statistikområdet är det två kommande aktiviteter som jag särskilt vill föra fram och rekommendera att de tas i beaktande vid planerandet. Först ut av aktiviteterna är den 43:e upplagan av Vinterkonferensen. I år hålls den i Orsa Grönklitt den 11 – 15 mars. Vanligen arrangeras Vinterkonferensen av Statistiska och Matematiska institutionerna vid Umeå universitet. I år är dessa institutioner upptagna med att arrangera den nordiska konferensen i matematisk statistik kommande sommar. Därför har Högskolan Dalarna ryckt in och står för 2012 års arrangemang. För information om ämne för konferensen och detaljer i upplägget hänvisar jag till sidan 12 i detta nummer av Qvintensen. Jag hoppas att många statistikfrämjare möter upp och träffas på Vinterkonferensen.

Strax efter vinterkonferensen kommer nästa aktivitet och möjlighet att träffa statistikfrämjare, nämligen vår förenings vårkonferens och årsmöte. Svenska statistikfrämjandets årsmöte 2012 är planerat att hållas den 23 mars i Alnarp. Dagen före, den 22 mars, håller medlemsföreningarna sina respektive årsmöten. Styrelsen arbetar nu med att sätta ihop ett föredragsprogram till vårkonferensen. Ett sammanhållande tema för föredragen är agrokulturellt relaterad statistik. Detta är givetvis ett mycket brett område som t.ex. rymmer tolkning av satellitbilder, uppskattning av storleken av fiskbestånd, officiell jordbruksstatistik och veterinärmedicinska experiment. Det är också ett område där statistikämnet har en lång historia. En speciell inbjudan till årsmötet skickas ut till Svenska statistikfrämjandets medlemmar och annonseras på föreningens webbplats.



Hans Nyquist,
ordförande.

»Ett sammanhållande tema för vårkonferensen är agrokulturellt relaterad statistik.«

Jag ser verkligen fram emot att träffa många statistikfrämjare i Alnarp på årsmötet.

En viktig punkt på vårkonferensen är att presentera årets statistikfrämjare. Jag vill uppmärksamma alla på att det nu är möjligt att lämna förslag till mottagare av denna hedervärda titel. En blänkare om detta finns på annan plats i detta nummer av Qvintensen.

Som jag tidigare har rapporterat arbetar styrelsen nu med att utveckla medlemsregistret och i och med det även proceduren för betalning av medlemsavgifter. Arbetet bedrivs huvudsakligen i en arbetsgrupp bestående av Joakim Malmelin, Pernilla Andersson, Annika Lindblom och Mikael Möller. En första målsättning är att skapa ett gemensamt register för alla medlemsföreningarna där det skall vara förhållandevis enkelt för medlemmarna att ändra sina adressuppgifter samtidigt som det skall vara enkelt för medlemsföreningarnas registeradministratörer att sköta sin del av registret och att nå sina medlemmar för föreningspecifik information.

Vi diskuterar även möjligheten att koppla betalningsfunktioner till det nya registret. Det innebär att vi kan få möjlighet till internetbaserad betalning av årsavgiften och att de som glömmat betala automatiskt får en påminnelse härom. Om man fortsätter att glömma hamnar man automatiskt på en reservlista med mer begränsad tillgång till Svenska statistikfrämjandets information (t.ex. utebliven Qvintensen och inga jobbannonser via e-post).

Statistik är ett livaktigt ämne. I det här numret finns mycket att läsa om hur statistiska tekniker påverkar tolkningen av resultaten, något som ofta kommer bort i det enorma massmediala flödet av siffror. Vi kan också läsa om statistikens betydelse för bekämpning av fattigdom, cancerforskning och mycket annat.

STATISTIKERTRÄFFEN

– SAKNADE VÄRDEN, CANCER OCH VISAR PISA ATT SVENSKA ELEVER HALKAR EFTER?

För femte året genomförde Statistikkonsulterna tillsammans med AstraZeneca, Göteborgs universitet och FMS Statistikerträffen i Göteborg. Det är ett medvetet val att låta Statistikerträffen spänna över olika områden för att på så sätt stimulera åhörarna att upptäcka nya möjligheter och infallsvinklar oavsett var man är verksam.

Årets statistikerträff hade stor spännvid: "Missing data", "Moderna metoder för att skatta cancerpatientöverlevnad", "Kritik av analysen i PISA-undersökningen", "Att använda interaktiv visualisering som ett stöd i analys av medicinska data", "Att utnyttja redan kända gener vid associationsstudier över hela genomet" samt "Att inkludera förkunskap utan att ta till det bayesianska maskineriet".

Magnus Kjaer inledde med att tala över ämnet missing data (saknade värden) i kliniska försök och belyste det genom områdena "Varför det plötsligt blivit viktigt", "Centrala dokument och guidelines", "Rekommendationer" samt med några intressanta exempel. Rekommendationerna, som många gånger är av förebyggande karaktär, känns igen från andra tillämpningsområden i statistik vilket inte är så konstigt men de är därför inte desto mindre centrala. Intressant var också att i exemplen se vilken skillnad som uppstod mellan effekt och placebo i en studie beroende på vilken metod som anammandes. Detta om något visar på behovet av förstärkta resurser inom detta viktiga statistikområde.

Therese Andersson och Sandra Eloranta, till vardags på Institutionen för medicinsk epidemiologi och biostatistik vid Karolinska Institutet fick oss intresserade av såväl traditionellt använda metoder för att skatta populationsbaserad cancerpatientöverlevnad som metoder som har utvecklats på senare tid. Bland annat belystes vikten av att hantera konkurrerande risker i den statistiska analysen i studier som syftar till att förbättra riskkommunikation mellan patient och läkare, samt en ny metod som kan användas för att skatta den förväntat förlorade livslängden bland cancerpatienter. Vi fick också höra om metoder som kan användas för att skatta andelen statistiskt botade patienter. Att en patientgrupp

är statistiskt botad innebär att det inte längre antas finnas någon överdödlighet som kan associeras till den sjukdom som studeras, dvs dödligheten i patientgruppen har konvergerat mot dödligheten i den generella populationen. Metoderna illustrerades med hjälp av tillämpningar på prostatacancer, tjocktarmscancer samt flera hematologiska cancer typer baserade på data från vårt nationella cancerregister. Föredraget var mycket bra och uppskattades av många.

Men ändå tilldrog sig Sture Holms presentation största intresset, om kritik av PISA-undersökningen (Programme for International Student Assessment). PISA är en internationell studie som syftar till att jämföra skolelevs resultat – och därmed indirekt utbildningsväsendet – i olika länder. Undersökningen, som består av utprovade tester inom matematik, läsförståelse, naturvetenskap och problemlösning, kritiserar på flera punkter. Tidigare i år i en artikel i DN (12/4 2011, "PISA-undersökningar kritiserar") intervjuas Svend Kreiner som har hävdade att beroende på hur man räknar kan man placera Sverige på mellan fjärde och trettioförsta plats i läsförståelse!

Sture Holm tog upp två principiellt viktiga invändningar. Resultaten räknas om genom användning av Rasch-modeller vars syfte är att skapa en "standardskala" med medelvärde 100. Enligt Sture Holm är manövern helt onödig då andelen rätta svar är tillräckligt för analys och dessutom enklare att tolka. Men än värre är att av rent beräkningstekniska skäl raderas extrema värden före analys. Detta betyder att om t.ex. polska barn, pga undervisningens upplägg, får extremt bra eller dåliga resultat på en viss uppgift så stryks alla polska barns resultat från den



uppgiften. Denna strykning motiveras med att Raschanalysen bli instabil om det finns många extrema data.

Alla elever skriver inte alla delprov, och inte heller alla uppgifter inom sitt delprov - det skulle helt enkelt ta för lång tid. Men före analys imputeras data för de icke svarande eleverna på samtliga uppgifter utifrån empiriska fördelningar över de svarande eleverna. Man kan på goda grunder ifrågasätta imputation – i PISA-undersökningen sker detta i mycket stor omfattning. Men som

om inte det vore nog imputeras alla data – alltså även för de elever som faktiskt gjort provet! Om en elev svarat på en fråga så kommer svaret att ersättas med ett slumpmässigt resultat, baserat på den empiriska fördelningen över alla som svarat på denna fråga.

»PISA-undersökningen används som måttstock på skolväsendet och citeras flitigt i media och av politiker.»

PISA-undersökningen används som måttstock på skolväsendet och citeras flitigt i media och av politiker.

Då en undersökning betraktas som "gold standard" ska kraven på vederhäftighet ställas högt. Sture sammanfattade sin kritik i (1) ta inte bort data och (2) ersätt inte data med slumptal. Även detta föredrag blev mycket uppskattat, Sture framförde det med skärpa och pedagogisk briljans och det ledde till en livlig och intressant diskussion.

Pia Tormalm, SAS Institute, demonstrerade

2011



interaktiv visualisering som ett stöd i analys av medicinska data med JMP Genomics. I genomgången visades prov på olika typer av intressanta visualiseringar. Genom relativt enkla grepp kan man få fram resultat som medför stora vinster i överskådlighet.

Malin Östensson talade sedan över temat "Att utnyttja redan kända gener vid associationsstudier över hela genomet". Malin gjorde en intressant inledning med en pedagogisk introduktion till sitt område som fick åhörarnas intresse att ytterligare växa.

Dagen avslutades med Thomas Svensson från SP Building Technology and Mechanics vars ämne var "Att inkludera förkunskap utan att ta till det bayesianska maskineriet". En del hade kanske väntat sig reaktioner på detta möjligen utmanande tema, men dessa uteblev och kanske har synen bland statistiker på bayesiansk kontra frekventistisk ansats lugnat ner sig och ersatts av ömsesidig respekt och förståelse för varandras styrkor beroende på situation. Föredraget uppskattades mycket, liksom hela dagens insatser av föredragshållare och engagerade åhörare.

Kvällen avslutades på restaurang Club Avanzes med förtäring av bästa slag och många uttryckte en uppriktigt lust att träffas igen 2012.

MATS RUDHOLM OCH
MAGNUS PETTERSSON,
STATISTIKKONSULTERNA

NOMINERING TILL

Årets statistik- främjare

Till Årets statistikfrämjare utses en eller flera personer bosatta i Sverige som gjort en slagkraftig insats på statistikområdet som under det senaste året gett publik uppmärksamhet.

Insatsen kan utgöras av en innovativ metod, en smart analys, en effektiv presentation eller att på ett framgångsrikt sätt ha utmanat en etablerad sanning. Belöningen består av ett prisföremål och ett diplom. Den utvalda blir inbjuden att hålla ett föredrag på årsmötet, blir presenterad i Qvintensen och får själv skriva en artikel om sin insats i Qvintensen.

Vi inbjuder härmed alla medlemmar att ge förslag på en eller flera personer som du tycker uppfyller ovanstående kriterier. Skicka förslaget med motivering till hans.nyquist@stat.su.se senast den 27 januari 2012.

Utdelningen av priset äger rum vid Svenska statistikfrämjandets årsmöte den 23 mars 2012.

Qvintensen rättar

■ Erik Leander lever sedan många år i Ljungsbro. Vi beklagar om läsare av Qvintensens förra nummer kan ha fått något annat intryck.



Erik Leander.

Länktips

■ Chefen för US Census Bureau, Robert Groves, gisslas i the Daily Show:
<http://www.thedailyshow.com/watch/thu-april-1-2010/2010-census>

■ Varför lägga en massa tid på att skriva rapporter? Använd rapportgeneratoren istället:
<http://pdos.csail.mit.edu/scigen/>

■ Och mer seriöst, SVT Pejls
<http://blogg.svt.se/svtpejl/>
"Vårt mål är att göra statistik, fakta och siffror mer tillgängliga och begripliga för alla."



Förståelse för skalan

– EN AV GRUNDPELARN I HANTERINGEN AV GEOGRAFISKA DATA

Att rulla ut ett långt ritpapper längs Vänerns strandlinje för att avbilda Väneren på en karta faller på sin egen orimlighet. För att kunna möjliggöra avbildningen av Väneren som en tvådimensionell modell så behöver sjön förminskas i förhållande till verkligheten och därmed har en kartografisk process startats.

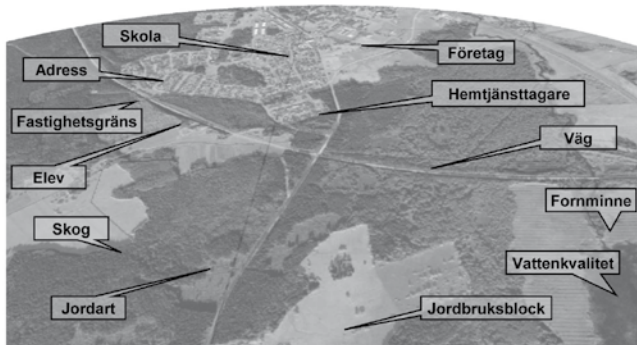


Bild 1a. Flygbild med utpekade platser för geografiska företeelser som kan behöva återges i en karta.

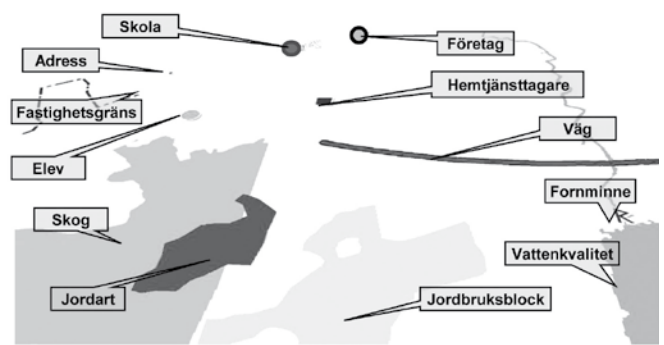


Bild 1b. En grafisk återgivning av de geografiska företeelserna i Bild 1a.

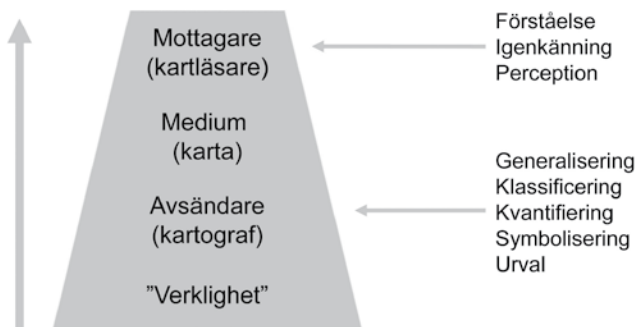


Bild 2. Den kartografiska processen.

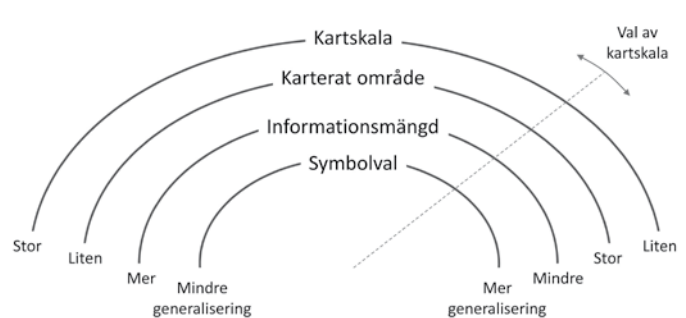


Bild 3. Förhållandet mellan kartskala och generaliseringsgrad.

På några kartserier går det att läsa: "En centimeter i kartan motsvarar 200 meter i verkligheten". Vanligare är att texten ersätts av ett uttryck med den äldre matematiska symbolen för division, i det här fallet 1:20 000. Förhållandet beskriver kartans skalfaktor. Alternativt förses kartan med ett skalstreck eller skalstock under kartan som visuellt ger läsaren en vägledning om förhållandet mellan kartbilden och verkligheten. Skalan beskriver därmed relationen mellan kartbilden och den verklighet som kartan är satt att avbilda.

Beroende på kartskala så har den som skapat kartan olika möjligheter att göra kartan läsbar. Objekt som återges i kartan har som syfte att kommunicera med mottagaren av kartan, se Bild 1a och 1b. På motsvarande sätt som vi kommunicerar med hjälp av språket så kan vi tala om en kartografisk kommunikation.

Den som tar fram en karta behöver generalisera objekten i förhållande till hur de ser ut i verkligheten. Samtidigt skall de ändå vara tydliga för mottagaren. I princip kan vi tala om ett antal generaliseringsverktyg som kartografen har till förfogande för att skapa kartan.

Ett urval av objekt från verkligheten kan anses så viktiga för budskapet i kartan

att de tilldelas egna symboler och lyfts därmed fram. Andra objekt sorteras in med flera andra i en gemensam klass. Ytterligare andra objekt bedöms mindre viktiga och tas inte med i kartbilden. Bild 2 visar hur den kartografiska processen ser ut i princip. Om verktygen används på rätt sätt kommer kartan att kunna kommunicera med mottagaren som utifrån sin egen förförståelse och perception skall kunna ta till sig kartans budskap.

Att generalisera verkligheten till en karta reducerar och förenklar den ursprungliga informationsmängden. Graden av generalisering styrs i stor utsträckning av kartskalan. Om en karta anges som storskalig är lägesnoggrannheten och detaljeringsgraden i kartan hög. Det är dock enbart mindre områden i taget som kan redovisas storskaligt för att kartan inte skall bli för stor i tryckt format. En småskalig karta är mer generaliserad och innehåller mindre information men täcker större sammanhängande områden. Bild 3 visar förhållandet mellan kartskala och generaliseringsgrad.

En del objekt som återges i kartan kan anses vara så viktiga för kartans budskap att dess storlek överdrivs i förhållande till verkligheten. Detta görs antingen för att de skall lyftas fram och synas i kartan eller för att de rent tekniskt inte kan ritas ut skalentligt. För att det mänskliga ögat skall ha en möjlighet att uppfatta objekten i en



Stefan Svanström.

FOTO: LENA BLÄSJO JANSSON

karta finns det ett antal minimistorlekar för att vanligt förekommande objekt i kartor skall vara tydliga. Se Bild 4.

En linje med bredden 0,1 mm motsvarar 5 meter i skala 1:50 000. En väg som är 0,75 mm bred på Terrängkartan (tidigare kallad Gröna kartan eller Topografiska kartan) som har skalfaktorn 1:50 000 skulle därmed vara 37,5 m bred i verkligheten. Det är den inte: den anges vara 5 meter bred. Vägar och många andra symboler på kartan är betydligt överdrivna i storlek för att de ska vara lätta att uppfatta med blotta ögat. Likaså är ytan mellan objekten i kartan viktig för att kartobjekten inte ska riskera att flyta ihop när man läser kartan.

I småskaliga kartor är det vanligt att till exempel väg och järnväg som ligger intill varandra flyttas isär för att lyfta fram dem båda som separata objekt i kartan. När symboler återges i färg i kartan kan det också påverka minimistorlekarna på

» » » »

Punkt	•	0,2 mm
Linje	—	0,1 mm
Fylld Fyrkant	■	0,4 mm
Tom fyrkant	□	0,6 mm

Bild 4. Minimistorlekar på objekt i kartan för att de skall vara synliga för ögat.

»»» FÖRSTÅELSE FÖR SKALAN

flera sätt. Måtten för minimistorlekar i Bild 4 gäller främst svarta objekt som är den mest lättlästa färgen.

Om objekten i kartan skall ha en annan färg kan storleken på symbolen behöva anpassas, alternativt kan den kontrasterande bakgrunden anpassas.

De symbolval som avsändaren har att jobba med i kartan är i sin grundutformning punkt, linje och ytor. För samtliga symboler finns det en mängd olika principer för hur objektens geometri kan förenklas, se Bild 5.

En av de vanligaste geometriska generaliseringarna är urval. Vid digitalisering av öar längs kusten i skala 1:50 000 måste ett antal kobbar och skär tas bort för att kartan skall vara läsbar för den valda skalan. Genom att förenkla reduceras antalet mätpunkter som bygger upp en linje. Det finns algoritmer som sparar de brytpunkter som bäst motsvarar linjens karaktäristiska drag (Douglas-Peuckers algoritm). Då symbolerna i kartan ofta är överdrivna i storlek kan de behöva flyttas isär för att bli tydliga. Till exempel vägar kan ritas med större avstånd från varandra än vad de har i verkligheten, vilket har skett i Bild 6 på nästa sida.

Om linjer är för kantiga kan de behöva utjämnas för att få en mjukare framtoning i kartan. Vissa objekt kan även behöva förstärkas i sin utformning för att tydlig-

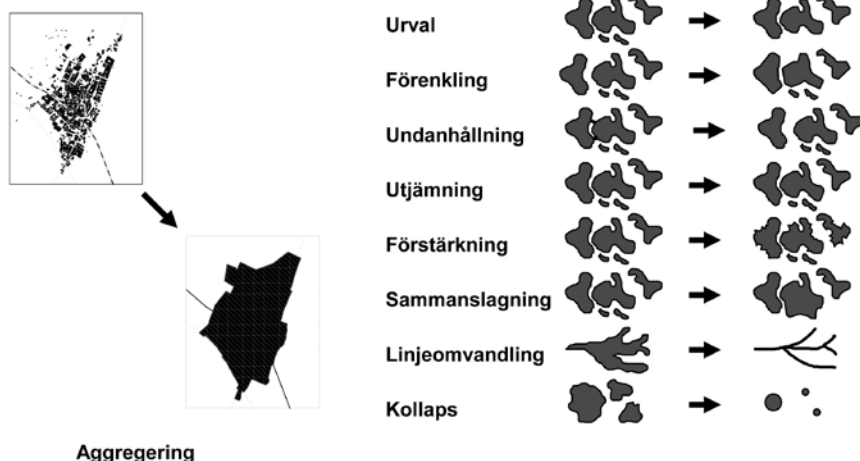


Bild 5. Exempel på geometriska generaliseringsprinciper. Förenkling innebär att kurvorna ersätts med styckevis räta linjer. Vid utjämning ersätter man till exempel de styckevis räta linjerna med splines (mjuka kurvor) för att objekten inte ska se kantiga ut.

Bild 7. Gräsö utanför Östhammars kommun återgivet i skala 1:10 000 (den fyllda gråa ytan) och skala 1:250 000 (ytan inom den svarta kantlinjen).

göras, det kan exempelvis handla om att illustrera en meandrande flod. Mindre öar som ligger nära varandra kan slås samman till ett större objekt. En vattendel med förgreningar kan i en skala vara en yta men återges som en linjedragning vid skalförändring.

När tätorter återges på en Sverige-karta är det bara storstäderna som behåller formen. Resterande tätorter har genomgått en kollaps och återges som punkter i kartan. Med aggregering reduceras ett antal mindre objekt till ett större areaobjekt som täcker motsvarande yta, till exempel kan det röra sig om byggnader i en tätort som ligger för tätt och blir svåra att uppfatta vid byte av skala.

Generaliseringsgraden är en av anledningarna till varför det uppstår svårigheter att analysera databaser där data har olika skalnivåer. Kartobjekt som kommer från småskaliga kartor riskerar att dra ner det totala resultatet för en geografisk undersökning. Det finns olika exempel på hur kartskalor i register anpassas efter flera behov. Ett sådant är SMHI:s sjöregister som innehåller runt 100 000 sjöar med en areal på över en hektar. Den vanligaste källan som används för att mäta upp arean för sjöar är Terrängkartan med skalfaktorn 1:50 000. För några av de större sjöarna återges

olika areal beroende på skalnivå och för till exempel Vänern anges både 5 519 och 5 648 kvadratkilometer.

I Bild 7 återges Gräsö utanför Östhammars kommun. Till ön kommer man normalt med bilfärja från Öregrund och ön har flera badplatser. Hur stor är då ön och hur lång strandlinje har Gräsö? Beroende på om man använder underlaget från kartan i skala 1:10 000 eller skala 1:250 000 så får vi helt olika längd på strandlinjen runt Gräsö. I skala 1:10 000 är strandlinjen 218 kilometer medan den är 137 kilometer lång i skala 1:250 000. Om den generaliserade svarta linjen i Bild 7 som anger skala 1:250 000 används för att räkna antalet byggnader (enligt Lantmäteriets fastighetsregister) inom strandskyddet (300 meter) kommer ett stort antal byggnader koordinatmässigt att ligga i vattnet, enbart beroende på de olika skalorna.

I Bild 8 illustreras hur valet av skalnivå påverkar en tematisk karta. Bilden visar andelen som röstade på Kristdemokraterna i valet 2006. I den vänstra kartan visas andelarna per valkrets, i den högra per kommun. Som synes blir intrycket av var KD är starka helt olika. De regionala variationerna suddas ut när statistiken redovisas på grövre administrativa indelningar, se till exempel Kuttainens med flera valdistrikt i Kiruna kommun.



Bild 6. Järnvägen flyttas ut från vägen när kartskalen blir mindre och det karterade området större. Källa: Geodata © Lantmäteriet.

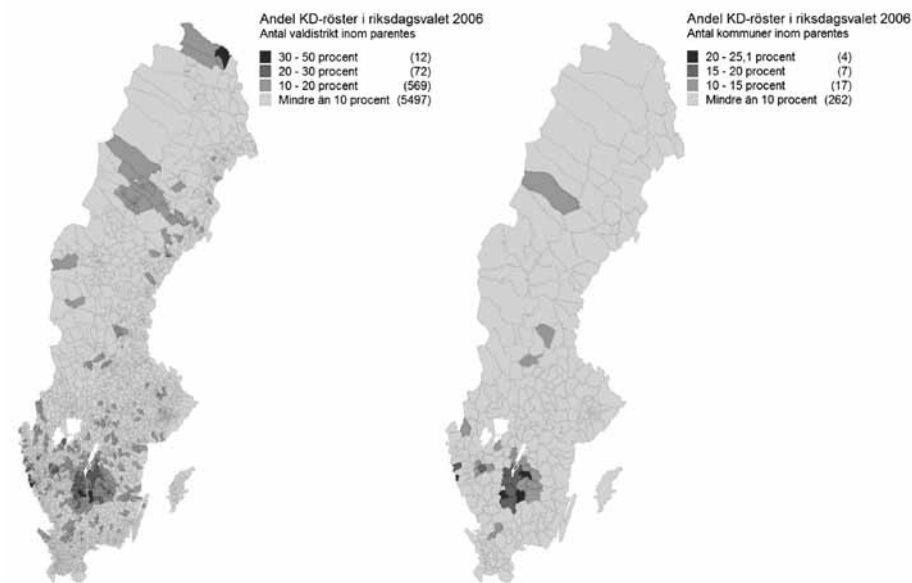


Bild 8. Skalnivåns betydelse i koropletkartor. Andel som röstade på Kristdemokraterna i valet 2006. I den vänstra kartan visas andelarna per valkrets, i den högra per kommun. På sid 32 finns bilden i färg och större version.



Bild 9. En oljefläck söder om Hoburgen som ser stor ut i förhållande till Gotland.

Att symboler måste överdrivas i storlek kan i en del fall ge intryck av att de verkliga objekten är större än vad de är. Man talar om att kartan har en reduktionsdistorion när detta uppträder. Vi kan ta ett exempel där kartan är ett stöd till en artikel om oljeutsläpp i farvattnen söder om Gotland. I artikeln går det att läsa att utsläppet är 13 kilometer långt och 75 meter brett. Detta illustreras med en karta där utbredningen av utsläppet ritas in på platsen för händelsen (Bild 9). Om oljeutsläppet avbildades i skalenlig storlek skulle det inte synas på kartan. Alltså ritas den med en flera gånger större utbredning, just i det här fallet genom att bredden ökas. Ett alternativ hade kunnat vara att illustrera utsläppet med en symbol som inte ger mottagaren intrycket att det som visas är oljeutsläpets verkliga utbredning. Till exempel

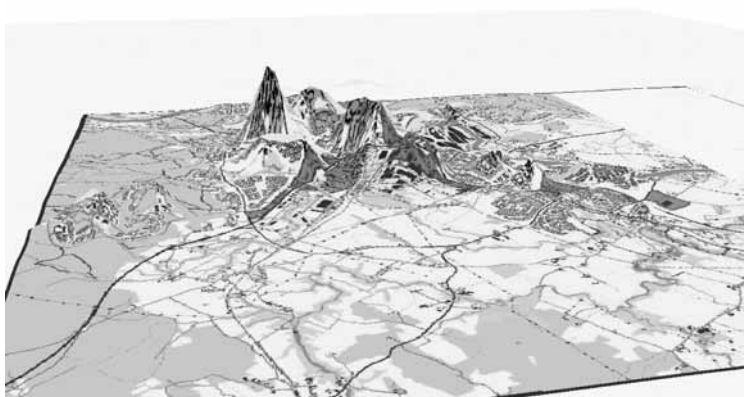
kan platsen markeras med en punkt-symbol eller kryssmarkering i kartan.

Ytterligare ett exempel på när man behöver vara medveten om kartskalorna är när en kommuninvånare ringer en miljöhandläggare i kommunen och frågar om möjligheterna att koppla in bergvärme i sin enplansvilla. Handläggaren kan då söka efter fastighetsbeteckningen i ett ärendehanteringssystem med kartstöd. Sedan kompletteras den kartan med tidigare beviljade ansökningar för bergvärme. Därefter kompletteras den med vattenskyddsområden för att se om tomten ligger innanför eller utanför inre och yttre skyddszon. Ibland kan det även vara bra att konsultera jordartskartan för att få en bild av vad det är för markförutsättningar på fastigheten. I det här fallet behöver handläggaren utnyttja flera olika underlag: Fastighetskartan med

ursprung Lantmäteriet (skala 1:10 000), tidigare bergvärmeborrningar med ursprung från kommunens eget register (skala 1:10 000 eller bättre), vattenskyddsområden med ursprung Länsstyrelsen (beroende på kartering skala 1:10 000 eller 1:50 000) och jordartskartan från Sveriges geologiska undersökning (skala 1:50 000).

Osäkerheten i svaret beror också på var fastigheten finns belägen i förhållande till övriga informationsmängder. Om fastigheten ligger på randen till ett skyddsområde eller på randen mellan flera jordarter blir osäkerheten i svaret större.

STEFAN SVANSTRÖM,
STATISTISKA CENTRALBYRÅN



Karta över befolkningens fördelning samt markanvändning i Borlänge skapad med GIS-verktyg. Höjden i kartan anger intensiteten i pendling mellan hem och arbete.

»The poverty maps we have produced have been the primary determinant for allocation of over USD1,000,000,000 of food aid alone.»

Vinterkonferens i rumslig statistik i Dalarna

Altsedan Mercator ritade sin världskarta 1567 har visualisering av rumsliga modeller haft stor betydelse för beslutsfattande på högsta politiska nivå. Den låg till grund för Portugal och Spaniens uppdelning av den "okända" världen under 1500-talet.

Verktyg för att visualisera digitala kartor har utvecklats kraftigt under de senaste två årtiondena med hjälp av geografiska informationssystem (GIS). Samtidigt har insamling av data förenklats drastiskt med GPS. Ett verktyg som används inom flera fält för att skapa kartor av hög kvalitet är ArcGIS. Programmet är mycket bra för visualisering men saknar flera verktyg för djupare statistisk analys. Men i R är det mesta möjligt. Det finns olika R-paket (t ex `sp`, `maptools`, `spdep`) med vars hjälp man kan jobba med data direkt exporterade från ArcGIS (filer i `shp`-format). En bra bok som ingående beskriver sådana R-paket är skriven av Roger Bivand et al. (Applied Spatial Data Analysis with R. Springer 2008). I boken beskrivs flera R-paket på ett lättillgängligt sätt.

Utvecklingen av bayesiansk statistik för att analysera rumsliga data har också varit stark men analyser av stora datamängder har begränsats av att beräkningsmetoder som behövs i bayesiansk statistik, såsom MCMC, är så beräkningskrävande. Håvard Rue med

medarbetare har utvecklat INLA (Integrated Nested Laplace Approximations) där rumslig modellering kan ske på ett beräkningseffektivt sätt med bayesiansk statistik. Nu finns också en R-version av INLA (www.r-inla.org).

Men insamlandet av stora datamängder innebär inte automatiskt ökad kunskap. Ett grundproblem kan vara dålig kvalitet på data eller utebliven försöksplanering. Själva modelleringen påverkas också av kunskap om rummets betydelse och val av skala för analys. För att få ut det mesta möjliga ur de fantasiska möjligheter som kombinationen GIS och GPS kan innebära behövs ett samarbete över ämnesgränser, t ex statistik, datateknik och kulturgeografi.

En vinterkonferens i rumslig statistik arrangeras av Högskolan Dalarna i Orsa Grönklitt 2012, 11-15 mars. Vi som ansvarar för planeringen är Lars Rönnegård (statistik), Pascal Rebreyend (datateknik) och Johan Håkansson (kulturgeografi). Vi har försökt lägga upp ett program med både teori och tillämpningar. Inbjudna föreläsare är Roger Bivand, Håvard Rue och Konstantin Krivoruchko (ansvarig för utveckling av statistiska verktyg i ArcGIS). För program och registrering (senast 15 januari) se www.du.se/spatialstatistics.

LARS RÖNNEGÅRD, JOHAN HÅKANSSON OCH PASCAL REBREYEND, HÖGSKOLAN DALARNA

In 2001, I held a Senior Fellowship on small area estimation at the Bureau of Labor Statistics in Washington DC. While I was there, BLS was approached by the Regional Headquarters of the United Nations World Food Programme (WFP) in Bangkok about undertaking small area estimation of poverty (also known as poverty mapping). Being outside BLS's brief, Dr John Eltinge, since 2004 Associate Commissioner for Office of Survey Methods Research at BLS, suggested WFP talk with me instead. Over the last ten years, the statistics group at Massey University in New Zealand has carried out poverty mapping exercises, or feasibility studies for them, in Bangladesh, Nepal, Cambodia, Timor-Leste and Bhutan for WFP, and in the Philippines for the World Bank.

The UN's Millennium Development Goals have as two targets to "halve the proportion of people whose income is less than one dol-

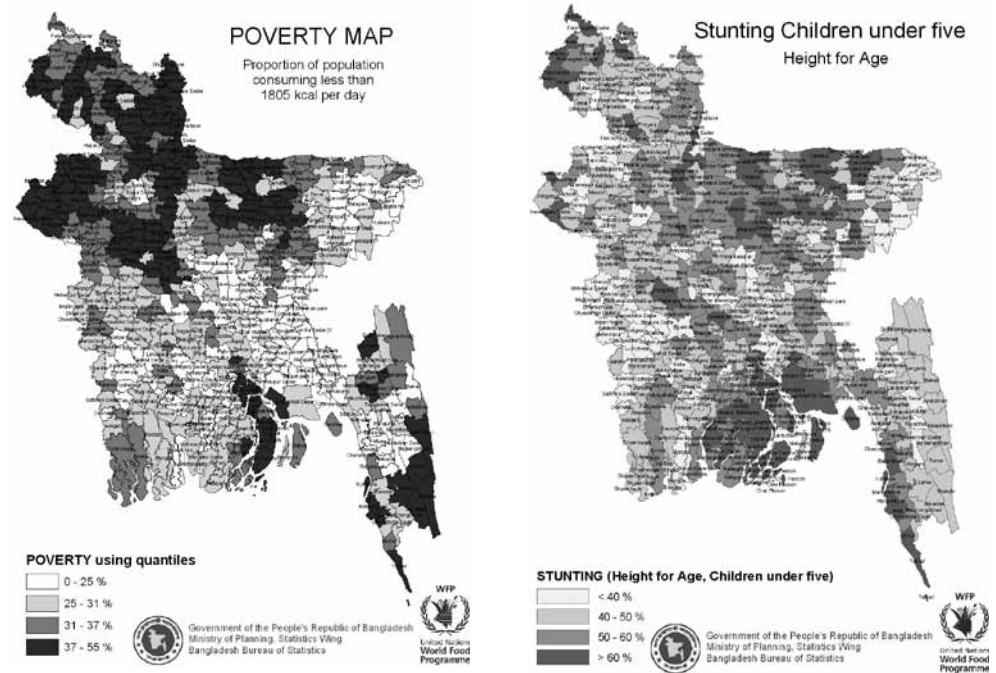


Figure 1: Expenditure poverty. Proportion of the population below the lower poverty line in Bangladesh 2003. På sid 31 finns bilden i färg och större version.

Figure 2: Stunting in children under five years of age in Bangladesh 2003. På sid 31 finns bilden i färg och större version.

POVERTY MAPPING

lar a day” and to “halve the proportion of people who suffer from hunger” between 1990 and 2015. To monitor these aims sound measures are required, and this is the reason that in Bangladesh and Nepal alone since 2006, the poverty maps we have produced have been the primary determinant for allocation of over USD1,000,000,000 of food aid alone. A high level of aid funding is also allocated via poverty mapping in the Philippines. Because the fine level estimates from small area estimation methods are the basis for the aid allocation, the underlying statistical methods and the estimates from them and their accuracy measures are central.

Poverty mapping is formally the name given to the maps produced from small area estimation of expenditure or income poverty. The poverty map for Bangladesh in 2003 in Figure 1 is an example. The technique has also been extended to kilocalorie consumption, and to stunting, underweight and wasting in children under five years of age. All child-based measures are referenced to the lowest

five percentile in a standard population, and based on low height-for-age (stunting), weight-for-age (underweight) and weight-for-height (wasting), respectively. Rates in third world countries are high, 50% and 48% in Bangladesh for poverty incidence and underweight, respectively. Importantly, as Figure 2 shows, the geospatial patterns for the poverty and stunting measures are not identical, so that using a single poverty measure would obscure important information even if it were to combine information from several particular measures.

Expenditure, income, and height, weight and age in young children are measured in national sample surveys, but not in censuses.

One important advantage of third world surveys is their high response rates – we have never had to deal with a survey with under a 95% response rate, and this is despite the considerable questionnaire length of some surveys. The National Living Standards Survey (NLSS-II, 2002/3) questionnaire in Nepal, for example, is over 70 pages long. Our experience is



Stephen Haslett.

echoed by that of the Demographic and Health Surveys (DHS) internationally – see for example Vaessen, Thaim and Lê (2005) who note that one reason for these very high response rates is “high level of co-operation on the part of potential respondents”.

The poverty mapping developed by Elbers, Lanjouw and Lanjouw (2003), commonly abbreviated to ELL, are used extensively in many third world countries. The ELL modeling procedure is econometric and was in 2003 a considerable advance in poverty mapping, a fact that deserves better recognition among statisticians. It has been implemented through Peter Lanjouw at the World Bank in the freely available PovMap software. Because the World Bank has a considerable part in allocating aid, and aid provision is often predicated by a poverty mapping exercise, the ELL method has now been applied in over 50 countries with various levels of success. The World Bank, as well as other agencies, has also run courses on ELL-based poverty mapping for third world statistical agencies. Even when ELL

»»»

»»» POVERTY MAPPING

has not been used, academics involved have often preferred to send their PhD students to third world countries to carry out the poverty mapping exercise. In our experience both are unwise strategies, as the fitting of models requires rather better understanding of the data, the relevance of the data collection mechanisms, and the problems of modeling such data than these strategies usually allow. We have invariably found that it needs local expertise as well as extensive modeling experience to get sound models and hence to get good aid allocation. Being able to produce a map and estimated standard errors using PovMap (or indeed any software) is, in itself, not sufficient.

„Methodology is interesting, but not as important as the life improvement that we all hope will be the consequence of better and fairer aid allocation to the world's poor.“

Problems may include: survey and census too far apart in time, differences in English wording of translated questionnaires not reflecting the questionnaire as implemented, problems with matching administrative area codes for survey and census, fitting separate models to subsets of the survey data, automatic fitting of regression models (especially where the number of variables including interactions approximates the sample size or subset size), use of the incorrect regression procedures and standard error estimates, using the percentage of variance explained by the regression model as justification for good model fit, poor estimation of variance components in mixed models, and not testing for (and where necessary fitting) an area level random effect. A number of these issues are not properly dealt with in ELL, and there can be additional complications where local knowledge is ignored or downplayed or where statistical modellers lack sufficient theoretical or practical experience. Over-fitting of models is relatively common. For further details see Haslett and Jones (2005; 2010).

Diagnostics for small area models are important, and need further research. However one useful and quick general diagnostic is

to look at the final maps to see how many adjacent areas have similar estimates (usually indicated by having the same colour across many adjacent small areas on the map). Unless models have involved spatial smoothing across the small areas (and this is not usually recommended because it can too easily lead to overconfidence in the estimates and maps), there is no reason on the basis of the small area model itself that this should be so. Thus large patches rather than a fine mosaic over the country as a whole is a very good, though not a sufficient reason for confidence in the model fitting, prediction at small area level, and even in the estimated standard errors for the small area level estimates themselves. Of course, it is also essential to recheck with local counterparts that such smoothness reflects inherent similarities, just as any smoothness anomalies should reflect known disjuncts in local conditions. Maps, while extremely useful, can however be misleading, since they can reflect not only the size of areas mapped but also the cutoff points used for the changes in colour level. Experimenting with different cutoffs and levels of aggregation to assess stability of what is mapped is strongly recommended. Even if comparatively objective measures like quantiles for cutoffs and administrative boundaries for areas are used, maps are not a substitute for detailed tabulations of small area estimates and their estimated standard errors. They are a graphical aid to our understanding of the information conveyed.

We are currently carrying out the poverty mapping exercise in Cambodia (including poverty incidence, gap, and severity, plus stunting, underweight and possibly wasting in children). This should be completed during 2012. We cooperate with the Cambodian national statistical office, the National Institute of Statistics <http://www.nis.gov.kh/>, which receives technical support from Statistics Sweden. Each time we carry out one of these studies we learn more and make improvements to our methodology. This is interesting, but not as important as the life improvement that we all hope will be the consequence of better and fairer aid allocation to the world's poor.

STEPHEN HASLETT,
MASSEY UNIVERSITY, NEW ZEALAND

Bibliography

- Elbers, C., Lanjouw, J. O., and Lanjouw, P. (2003) Micro-level estimation of poverty and inequality. *Econometrica* Vol 71, pp 355-364.
- Haslett, S., Isidro, M. and Jones, G. (2010) Comparison of survey regression techniques in the context of small area estimation of poverty, *Survey Methodology*, Vol 36, No 2, 157-170.
- Haslett, S. and Jones, G. (2004a) *Estimation of local poverty in the Philippines*, November 2004, Published by the World Bank and Philippines National Statistics Co-ordination Board, 70 pages.
- Haslett, S. and Jones, G. (2004b) *Local estimation of poverty and malnutrition in Bangladesh*, May 2004, Published by United Nations World Food Programme and Bangladesh Bureau of Statistics, 43 pages.
- Haslett, S. and Jones, G. (2005) Small area estimation using surveys and censuses: some practical and statistical issues, *Statistics in Transition*, Vol 7, No 3, 2005, pp 541-555.
- Haslett, S. and Jones, G. (2006) *Small area estimation of poverty, caloric intake and malnutrition in Nepal*, September 2006, Published by the World Food Programme, the World Bank and the Nepal Central Bureau of Statistics, 218 pages, ISBN 99937018-5.
- Haslett, S. and Jones, G. (2010) Small area estimation of poverty: the aid industry standard and its alternatives, *Australian and New Zealand Journal of Statistics*, Vol 52, No 4, 341-362.
- Haslett, S., Jones, G., Noble, A. and Ballas, D. (2010) More for less? Using existing statistical modeling to combine existing data sources to produce sounder, more detailed, and less expensive Official Statistics, *Official Statistics Report Series*, Vol 3, <http://www.statisphere.govt.nz/official-statistics-research/series/2010/page1.aspx>, ISSN 1177-5017, 75 pages.
- Hoogeveen, J., Emwanu, T., and Okwi, P. (2003) *Updating small area welfare indicators in the absence of a new census*, Preprint.
- Isidro, M. (2010) *Intercensal updating of small area estimates*, unpublished PhD Thesis, Massey University, New Zealand, <http://muir.massey.ac.nz/handle/10179/2053>
- Isidro, M., Haslett, S. and Jones, G. (2010a) Comparison of intercensal updating techniques for local level poverty statistics, *Proceedings of Statistics Canada Symposium 2010*, 10B-4, Social Statistics: The Interplay among Censuses, Surveys and Administrative Data, <http://www.statcan.gc.ca/conferences/symposium2010/abs-res-eng.htm>.
- Isidro, M., Haslett, S. and Jones, G. (2010b) Extended structure-preserving estimation method for updating small-area estimates of poverty, *Proceedings of the American Statistical Association 2010: Section on Government Statistics*, Joint Statistical Meetings, Vancouver, Canada, pp.1795-1807, <http://www.amstat.org/meetings/jsm/2010/onlineprogram>.
- Jitsuchon, S. and Lanjouw, P. (2005) *Updating poverty maps: emerging experience with available methods and data from Thailand*, Preprint.
- Lanjouw, P. and van der Wiede, R. (2006) *Determining changes in welfare distributions at the micro-level: updating poverty maps*, Powerpoint slides, NSCB/WB Workshop for Intercensal Updating Project.
- Molina, I. and Rao, J. N. K. (2010) Small area estimation of poverty indicators, *Canadian Journal of Statistics*, Volume 38, Issue 3, pp 369-385.
- NSCB (2009) 2003 City and municipal level poverty estimates, NSCB, Philippines.
- Rao, J. N. K. (2003) *Small Area Estimation*, Wiley Series in Methodology, New York.
- Tarozzi, A. and Deaton, A. (2009) Using census and survey data to estimate poverty and inequality for small areas, *The Review of Economics and Statistics*, Vol 91, No 4, pp 773-792.
- Vaessen, M., Thaim, M. and Lê, T. (2005) The Demographic and Health Surveys, Chapter XXII, in *Household Sample Surveys in Developing and Transition Countries*, United Nations, ISBN 92-1-161481-3, retrievable from http://unstats.un.org/unsd/hhsurveys/pdf/Chapter_22.pdf
- Zhao, Q. (2006) *User manual for PovMap*, The World Bank <http://siteresources.worldbank.org/IN-TPGI/Resources>.



On the controversy on poverty mapping methods

The small area methods of Rao (2003) are extensive but do not cover or even mention the standard method of poverty mapping developed by Elbers, Lanjouw and Lanjouw (2003), commonly abbreviated to ELL (see Bibliography on p 14). The ELL method, rather than simply fitting a linear or non-linear mixed model to survey data as in Rao (2003) also applies the model to census data, making predictions for every household based on the survey model and summing these predictions to small area level. By predicting the various random effects in such models, and repeating this process many times for every census household, standard errors conditional on the model being correct can also be estimated. Molina and Rao (2010) have extended the methods in Rao (2003) to include such predictions via censuses, but they have only applied it for the special case where there is some sample in every small area, about 30 in the case of Spain. When there is no sample in a good proportion of small areas (as is the case for small area estimation for the approximately 1000 ilaka in Nepal, and is almost invariably the case for small area estimation of poverty in the third world), the Molina and Rao (2010) method reduces to a synthetic estimator very similar to ELL but with some relatively minor though very important changes. The required modifications

are detailed for example in the earlier papers of Haslett and Jones (2010) and Haslett, Isidro and Jones (2010). See also Tarozzi and Deaton (2009).

ELL-type methods have stronger data requirements than those for the Rao (2003) methods. Survey and census must be contemporaneous, and survey-based models must only include variables collected (and equivalent) in both survey and census, although contextual variables such as area or cluster-level means can also be added. The ELL methods are also more sensitive, and this, together with some modeling and related issues in the original ELL method that are listed below, has led to strong criticism from Jon Rao and others that the ELL method as originally formulated underestimates standard errors for its small area estimates. ELL also has strong links to spatial microsimulation methods used by geographers and policy planners (see Haslett, Jones, Noble and Ballas 2010 for details of this comparison).

The ELL method is a unit-level (e.g. person, child, household or whatever the unit is) model, that is, the model is fitted at the level of the smallest unit. This has two major advantages. Firstly, it allows differentiation of areas with similar poverty levels where some have a high proportion of people near the borderline and other

areas do not. (The first type should have a higher estimated standard error than the second.) Further providing the model is correct it usually gives sound but much smaller standard error estimates than area-level models, which are fitted at a higher hierarchical level. That the model is correct is a critical issue however. Poor model fitting and lack of understanding of the critical points or issues in the data, data collection and analysis can lead to poor small area estimates of poverty, whether these be poverty incidence, gap and severity via log expenditure or log income modeling, or measures of child poverty, as well as to concomitant problems with underestimated standard errors.

Area-level models are however more easily implemented when there is a repeated survey but no contemporaneous census for every repetition. Nevertheless, methods do exist for updating ELL-type estimates given a new survey. The most promising of these, to judge a validation exercise in the Philippines and a direct comparison with other available updating methods, is outlined in Isidro (2010) and in summary form in Isidro, Haslett and Jones (2010 a and b).

STEPHEN HASLETT,
MASSEY UNIVERSITY, NEW ZEALAND

Det finns termer och det finns termer. En del består av enkla ord: hus, grafit, andra av sammansättningar: anhörigvårdare, pallborrningsaggregat, stödreaktion. En tredje kategori är de som består av flera ord: god man, byggnad med särskilt bevarandevärde, tvåkupig plankorrigerad takpanna.

Om termer i allmänhet och likelihood i synnerhet

Termer kan också utgöras av förkortningar, som t.ex. aids, laser, las. Det finns termer som är av fornnordiskt ursprung och andra som bygger på utländska morfem eller är inlånade helt och hållet från ett annat språk: Hus är ett inhemskt ord, ibland också kallat arvord, som antas ha funnits i språket sedan mycket länge och som har bildats oberoende av påverkan från andra språk. Grafit (ett mineral som består av kristallint kol) kommer av grekiskans grafein som betyder skriva. Skanner är inlånat och lätt anpassat till svensk stavning.

Vad är det då som gör att en term är bra och en annan dålig? Ja, ursprunget är bara ett av kriterierna att ta i beaktan. Och för att kunna göra en bra bedömning måste man gå systematiskt till väga.

Kortfattat kan man säga att systematiskt terminologiarbete handlar om att å ena sidan identifiera och precisera begreppet i en definition och att å andra sidan välja en lämplig term för detta begrepp. Arbetet rör sig mellan dessa två sidor: fånga innehållet i begreppet och bestämma det språkliga uttryck – den term – som blir mest lämpligt för fackområdet i fråga.

Det finns en rad kriterier att ta ställning till när man väljer en lämplig term bland flera möjliga.

Sannolikhet, trolighet eller likelihood.

En lämplig term ska vara:

- Etablerad i fackexperters språkbruk, ibland också internationellt.
- Anpassad till svenska språkets struktur.
- Entydig, dvs. inte förekomma i flera sammanhang med olika betydelser.
- Genomsynlig, dvs. man ska kunna förstå det bakomliggande begreppet bara genom att titta på termen.
- Precis.
- Språkekonomisk, dvs. en kortare term är bättre än en längre.
- Inte missvisande.

Det går inte att tillämpa samtliga kriterier i alla termvalssituationer, utan man måste väga dem mot varandra. Många termer är till exempel inte genomsynliga, andra går inte att korta av. De flesta är inte helt entydiga, utan tvärtom kan samma språkliga uttryck förekomma såväl inom ett och samma fackområde som i olika fackområden men med olika betydelser.

Den statistiska termen likelihood förekommer i svenskan i statistiska och matematiska sammanhang och den är väl etablerad. Det finns också sammansättningar som likelihoodkvot och maximum-likelihood-metoden.

Kriteriet "etablerat i fackexperters språkbruk" är normalt ett tungt vägande skäl till att behålla termen som den är, bra eller

dålig. Likelihood är inte anpassad till svensk stavning eller böjning, och många tvekar säkert en stund innan de skriver likelihooden i bestämd form. Det går då lika bra att skriva likelihoodfunktion, om man avser funktionen och inte ett värde på funktionen. För den som är kunnig i engelska kan termen vara genomsynlig, kanske också precis. I sitt statistiska och

matematiska sammanhang är den entydig och inte missvisande.

Men den viktigaste frågan som man bör ställa då man utreder termerna likelihood, sannolikhet och trolighet är om det inom statistikområdet finns en samsyn om hur dessa termer förhåller sig till varandra, och hur de i sin tur förhåller sig till engelskans likelihood och probability. Det är lika viktigt att ta reda på om det finns en samsyn vad gäller förhållandet mellan de engelska termerna.

I en svensk standard [1] finns likelihood med definitionen "sannolikhetsfunktions eller frekvensfunktions värde för en given mängd av observerade värden". Som synonym anges trolighet.

En viktig fråga för den som utreder termerna är om den engelska termen likelihood används om samma begrepp som den svenska termen likelihood. Andra viktiga frågor att ställa är om man har behov av att inom statistiken uttrycka två begrepp eller ett. Behöver man två måste man också ta ställning till om de bäst benämns med sannolikhet och trolighet eller sannolikhet och likelihood eller något annat lämpligt termpar. Har man bara behov av att uttrycka ett begrepp så måste man ta ställning till vilken av de tre existerande termerna som man bör välja.

Det är också viktigt och bra att slå fast om man väljer mellan sannolikhet och likelihood mer beroende på kontext än på deras definitioner.

Det är inte termen likelihood i sig som behöver vara ett problem. Det är ett mycket större problem om begreppet bakom termen är oklart och inte preciserat.

ANNA-LENA BUCHER
TERMINOLOGICENTRUM TNC
WWW.TNC.SE

Referenser

- [1] Statistik - Terminologi och beteckningar Svensk standard, (SS 014201). SIS.



Dags att ta gemensamt ansvar för våra enkätstudier

Våren 2005 sökte jag till en forskarskola vid Institutionen för medicin och hälsa vid Linköpings universitet med en idé om att arbeta med respondentens perspektiv i hälsorelaterade enkätstudier. Idén väckte stor förundran. Enkätstudier är ju varken nytt eller särskilt spännande. Eller som en av mina doktorandkollegor uttryckte det: "Det är ju inte ens en bra metod".

Men ju mer jag läste och ju mer jag pratade med olika personer som arbetar med enkätstudier, desto mer övertygad blev jag att enkätmetodik är ett eftersatt forskningsfält i Sverige. Speciellt om man tar hänsyn till hur många enkätstudier som görs.

Fler studier men färre som vill svara

Det genomförs ett stort antal enkätstudier varje år. Jag har också en känsla av att många studier blir mer och mer omfattande för att resultaten ska kunna brytas ner på mer detaljerade redovisningsgrupper. Bara den nationella patientenkäten omfattade år 2010 ett urval av 518 976 personer. Arbetsmiljöverket genomför arbetsmiljöenkäter, CSN frågar studenter om deras situation, Diskrimineringsombudsmanen ställer frågor om allmänhetens inställning till diskriminering, Energimyndigheten skickar enkäter till småhusägare, Försäkringskassan genomför kundundersökningar och Högskoleverket genomför studien Studentspegeln. Sedan kom jag inte längre i min myndighetslista som består av ytterligare ett 60-tal myndigheter. Jag kom inte heller till enkätstudier som genomförs vid universitet, landsting, kommuner och företag. Jag har inte gjort någon samlad översyn av svarsfrekvenser i enkätstudier men bland dem jag sett redovisar de allra flesta sjunkande svarsfrekvenser. I takt med att andelen svarande minskar, ökar ansträngningarna att få fler



Marika Wenemark.

att svara genom allt intensivare uppföljningsmetoder. Det verkar råda stor samstämmighet kring metoderna att höja svarsfrekvensen – att skicka fler skriftliga påminnelser och att komplettera med påminnelser eller intervjuer per telefon.

Vår gemensamma kohort

Hur reagerar människor på dessa allt mer intensiva metoder för att höja svarsfrekvensen? Negativa reaktioner bland dem som tillfrågas skulle ju kunna leda till mindre vilja att delta i framtida studier. Om var och en av oss som genomför enkätstudier bara tänker kortsiktigt för att få upp svarsfrekvensen i vår egen studie bidrar vi ju själva till ett allt sämre undersökningsklimat i befolkningen. En kohortstudie blir självklart inte framgångsrik om man använder frågor eller datainsamlingsmetoder som gör de tillfrågade arga och irriterade. I läroböcker om longitudinella studier står det därför betydligt mer om respondenternas perspektiv eftersom forskaren ska kunna återkomma till samma personer med nya frågor. Men Sveriges befolkning är inte en oändlig population. Tänk om vi kunde se på Sveriges befolkning som vår gemensamma kohort som vi alla måste vara väldigt rädda om. Då blir respondenternas perspektiv lika viktigt i alla studier. Det är dags att ta ett gemensamt ansvar för att inte förbruka vår gemensamma kohort.

Befintliga verktyg för bättre enkätstudier

Det finns redan många bra verktyg tillgängliga som vi skulle kunna utnyttja bättre. Det mest självklara är kanske traditionell enkätmetodik. Det handlar till exempel om att ställa en fråga i taget och att undvika dubbla negationer. Trots det används många frågor med uppenbara brister för att de till exempel ingår i en etablerad skala eller för att man vill jämföra med tidigare studier. Men den etablerade skalan blir ju tämligen värdelös om ingen vill svara på den och att jämföra över tid får tveksamt värde om begreppen som används är otidsenliga eller till och med betyder något annat för den nuvarande befolkningen. Många gör det lätt för sig och använder instrument som redan är validerade. Frågan är bara om ett instrument som validerats i ett annat land 20-30 år tillbaka i tiden kan anses validerat för att användas i Sverige idag.

Ett verktyg som vi skulle kunna använda mer är responspsykologi. Responspsykologi handlar om hur respondenten gör när han eller hon besvarar våra frå-



HÄNT OCH HÖRT

■ **THOMMY PERLINGER** vid Statistiska institutionen i Uppsala har tilldelats UPS (Uppsala Politicesstuderande) pedagogikpris 2011.

■ **ELISABET BORG** vid Psykologiska institutionen utsågs till Årets lärare 2011 vid Stockholms universitet för sin lustfyllda och entusiasmerande undervisning i statistik. Hon skapar ett kreativt lärandeklimat genom att kombinera höga förväntningar med stor tilltro till studenternas förmåga att lära sig på djupet och arbeta hårt. Elisabet tillgängliggör även området genom att utarbeta kurslitteratur och kursmaterial.

■ **ANNIKA TILLANDER** vid Statistiska institutionen, Stockholms universitet, försvarade den 26 oktober sin licentiatavhandling *Supervised Classification in a High-Dimensional Framework*. Diskutant var professor Timo Koski, KTH i Stockholm.

Har du flyttat?

■ Meddela din nya adress genom att skicka ett e-brev till Chandra Adolfsen och Pernilla Andersson på sekrfram@gmail.com.

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 bli medlem i Svenska statistikfrämjandet skicka ett e-brev till Chandra Adolfsen 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.

»»» DAGS ATT TA GEMENSAMT ANSVAR FÖR VÅRA ENKÄTSTUDIER

gor. Jag är förvånad över att många som arbetat lång tid med enkätstudier i Sverige inte känner till detta kunskapsfält. Kanske har det att göra med att det finns väldigt lite skrivet om det på svenska? Jag tror att responspsykologi kan användas mycket mer för att konstruera frågor som blir så smidiga som möjligt att besvara för respondenterna. När jag föreläser om responspsykologi blir många överraskade över att få se sina frågor från ett nytt perspektiv och häpna över att de aldrig reflekterat över svarsprocessen för sina frågor. Ibland kan man med ganska enkla medel förbättra frågorna så att de blir betydligt lättare att besvara.

Det sista verktyget jag vill slå ett slag för är kognitiva intervjuer för att pröva ut frågor och frågeformulär. Kognitiva intervjuer har snabbt blivit en av standardmetoderna internationellt. (Se även Andreas Perssons och Fredrik Schefers artikel i Qvintensen [1].) Intervjun ger en inblick i respondentens svarsprocess och det är ofta mycket lärorikt för enkätkonstruktören. Det som vi uppfattar som inkonsekventa svar får ofta en helt naturlig förklaring när man förstår hur respondenten tänkte. Personligen tycker jag att det är synd att avgränsa intervjuerna till den kognitiva processen. Jag skulle hellre vilja kalla metoden för respondentintervjuer och låta intervjun även omfatta respondentens uppfattning om till exempel frågors relevans och formulärets helhet såväl som tankar och känslor frågorna väcker.

Minskar vi motivationen genom att fokusera på de som inte vill svara?

I min avhandling [4] har jag använt dessa redan etablerade verktyg men jag har också inspirerats av andra forskningsfält. Det finns mycket forskning kring anledningarna att inte vilja delta i enkätundersökningar men betydligt mindre forskning om respondenterna som faktiskt deltar. Jag blev nyfiken på vad som bidrar till respondentens motivation i enkätstudier. Då tänker jag inte enbart på motivation för att gå med på att delta, utan också för att bli engagerad och besvara frågorna noggrant. Motivation är ett stort forskningsfält som tillämpats inom till

exempel utbildningsforskning och idrottsforskning. Jag har i min avhandling valt att använda en motivationsteori som kallas self-determination theory (SDT), se Bild 1. Enligt SDT kan motivation variera från amotivation (avsaknad av motivation) till intern motivation. Intern motivation är förknippad med starkare engagemang, bättre resultat och mer positiva erfarenheter. Enligt SDT kan motivationen öka eller minska beroende på vilka stimuli som används. Om man till exempel ger ett barn pengar för att läsa sina läxor kan barnets interna motivation minska och studieresultaten på sikt försämrast. De flesta faktorer vi normalt använder för att höja svarsfrekvensen (till exempel belöningar eller påminnelser) stimulerar den lägsta formen av motivation. Enligt SDT skulle en studiedesign som anpassas till de som har en väldigt låg motivation kunna göra deltagandet mindre intressant och minska motivationen för de som redan har en högre motivation.

»Om man ger ett barn pengar för att läsa sina läxor kan barnets interna motivation minska och studieresultaten på sikt försämrast.»

SDT specificerar också faktorer som är förutsättningar för att stimulera intern motivation. I en av studierna i avhandlingen har vi försökt att stimulera sådana faktorer i studiedesignen [2]. Resultaten talar för att det är möjligt att öka svarsfrekvensen i enkätstudier på ett sätt som samtidigt bidrar till en god

datakvalitet och positiva erfarenheter för respondenterna. Resultaten i avhandlingen stödjer SDT som användbart teoretiskt ramverk för att studera motivation i urvalsundersökningar och som en källa till idéer om hur enkätstudier kan designas med potentialen att motivera respondenterna.

Mycket att lära från kvalitativ metodik

Ett annat forskningsfält som har fungerat som inspiration för studierna i avhandlingen är kvalitativ metodik. Det är ju spännande att fundera på varför kvalitativa studier inte verkar ha lika stora problem att rekrytera informanter. När man slår upp en lärobok i kvalitativ metodik så handlar det mycket om hur man ställer frågor som inbjuder till engagemang, eftertanke och berättande. Det blir då uppenbart hur lite fokus vi som kvantitativa forskare lägger på dem som ska besvara våra

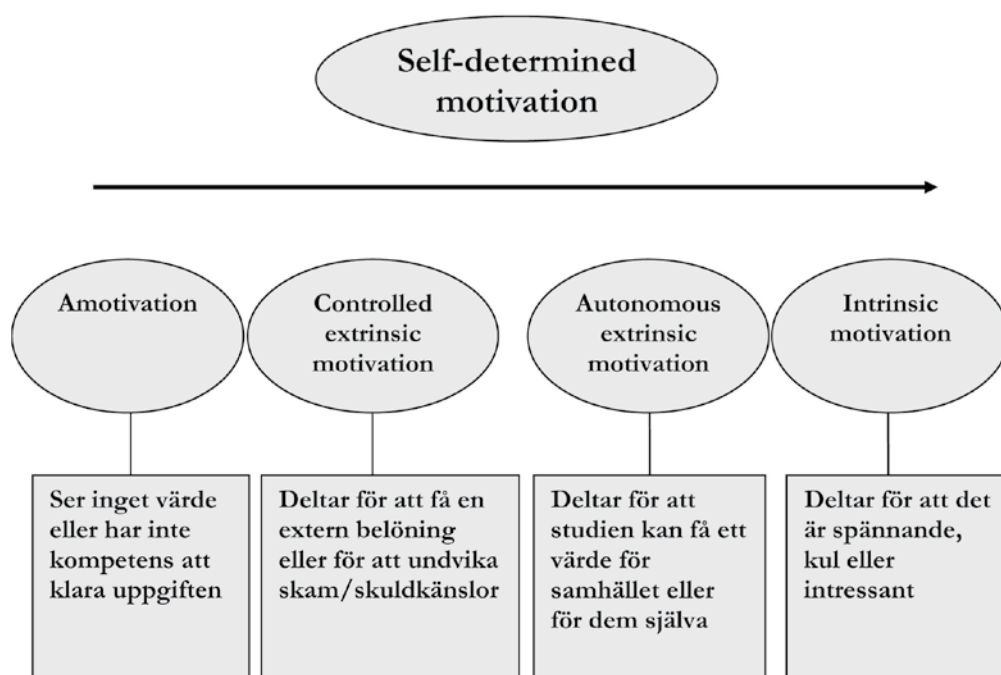


Bild 1. Förenklad bild av self-determination theory med tillämpning på deltagande i en enkätstudie

Referenser

- [1] Persson A. och Scheffer F. (2011) Att utvärdera frågor i en statistisk undersökning. *Qvintensen* nr 1/2011, sid 14-17.
- [2] Wenemark M., Persson A., Noorlind Brage H., Svensson T. and Kristenson M. (2011). Applying motivation theory to achieve increased response rate, respondent satisfaction and data quality. *Journal of Official Statistics* 27, 393-414.
- [3] Wenemark M., Hollman Frisman G., Svensson T. and Kristenson M. (2010). Respondent satisfaction and respondent burden among differently motivated participants in a health-related survey. *Field Methods* 22, 378-390.
- [4] Wenemark M. (2010). The respondent's perspective in health-related surveys. The role of motivation. Linköping University Medical Dissertations No. 1193. Avhandlingen finns att ladda ner i fulltext liu.diva-portal.org/smash/get/diva2:355603/FULLTEXT01

enkätstudier. Vi är också betydligt sämre på att följa upp hur respondenterna upplevt deltagandet. En annan intressant skillnad är hur man i kvalitativa studier ibland låter informanterna kvalitetssäkra datamaterialet.

Jag har i avhandlingen intervjuat respondenter om deras erfarenheter och upplevelser av att ha deltagit i en hälsorelaterad enkätstudie [3]. Resultaten visar ett brett spektrum av såväl positiva (respondenttillfredsställelse) som negativa (respondentbörda) aspekter av att delta. Respondenterna betonade vikten av att forskaren visar intresse att lyssna på dem och de uttryckte uppskattning av frågor som leder till reflektion och nya insikter. Respondenterna efterfrågade också ofta möjligheter att kunna ta större ansvar för att svaren ska bli rättvisande. Många av de önskemål respondenterna uttryckte tillgodoses betydligt bättre inom ramen för den kvalitativa metodiken. Jag tror att vi behöver vara mer lyhörda för dessa önskemål. Vi kan till exempel kommunicera med respondenterna genom att använda öppna frågor som komplement till flervalsfrågor. I kvalitativ metodik står informanten i centrum. I kvantitativ metodik framställs respondenterna ibland snarare som hinder när de inte vill delta, inte följer instruktioner, svarar slarvigt eller missuppfattar våra frågor.

Tradition och tid hinder för respondentfokus?

Men varför arbetar inte alla mer respondentanpassat? Jag tror att det handlar mest om två saker: vi gör som vi alltid har gjort och vi anser oss inte ha tid. Många av myndigheternas

studier är direkta regeringsuppdrag. Regeringen tillsätter en utredning för att få in faktaunderlag genom en enkätstudie, myndigheten planerar studien enligt departementets önskemål och datainsamlingsföretaget genomför datainsamlingen utifrån myndighetens uppdragsbeskrivning. Alla tar ansvar för sin del i kedjan och alla gör precis vad som förväntas. Men vem har då ansvar för att förändra? Hur ska politikerna få veta att en enkätstudie kanske inte alla gånger är det bästa alternativet att få fram ett faktaunderlag? Eller hur ska myndigheterna kunna utnyttja datainsamlingsföretagens erfarenheter i sina uppdragsspecifikationer? Vi måste alla ta ansvar för den del i kedjan vi befinner oss i genom att våga ifrågasätta och tänka nytt.

Jag får ofta frågan om det inte tar mer tid att till exempel testa frågor noga, övertvåga alternativa datainsamlingsmetoder och kommunicera med respondenterna genom öppna frågor. Jag tror att det tar mer tid, men att det är väl investerad tid om respondenterna känner att de kommer till tals och om det skapar trovärdighet för studien. När jag började arbeta med enkäter satt man fortfarande och stansade in enkäter manuellt. Vart tog den tiden vägen? Hur kommer det sig att vi har råd med urval på en halv miljon personer men inte med utvecklingsarbete?

Låt oss leka med tanken att forskarna och respondenterna skulle dela arbetsinsatsen lika i en studie. Respondenternas tid för ifyllande står för halva tiden och forskarnas tid för sammanställning, analys och implementering av resultaten står för halva tiden. Enkäten i studien Grodan (om skolbarns psykiska hälsa) fylldes i av

172 000 skolbarn under en 40 minuters lektion. Det motsvarar arbetstid för ungefär 65 heltidsanställda under ett år. Om alla patienter skulle fylla i den nationella patientenkäten (30 minuter) skulle det motsvara ungefär 144 heltidsanställda under ett år. Den tiden skulle gott och väl räcka till för att planera studien riktigt bra och till att se till att resultaten kommer till verklig nytta för barnen i skolan eller för patienterna. Vi förväntar oss att respondenterna ska bidra med väldigt många gratis arbetstimmar. Då borde de i sin tur kunna förvänta sig väldigt bra kvalitet i alla delar av studien.

Slutligen skulle jag vilja lansera en idé om att följa upp respondentens upplevelser av att delta. Tänk om alla enkäter avslutades med tre frågor: om man har kunnat ge en rättvisande bild av sin hälsa (eller det studien handlar om), om man blivit just bemyndigad och om man skulle kunna tänka sig att delta igen. Det skulle ge ett väldigt intressant komplement till svarsfrekvensen i bedömningen av studiens kvalitet. För vem skulle kunna presentera resultat från en studie där respondenterna själva inte tycker att de har kunnat lämna rättvisande data?

Detta var några personliga reflektioner efter mina år med avhandlingsarbetet. Glädjande nog har jag betydligt fler idéer om design av respondentanpassade enkätstudier idag än när jag började. Att arbeta respondentanpassat är inte lättare men mycket, mycket roligare.

MARIKA WENEMARK,
LANDSTINGET I ÖSTERGÖTLAND OCH
LINKÖPINGS UNIVERSITET

I del 1 i Qvintensen nr 3/2011 berättade Joakim Westerlund bland annat om tidskriften Journal of Parapsychology. Många skulle bli förvånade om de läste ett nummer av den: "de allra flesta artiklar presenterar signifikanta resultat med hyfsade effektstyrkor 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." Nu fortsätter berättelsen...



Joakim Westerlund.

FOTO: ORASIS FOTO/MÅ

DEL 2.

Mina äventyr i parapsykologins förtrollade värld

Men efter hand började jag upptäcka att en viss typ av händelseförlopp i den kollektiva parapsykologiska forskningsprocessen brukade återkomma och upprepa sig genom parapsykologins 80-åriga historia. Det börjar med att någon parapsykolog kommer på ett nytt sätt att studera psi-fenomenen på. Ofta brukar den nya metoden vara hämtad från det senaste och "häftigaste" inom den vanliga psykologin och de första resultaten är väldigt starka och slående. Ett antal andra parapsykologer blir väldigt entusiastiska över dessa resultat och den nya metoden och gör egna experiment men normalt inte exakta replikeringar av det ursprungliga experimentet. I stället gör de en liten förändring i upplägget, antingen för att de tror sig ha kommit på något som borde göra effekterna ännu starkare eller för att de upptäckt en potentiell felkälla i det ursprungliga experimentet. Ibland får även dessa parapsykologer positiva resultat och så kommer nya parapsykologer som upprepar processen. Men ganska snart börjar resultaten bli allt svårare att re-

producera och efter ett antal misslyckade försök glöms till slut metoden bort och ersätts av någonting annat nytt och spännande och så upprepas det hela.

Byrålädsproblemet

Jag tycker att det finns en rimlig förklaring. Även om parapsykologin försöker minska risken att endast undersökningar med positivt resultat publiceras ("byrålädsproblemet") genom att the Parapsychological Association har som policy att i sina tidskrifter även publicera icke-signifikanta resultat så är förstås enskilda forskare inte lika motiverade att skriva artiklar utifrån undersökningar som har misslyckats. Så en del av dessa nya och häftiga metoder som parapsykologer hittar på kommer helt enkelt aldrig att bli en del av forskningsfältet. En annan förklaring är förstås att när en metod är helt ny är risken större att den

innehåller brister som ännu inte upptäckts och som skulle kunna förklara de positiva resultaten.

Det bästa sättet att få en uppfattning om bevisläget inom parapsykologin är att sätta sig in i de metaanalyser som gjorts på området. När det gäller mikropsykokinesi (se Del 1 i Qvintensen nr 3/2011) gjorde den tyske parapsykologen Holger Bösch en metaanalys som publicerades i Psychological Bulletin 2006 där han enligt min uppfattning övertygande visade att de väldigt svaga men signifi-

kanta effekterna faktiskt kan förklaras av byrålädsproblemet. Psykokinesi har man undersökt ordentligt och kommit fram till att det inte finns någon som helst anledning att anta att det finns.

»Men ganska snart börjar resultaten bli allt svårare att reproducera och efter ett antal misslyckade försök glöms till slut metoden bort«

Ganzfeld

För ESP-fenomenen är bevisläget inte lika solklart. Effektstyrkorna brukar vara

större och det finns några metaanalyser med positiva resultat. Tyvärr innehåller flera av dessa för få studier för att vara särskilt övertygande. Det finns dock en alldeles särskild metod för att undersöka ESP och i synnerhet telepati, som står ut från alla andra metoder och som har använts i snart 40 år och som under denna tid utvecklats och förfinats, både när det gäller att ge eventuella telepatisignaler den allra bästa chansen att kunna visa sig (minska risken för typ II-fel) och när det gäller att eliminera alla potentiella felkällor (minska risken för typ I-fel). Denna metod kallas ganzfeld. En grov gissning är att ganzfeldmetoden använts vid ett dussintal laboratorier, av ett femtiotal försöksledare och på ca 3000-4000 försökspersoner. Så om man vill få en uppfattning om bevisläget för det minst orimliga av psi-fenomenen, telepati, bör man definitivt ta reda på vad man kommit fram till med ganzfeldmetoden.

I ett ganzfeldexperiment ligger mottagaren i en mjuk fåtölj eller säng med hörlurar på sig och lyssnar på ett homogent ljud (vitt brus, skärt brus, ljud från vågor som skvalpar eller ljudet av ett vattenfall är vanligast). Framför ögonen brukar hon ha ett par "glasögon" bestående av halva pingisbollar som belyses med rött ljus. Poängen med detta är att försätta mottagaren i ett tillstånd av mild sensorisk deprivation genom att minska på intrycken från sinnesorganen. Tillståndet har använts i den konventionella psykologiska forskningen och brukar ofta leda till att försökspersonen får livliga hallucinationsliknande fantasier. En tanke är att den förmodat svaga psi-signalen borde vara lättare att upptäcka om man minskar "bruset" från de konventionella sinnesorganen.

När väl mottagaren försatts i ganzfeldtillståndet börjar en filmsnutt visas för en mänsklig sändare som befinner sig i ett rum en bit därifrån. En slumpprocedure har från ett bibliotek om exempelvis 100 filmsnuttar, indelade i grupper om fyra, först valt ut vilken av grupperna av filmsnuttar som skall vara aktuell och därefter valt vilken av de fyra filmsnuttarna i den aktuella gruppen som skall visas för sändaren. Filmsnuttarna är ofta en eller ett par minuter långa och visas för sändaren om och om igen under ca 20 minuter. Samtidigt ligger mottagaren i ganzfeldtillståndet och pratar högt om de tankar och upplevelser hon har. Detta prat brukar spelas in på band och kallas mentation. När filmvisningen är klar inne hos sändaren "väcks" mottagaren och därefter får mottagaren titta på de

fyra filmerna i den aktuella filmgruppen och försöka bestämma sig för vilken av de fyra filmerna hon tror det var som visades för sändaren.

Studier av utförda ganzfeldexperiment

Det har gjorts ett antal metaanalyser av ganzfeldexperimenten. Flera av dessa har publicerats i den "fina" psykologiska tidskriften *Psychological Bulletin* och många parapsykologer anser att dessa metaanalyser ger starkt stöd för att psi verkligen är ett reellt fenomen. Andra, och däribland undertecknad, är mer skeptiska. Jämfört med ovan nämnda metaanalys av mikropsykokinesi kan sägas att när det gäller ganzfeld är effekterna starkare och mer homogena och någon byråladseffekt verkar det inte vara tal om. När man gjort blindade skattningar av undersökningarnas metodologiska kvalitet har man heller inte funnit några korrelationer mellan kvalitet och effektstyrka. Men visst finns det även svagheter som man kan läsa sig till. För det första hade de tidigaste ganzfeldexperimenten klara metodologiska brister och dessa studier kan man nog med gott samvete bortse från. För det andra kunde en av metaanalyserna, av Milton och Wiseman, som omfattade perioden 1987-1997, inte visa på någon signifikant genomsnittlig effektstyrka. För det tredje verkar det som att vissa försöksledare oftast får positiva resultat medan andra oftast får negativa resultat. Min skepsis beror dock på mer personliga erfarenheter av parapsykologisk forskning i största allmänhet och av ganzfeldexperiment i synnerhet.

Mitt eget ganzfeldexperiment

När jag började med parapsykologi stod ganzfeldmetoden på sin topp. En metaanalys med positiva resultat hade just publicerats i *Psychological Bulletin* och många parapsykologer trodde nog att det bara var en tidsfråga innan parapsykologin skulle få sitt stora genombrott. Jag var 1997 på min första parapsykologiska konferens i Brighton och det var där jag första gången fick höra dessa anekdoter som så ofta kan berättas av personer som arbetat som försöksledare i ganzfeldexperiment. Det var en av pionjärerna inom ganzfeldmetodiken som över en lunch berättade att de en gång, på den tiden de använde sig av bilder i stället för videofilmer, hade visat en bild på en spindel för sändaren. Mottagaren hade då sagt ordet "spider" 46 gånger och resten av tiden bara pratat om spindelnät och andra relaterade saker. Kort efter



att jag hade kommit hem till Stockholm fick jag ett videoband av Adrian Parker som höll på med ganzfeldexperiment på Psykologiska institutionen i Göteborg. Videobandet var en sammanställning av påfallande likheter mellan filmen som visats för sändaren och mentationen (pratet) från mottagaren som hade uppstått i ganzfeldexperiment i Edinburgh. Sammanställningen hade gjorts av en av försöksledarna vid dessa experiment, Kathy Dalton. Detta band gjorde djupt intryck på mig och min kollega Jan Dalkvist. Likheterna mellan film och prat var vid flera tillfällen helt förbluffande. Det var särskilt en scen, med tillhörande mottagarprat, som vi blev imponerade av. På filmen ser man hur en man springer genom en skog, samtidigt säger mottagaren "trees, people running". Plötsligt faller mannen ner i en djup gytjepöl och mottagaren säger "falling..., muddy". Kameran zoomar in på mannens ansikte och mottagaren säger "blond hair, 70s' hairstyle, curly-ish" – vilket är precis med innehållet överensstämmande. Sedan, i exakt samma ögonblick som Blixt Gordons näsa hamnar under ytan säger mottagaren "Dead man in the water"! Hur stor är egentligen sannolikheten för att sådana här frapperande överensstämmelser mellan filmen som visas för sändaren och pratet från mottagaren skall kunna inträffa av en slump?

Om dessa överensstämmelser berodde på telepati borde man kunna utnyttja dem till att göra bättre ganzfeldexperiment. På den tiden spelades mottagarens prat in på band och om mottagaren exempelvis vid ett tillfälle under ganzfeldtillståndet sa att hon såg ett ansikte så skulle hon vara föga hjälpt av detta när hon sedan tittade igenom de fyra filmerna på jakt

Ganzfeldexperimentet försätter mottagaren i ett tillstånd av mild sensorisk deprivation genom att minska på intrycken från sinnesorganen.

»»»»

»»» MINA ÄVENTYR I PARAPSYKOLOGINS FÖRTROLLADE VÄRLD

efter likheter mellan dessa och hennes upplevelser eftersom det antagligen skulle förekomma ansikten i alla fyra filmerna. Men, om hon samtidigt som hon tittade på filmerna kunde höra sitt eget prat synkroniserat med den faktiska filmen som visats för sändaren och om hon då kunde se och höra att precis samtidigt som hon sa att hon såg ett ansikte så dyker det upp ett ansikte i närbild i film nr 3, då skulle detta kunna hjälpa henne att välja rätt film.

Jag gjorde därför ett dataprogram som utnyttjade detta och som användes för att köra ganzfeldexperiment i Göteborg 128 gånger (vilket är en ovanligt stor datainsamling i sådana här sammanhang; ofta brukar en experimentserie omfatta ca 30 enskilda försök). Till dataprogrammet hade jag digitaliserat 100 filmer som alla var exakt 2 minuter långa (detta var på den tiden då det fortfarande var ganska ovanligt att visa filmer på datorn). När ett ganzfeldförsök startades valdes en av 25 grupper om fyra filmer ut slumpmässigt av datorn och av dessa valdes slumpmässigt en av filmerna ut som "target", det vill säga som den film som visades för sändaren. Filmen visades för sändaren 7 gånger och samtidigt spelades mottagarens prat in och lagrades i en ljudfil vars namn avslöjade vilken av de 25 filmgrupperna som hade använts men förstås ingenting om vilken av de fyra filmerna som visats. När försöket var klart skickades ljudfilen upp på en server och sedan

kunde jag, hemifrån Stockholm, ladda ner ljudfilen till min egen dator och lyssna på denna synkroniserad med de fyra aktuella filmerna. Utifrån de likheter jag hittade försökte jag gissa vilken av de fyra filmerna som hade visats för sändaren.

En stor del av nöjet med att utforma detta och andra parapsykologiska experiment bestod i att i förväg tänka igenom vilka möjliga kryphål som kan finnas för information att på icke-paranormal väg överförs och påverka resultaten.

Eftersom jag själv hade fått vara med

att utforma experimentet kände jag att jag hade koll på alla detaljer och att om experimentet skulle ge positiva resultat så skulle jag på allvar börja tro på möjligheten att det verkligen finns något paranormalt fenomen. Nu blev huvudresultatet av detta experiment att jag gissade rätt film i 30 av 128 fall vilket är aningen sämre än vad som kan förväntas av en gissare som bara går på chans. Men, ännu viktigare för mig, var resultatet av ett annat test vi gjorde.

Jag hade nämligen bestämt mig för att jag medan jag jämförde ljudfilerna med de olika filmerna, och då ju inte visste vilken som var den film som faktiskt hade använts i experimentet, skulle notera om det uppstod någon riktigt anmärkningsvärd överensstämmelse mellan mottagarens prat och det som visades på någon av filmerna. Jag bestämde mig för att välja ut högst en sådan överensstämmelse per ljudfil.

Jag minns fortfarande den allra första gången jag lyssnade igenom pratet från den första mottagaren och jämförde detta med innehållet i de fyra filmerna. Jag hittade massor med överensstämmelser mellan pratet och alla fyra filmerna!

När hela datainsamlingen var klar, vilket tog ungefär ett år, satt jag där till

slut med 20 frapperande sekvenser med överensstämmelser framför mig. Exempelvis kunde man för en sådan sekvens höra mottagaren säga: "Ett gammalt ånglok kommer åkandes".

Samtidigt ser man på en av filmerna hur ett gammalt ånglok kommer åkandes. Sedan hör man mottagaren säga "tjuter" och precis samtidigt kan man på filmen höra hur ångloket tjuter!

Om dessa fantastiska överensstämmelser berodde på telepati så borde man förvänta sig att betydligt fler än 25% (dvs 5 av 20) av dem var överensstämmelser mellan mottagarens prat och den film som faktiskt hade visats för sändaren. Nu visade det sig att detta bara var fallet för sex av överensstämmelserna, alltså återigen ett resultat väldigt nära vad man

kan förvänta sig att få av en slump. Exempelvis var ovan nämnda film med ångloket som kom åkande och tjöt en av de tre kontrollfilmerna och hade alltså inte visats för sändaren i det aktuella experimentet!

Slutsatser

Av det här experimentet drog jag två slutsatser. För det första att hjärnan är en maskin som inte kan låta bli att hela tiden väva ihop de intryck den får till en sammanhängande och meningsfull "berättelse". Prova själv att sätta på en tv men stäng av ljudet. Låt sedan en annan tv stå på men inställd på en annan kanal så att du kan höra ljudet från den andra kanalen. Titta på den första tv:n och lyssna på den andra. Snart kommer du att uppleva att ljudet verkligen hänger ihop med bilden på ett sätt som inte känns helt slumpartat.

Den andra slutsatsen jag drog är att de mest otroliga sammanträffanden faktiskt inträffar rätt ofta av en slump.

En annan sak jag har förstått efter min tid inom parapsykologin och efter mina erfarenheter av bland annat ovan nämnda experiment, är att det finns oerhört många detaljer som kan vara avgörande för vilka resultat man får och vilka slutsatser man kan dra och jag tänker då på detaljer som aldrig någonsin hamnar i de publicerade artiklarnas metodavsnitt. Väldigt mycket av forskningsarbetet är ett hantverk som sitter i ryggmärgen. Exakt hur man gör med dessa detaljer beror förstås på utbildning, intelligens, erfarenhet och sådana saker. Men mycket påverkas nog av ens personlighet, om man är pedantisk och noggrann eller om man är mer impulsiv till exempel. Min poäng är att jag tror att det finns massor med undersökningar (och däribland parapsykologiska sådana) som verkar alldeles utmärkta när man läser om dem, men som skulle visa sig innehålla allvarliga brister om man kunde följa dem som en fluga på väggen och ha möjlighet att få information om minsta lilla detalj.

JOAKIM WESTERLUND,
STOCKHOLMS UNIVERSITET

»Min skepsis beror dock på mer personliga erfarenheter av parapsykologisk forskning i största allmänhet.»

Joakim Westerlunds fascinerande berättelse om sin parapsykologiska forskning väcker många frågor. En fråga som jag undrat över är varför statistiska resonemang blir så väsentliga i just pseudo-vetenskap.

Eftersom vi lever i en debattkultur där man förväntas ha invändningar mot allt möjligt kanske jag måste påpeka att Joakim Westerlund och jag tycker i stort sett likadant om kommentarerna nedan.

Kommentarer till forskningsområdet parapsykologi

Vad PSI-fenomen är antyds endast med ett slags restdefinition: när man uteslutit allt annat som är "naturligt" då återstår något, och där någonstans äger PSI-fenomenet rum. Det är inte en signal som kan mätas eller upptäckas med dagens instrument. Det är inte något som kan uppfattas med våra fem sinnen. Det är inte något annat som skulle utgöra en "naturlig" förklaring. Man har vad jag vet ingen arbetshypotes om vad för signal man letar efter när man undersöker PSI-fenomen. Det som då återstår är att göra experiment med sändare och mottagare och leta efter någon avvikelse från förväntat värde i en statistisk fördelning som vore konsistent med att mottagaren gissar. Varje statistiskt signifikant avvikelse tolkas som att den kan ha orsakats av PSI. Kan man säkert dra slutsatsen att den måste ha orsakats av PSI? Då måste man först kunna utesluta varje annan förklaring. Det finns enormt många andra förklaringar, kända och okända, och samtliga dessa måste på säker grund kunna uteslutas för att man ska kunna dra slutsatsen att det som observeras beror på PSI. Det är knappast möjligt.

Finns det en tydlig arbetshypotes om vad som ger förmåga till PSI, till exempel ett elsinne vars existens Joakim Westerlund helt korrekt påpekar inte kan uteslutas, då kan hypoteser om elsinnets natur formuleras och studeras. I vilket organ sitter detta elsinne? Hur fungerar detta organ? Man kan tycka att framsteg borde kunna göras – om det finns något att se. Jämför med ett starkt forskningsområde, där det finns en stabil och fruktbar teori som kan generera hypoteser som faktiskt borde vara sanna om teorin stämmer. Då kan dessa hypoteser undersökas. Ta evolutionsteorin som exempel. Om den teorin stämmer, då borde det finnas arter med tillbakabildade organ (därför att deras ursprungsart

hade ett organ som inte längre behövs i den miljön den nya arten lever i) och det borde gå att hitta en utvecklingskedja av arter i jordlager av olika ålder. Passa på att se Richard Dawkins om dessa två fall (se adresser i referenslistan [1], [2]) och njut av hans pedagogiska förmåga.

PSI är en onödigt förklaring

Man kanske kan tycka att om mottagaren gissar rätt på 53 % av sändningarna i ett ganzfeldexperiment och om det är statistiskt signifikant relativt vald nivå på testet, då ska det tolkas som evidens för PSI. Ett sådant resultat är emellertid av begränsat värde. Det kan lika gärna bero på en felkälla som man missat.

I vetenskapliga sammanhang föredrar man enkla förklaringar framför komplexa, krångliga förklaringar eller förklaringar som innehåller okända saker, om de i övrigt är likvärdiga. Den forskningsstrategin kallas Occams rakkniv. Occams rakkniv har att göra med målet att nå intersubjektiv kunskap (se [3] och [4]), alltså gemensam kunskap som de flesta av oss kan vara överens om trots att den innehåller subjektiva satser. Det är bara om man gör den minimala mängden antaganden, som tvingas fram av fakta, som vi har en chans att vara överens. Så fort man gör ytterligare antaganden, som inte är nödvändiga, då uppstår diskussioner om de är riktiga och om man inte borde göra andra antaganden istället. Alltså blir kunskapsmassan mer kraftfull med Occams rakkniv.

Med Occams rakkniv förklarar man hellre 53 % rätt med att det finns någon felkälla man missat eller att en händelse med låg sannolikhet inträffade den här gången än att vi har sett ett fenomen som vi inte kan förstå.

Lägg märke till att detta sätt att resonera står i samklang med vardagslivets sunda resonemang.

Om jag saknar en strumpa efter att ha tömt tvättmaskinen, vilken av följande tänkbara orsaker börjar jag med att undersöka:

1. Strumpan sitter fastklistrad i trumman i tvättmaskinen.

2. Den har hamnat någon annanstans; hemma hos mig kan den t.ex. ligga kvar otvättad i dammet under sängen.

3. Ett spöke har tagit den.

Vem börjar med att undersöka alternativ 3? Även om jag tror att jag har uteslutit alla "naturliga" förklaringar, vem tror att det är mest sannolikt att ett spöke har tagit strumpan?

Nonsens-signifikans

Letandet efter någon signifikant avvikelse utan någon tydlig hypotes om vad det är man letar efter leder lätt till endera av tre ting: man tröttnar, man hittar falska signifikanser eller så blir det ett ändlöst insamlande av data med magra slutsatser. I sin fascinerande artikel om skallmätningar beskriver Olof Ljungström hur mätningarna föreslogs utökas med ännu flera mått när de mätningar som redan gjorts inte visade på det man ville att de skulle visa: att formen på människors kranium är associerad med ras [5]. Hade de mätningarna gjorts idag, nu när teorin för statistiska test finns utvecklad, och hade man tillåtit sig att formulera hypoteser efter man sett data, då hade någon statistiskt signifikant association lätt kunnat hittas någonstans [6].

Skevheter i redovisning

Termen byrålädsfenomenet (publication bias) avser att det publiceras fler artiklar som redovisar evidens på något samband eller någon effekt än artiklar som undersökt något och inte funnit något samband eller effekt. Man brukar

»»»»

»»» KOMMENTARER TILL FORSKNINGSOMRÅDET PARAPSYKOLOGI

säga att fler positiva än negativa artiklar kommer ut. Som Joakim Westerlund påpekar är ett banalt skäl att det helt enkelt är inte så kul att fortsätta sin mödosamma forskning när resultaten blir negativa. Med positiva resultat sporras däremot forskaren att fortsätta arbetet. Även om många tidskrifter, däribland *Journal of Parapsychology*, uppmuntrar artiklar med negativa resultat kan det finnas anledning att tro att negativa artiklar ibland refuseras eftersom de inte visar något nytt. Det har gjorts försök att studera skevheten som orsakas av icke-publicerade artiklar [7] och det finns faktiskt tidskrifter som koncentrerar sig på negativa resultat, t.ex. [8].

Byrålädsfenomenet är besvärande i den typ av tillämpad forskning där statistiska test är centrala. Inom vissa forskningsområden har man därför byggt upp en infrastruktur där pågående experiment dokumenteras och redovisas så att man kan se vilka experiment som gjorts även om inte alla av dem publicerats i forskningstidskrifter [9]. Inom många områden saknas en sådan infrastruktur. Det gäller t.ex. psykologi. Inom de områdena kan man klara sig ganska bra ändå eftersom experimenten kan göras billigare och därför upprepas oftare. Styrkan i forskningsområdet ges av den samlade bilden av många experiment och annan empiri som stöds av teoretiska resonemang.

Byrålädsfenomenet har använts för att försöka avfärda studier som visat på samband. Till exempel har Singer [10] fört fram att bara för att det finns en mängd studier som visar på hälsorisker med passiv rökning kan vi inte dra slutsatsen att passiv rökning är negativt för hälsan; det kan finnas icke-publicerade studier som inte påvisat samband och sammantaget skulle då bilden om riskerna vara osäker! Att det alls går att lägga fram en så konstig argumentation visar på ett grundläggande problem som byrålädsfenomenet för med sig.

Vår dåliga förmåga att bedöma sannolikheter

Om ett experiment visar på en effekt fast det inte finns någon brukar man säga att man fått ett fel av typ 1. Testet

har nivån 5% är ett sådant fel faktiskt inte osannolikt men omtalas ofta som om det vore nästintill omöjligt. Massmedia brukar inte rapportera testets nivå när det meddelas att "ny forskning visar det-eller-det". Om experimentet dessutom upprepas och byrålädsfenomenet verkar, då blir fel av typ 1 mer sannolikt än vad många intuitivt skulle gissa. Särskilt stor är sannolikheten att något osannolikt inträffar om utfallsrummet är större än vad man tänker sig. Hur stor är sannolikheten att jag helt oväntat träffar på en arbetskamrat i en by i Toscana? "Tänk att jag skulle dyka på Anders just här, i den här trattorian!" Det känns helt otroligt. Men den händelsen ska ställas i relation till alla de relevanta händelser som hade kunnat inträffa, dvs det vi kallar utfallsrummet. Vilket är utfallsrummet i det här fallet? Räknas alla platser utanför Sverige och alla som jobbar på samma jobb? Ingår i utfallsrummet oväntade sammanträffande var som helst på "oväntad" plats med vem som helst som jag är bekant med? I så fall blir det många möjligheter att något "osannolikt" ska inträffa. De överensstämmelser mellan filmer och "prat" som Joakim Westerlund beskriver är ett ovanligt tydligt exempel på odefinierat utfallsrum, därför inträffar som han påpekar till synes otroliga händelser ofta av en slump.

Vi är väldigt bra på att se mönster. Vi har också en intuitiv ovilja mot att tolka mönster och sammanträffanden som slumpmässiga. Det rapporteras att slumpmässiga spellistor i iPod-apparater inte är slumpmässiga: Med en enkel slumpmekanism skulle nämligen samma låt ofta komma två gånger i rad. Urvalsmekanismen måste göras mindre slumpmässig för att framstå som mer slumpmässig, som Steve Jobs uttryckte det [11].

Sammantaget har parapsykologin som forskningsområde problem, inte minst beroende på avsaknaden av bärkraftig teori (som jag gissar i sin tur beror på att man forskar på något som inte finns). I ett starkt forskningsområde växelverkar teori med empiri och ger fördjupade och utökade kunskaper.

DAN HEDLIN

Referenser

[1] Dawkins, R. (odaterad). *Vestigial Organs: The Wings of the Flightless Cormorant*. Film sedd den 13 nov 2011 på <http://richarddawkins.net/rdf Productions/vignettes>. The Richard Dawkins Foundation for Reason and Science.

[2] Dawkins, R. (odaterad). *Diatoms: The Evolution of a New Species - Nebraska Vignettes #5*. Film sedd den 13 nov 2011 på <http://richarddawkins.net/rdf Productions/vignettes>. The Richard Dawkins Foundation for Reason and Science.

[3] Hansson, S.-O. (1987). *Vetenskapens kunskapssyn*. Folkvett nr 1/1987, sid 11-14.

[4] Hansson, S.-O. (1989). Om nyttan av det enkla. *Vetenskapens kunskapssyn*. Folkvett nr 1/1989, sid 11-12.

[5] Ljungström, O. (2010). Skallmätningarnas tidiga historia. *Qvintensen* nr 3/2010, sid 12-14.

[6] Hansson, S.-O. (2010). Signifikansmissbruket i pseudovetenskap. *Qvintensen* nr 3/2010, sid 9-10.

[7] Copas, J.B. and Shi, J.Q. (2001). A sensitivity analysis for publication bias in systematic reviews. *Statistical Methods in Medical Research*, 10, 251-265.

[8] *Journal of Articles in Support of the Null Hypothesis*. www.jasnh.com

[9] Huitfeldt, B. (2011). Oberoende eller kompetens. *Qvintensen* nr 1/2011, sid 7-11.

[10] Singer, F.S. (2010). Secondhand Smoke, Lung Cancer, and the Global Warming Debate. *American thinker*. Artikel hämtad den 13 nov 2011 från http://www.americanthinker.com/2010/12/second_hand_smoke_lung_cancer.html

[11] Mlodinow, L. (2008) *The drunkard's walk. How randomness rules our lives*. Penguin Books.



ORDFÖRANDEN HAR ORDET

Fel, fel, fel!

Fel, fel, fel, sa Brasse till Magnus, när han triumferande viftade med hatt och armar. Vi barn förstod med ens vad ett fel var: Det var något som inte var rätt, eller om man så vill, något som inte var sant. Man kan undra hur Brasse skulle ha burit sig åt för att förklara för Magnus att ett statistiskt fel i egentlig mening inte är ett fel.

Att termen "fel" i statistiska sammanhang är ett olyckligt ordval är det många som anser. Vad icke-statistikern då hör är att om det finns ett "fel" så måste det finnas ett "rätt", snarare än en skattning av ett "rätt". Än knepigare blir situationen om vi försöker förklara att det oftast bara är urvalsfelet vi kan, eller väljer att, uttala oss om, inte allt det där andra, det som tillsammans med urvalsfelet kallas det totala felet. Det är något vi har svårt för att prata om. Varför är det så?

Förmodligen för att vi har en kommunikationsklyfta att ta oss över. Vi har svårt för att kommunicera sådant som inte är helt och hållet mätbart eller innehåller ett mått av sannolikhet. Med en viss erfarenhet av statistiska undersökningar får vi en uppfattning om, en känsla för, hur viktigt det är att frågorna

är korrekt formulerade, att intervjun gjorts på rätt sätt och att insamlade data hanterats så att ingen information gått förlorad eller förvanskats.

Det knepiga är att det inte är helt orimligt att tro att den vardagliga statistik användaren tar för givet att hela denna kedja fungerat klanderfritt om enbart urvalsfelet diskuteras i undersökningsresultatet. Och varför skulle hon inte det?

Min erfarenhet är att det är nog svårt att kommunicera vad ett urvalsfel är. Det är även svårt att förklara vad ett sannolikhetsurval är och varför vi inte kan jämföra två olika undersökningsresultat om de använda metoderna inte tillåter det. Men ska vi ge tappt fördenskull?

Vad görs inom forskningen? Den internationella sammanslutningen International Total Survey Error Workshops (ITSW) träffas årligen och i år hölls träffen i Kanada. För ett par år sedan, 2009, hölls träffen i Sverige under ledning av Lars Lyberg. Han presen-

terade också ett papper, Total survey error: Past, present, and future, tillsammans med Robert M. Groves som året därpå publicerades i Public Opinion Quarterly (nummer 5, sidorna 849 - 879) och som kan tjäna som

en bra ingång till problematiken ur ett totalfelperspektiv. Läsaren tas med på en odysse genom totalfelets historia med början år 1944 (In Errors on Surveys av Edwards Deming). Den stickprovsteoretiska grunden som lagts årtiondet dessförinnan innehöll starka antaganden om att de icke-urvalsbaserade felen var små.

Med Deming börjar utvecklingen av konceptet total survey error och vi får lära hur inte bara språkbruket och totalfelets beståndsdelar förändrats genom åren, utan också hur andelen sidor i teoriböckerna som ägnats totalfelet varierat.

En sista fråga: Är fem urvalsfel fler än fyra totalfel?

»Vi har svårt för att kommunicera sådant som inte är helt och hållet mätbart eller innehåller ett mått av sannolikhet.«

PS I förra numret av Qvintensen skrev jag om öppna data. Någon vecka efter att det förra numret kommit ut får jag under rubriken "Stockholm goes open data" reda på att Stockholms stad har lanserat webbsidan <http://open.stockholm.se/>. DS

JOAKIM MALMDIN
ordfsurvey at gmail.com



FOTO: MONICA HOLMBERG

»Poängen var att visa uppdrags-
givaren hur fel det kan slå om
man använder en billig metod.»

BORTFALL

– SEMINARIUM MED SURVEYFÖRENINGEN

Det är nästan alltid svårt att få in svar från alla som dragits i ett slumpmässigt urval. Andelen som utgör bortfall har stadigt ökat i undersökningar av Sveriges befolkning.

Surveyföreningen anordnade tidigare i år ett seminarium om bortfall, ett ständigt och alltid lika aktuellt problem i surveyundersökningar. Seminariet belyste bortfallsproblematiken ur ett brett perspektiv – vilka konsekvenser det kan få i en undersökning samt hur man kan arbeta för att reducera och justera för bortfall. Bortfallet är besvärande därför att det stör slutledningarna som kan dras från det slumpmässiga urvalet. Åtgärderna är i princip att göra undersökningen trevligare och mer intressant att delta i eller att få in fler eller bättre svar genom förbättringar av bland annat kontaktstrategier. När alla svar samlats in finns möjligheter att justera skattningarna så att bortfallens störande effekt blir mindre.

Jan Hörngren inledde med att presentera SCB:s paraplyprojekt för åtgärder att minska bortfallet i individ- och hushållsundersökningar. Han redovisade vad som kan ligga bakom bortfall och delade upp det i externa faktorer (t.ex. surveyklimatet i samhället och konkurrens om uppgiftslämnarnas tid) och interna faktorer (t.ex.

vilka strategier SCB använder för att kontakta uppgiftslämnare eller beläggningen på SCB:s intervjuarenhet). Han visade sedan några bilder på bortfallsutvecklingen som tydligt åskådliggjorde att bortfallet stigit kraftigt de senaste åren. Denna utveckling gör bortfallsfrågan extra angelägen just nu.

Jan Hörngren gick igenom olika åtgärder. En av dem är ett nytt introduktionsbrev till utlandsfödda i intervjuundersökningar där alla får ett brev på svenska och åtta andra språk som är vanliga i Sverige, samtliga skrivna på klarspråk, alltså enkelt, korrekt och tydligt språk. Andra åtgärder är att utveckla bättre kontaktstrategier, förbättra spåringsrutiner av telefonnummer och arbete med SCB:s varumärke. Jan Hörngren redovisade ett experiment med nya kontaktstrategier som gav samma andel genomförda intervjuer fast med färre kontaktförsök, vilket innebär att undersökningen genomförs till lägre kostnad. En metod som studeras på SCB kallas responsive design. Den innebär att resurser styrs om under genomförandet av undersökningen så att mängden av svar bättre representerar populationen, enligt det mått på representativitet som man har valt. Det kan till exempel leda till att man försöker samla in svar i kategorier med särskilt högt bortfall.

Mer kortsiktiga åtgärder inkluderade bland



Daniel Thorburn.

FOTO: MARIA SANDQVIST



Åke Wissing.



Jan Hörngren.

FOTO: MONICA HOLMBERG

annat olika inbäddade experiment med förebrev innan enkätutskick, olika utskicksdagar, olika påminnelsestrategier och belöningar för deltagande. Ett experiment visade att ett förebrev ökade svarsfrekvensen med tre procent. En fråga som kvarstod var dock om det var förebrevet i sig eller endast att förebrevet innebär en extra kontakt som orsakade den högre svarsfrekvensen. För att undersöka den saken genomfördes ett experiment till med tre grupper: 1) standardupplägg, 2) standardupplägg + förebrev 3) standardupplägg + extra påminnelse i slutet. Med denna design kan förebrevet och den extra påminnelsen jämföras med standarddesignen men också med varandra. Preliminära resultat visade att förebrevsgruppen hade 3-4% högre svarsfrekvens än standardupplägget. Det troliga är att förebrev kommer att bli standard i SCB:s enkätundersökningar.

SCB publicerade länge en årlig rapport om bortfallet i flera viktiga undersökningar, den så kallade Bortfallsbarometern. Ytterligare åtgär-

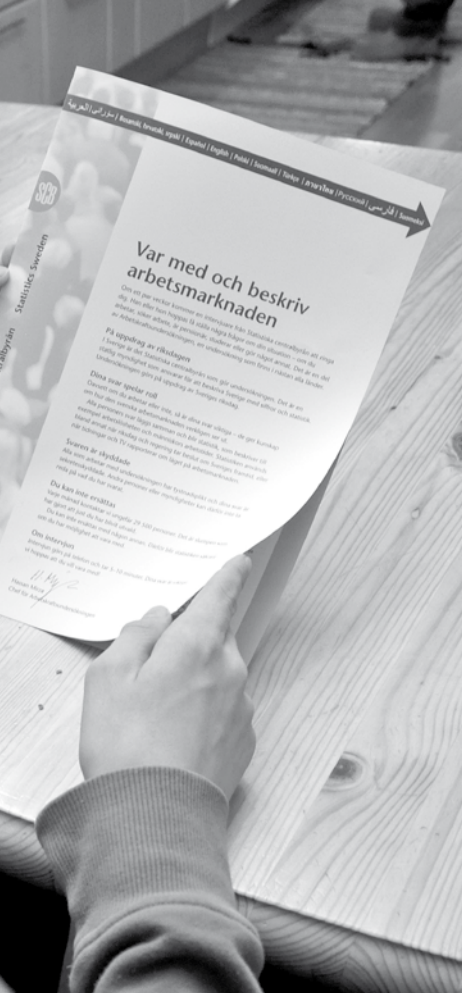


FOTO: LINNEA HEDLIN

Ett nytt introduktionsbrev för intervjuer i Arbetskraftsundersökningen.

Tävling om bästa uppsats

Kom ihåg att anmäla goda uppsatser inom surveyområdet till tävlingen om bästa uppsats!

Uppsatsen ska ha lagts fram mellan den 25 januari 2011 och den 24 januari 2012. Det senare datumet är sista dag för nominering. Nominerade uppsatser bedöms utifrån graden av originalitet och nyskapande, relevans i metodansats med avseende på vald problemställning, genomförande, och klarhet i framställningen.

Vinnande uppsats utses av en jury bestående av styrelsen för Surveyföreningen. Prissumman är på 7500 kronor och överlämnas i samband med Surveyföreningens årskonferens den 22 mars 2012. Syftet med tävlingen är att stimulera intresset för samhällsstatistisk undersökningsmetodik bland studenter och handledare i grundutbildningen.

Skicka din nominering till Joakim Malmdin på ordfsurvey@gmail.com

der som Jan Hörngren nämnde är att återinföra Bortfallsbarometern, att arbeta mer teoretiskt med orsaker till bortfall samt det internationella samarbetet. Bortfallsproblemet finns i hela världsliden. Jan Hörngren berättade också att den bortfallsbrevlåda som projektet inrättat blivit en succé. Framförallt intervjuarkåren har lämnat många bra förslag som även påverkat inriktningen i några utvecklingsprojekt.

Lätt- och svårnådda

Åke Wissing från företaget ISI Wissing fortsatte därefter med att berätta om praktiska konsekvenser av bortfall. En undersökning av marknadsandelar för hemelektronik som genomfördes vid årsskiftet 2010/2011 fick stå som exempel. Urvalet bestod av 2 300 vuxna personer som slumpvis dragits från befolkningsregistret. Nära 1 400 personer intervjuades och Åke visade med hjälp av bortfallssnurren (se nedan) hur man räknar ut andelen bortfall. Undersökningens frågor redovisades och respondenternas svar på dem fördelades i tre grupper: lättådda (telefonnummer hittades genom datakörning/maskinell sökning), svårådda (komplettering genom att manuellt söka telefonnummer och även kontakta personerna med brev) och återkontaktade vägrare. Vid den maskinella sökningen hittade man 70 procent av personerna vilket ökade till cirka 90 procent genom den manuella sökningen. I den bortfallsanalys som Åke sedan presenterade gick det, i flera av frågorna, att se ett gemensamt

mönster för hur de tre grupperna svarade. Sammanfattningsvis gick det att säga att det krävdes mycket stora resurser för att nå de svårådda, och att resultatet för den gruppen skiljde sig på flera punkter från gruppen lättådda. Även svaren som gavs av de återkontaktade vägrarna skiljde sig åt från den lättådda gruppens svar. Åkes slutsats var att den högre totala svarsfrekvensen som uppnåddes efter återkontakten påverkade resultatet.

Som avslutning visades en metodstudie över hur svaren fördelade sig vid en undersökning där formulären skickades med en kundtidning kontra en läsvärde mätning med slumpmässigt urval. Poängen var här att visa uppdragsgivaren hur fel det kan slå om man använder sig av en billig metod (medskick i tidning). Enbart genom att titta på respondenternas profiler kunde vi se att de skiljde sig åt mellan de olika undersökningarna. Skillnaderna gick också att se i hur respondenterna besvarat de olika frågorna.

Bortfallsproblematiken ur ett metodperspektiv

Daniel Thorburn, Stockholms universitet, inledde med att kort diskutera många olika aspekter av problematiken, allt från psykologiska faktorer hos respondenten till bortfallsuppföljning. Därefter koncentrerade sig Daniel Thorburn på två av aspekterna: "bortfall versus andra fel" och "Vad göra med ett stickprov med bortfall?". I det första fallet fick vi exempel på ramfel, mätfel,

aktualitet och ett resonemang om hur måttet på bortfallsandelen inte alltid säger allt. Ska vi hellre försöka minimera totalfelet eftersom en högre svarsfrekvens nödvändigtvis inte innebär att totalfelet blir mindre?

Hur ska vi då förhålla oss till bortfallet när det är en realitet? Daniel punktade upp fyra åtgärder; gör bortfallsredovisning, redovisa tänkbara effekter (även på undergrupper), gör bortfallsstudier och utför en kvalitetskontroll eftersom det alltid är bra att ha en viss uppfattning om totalfelet.

Hur gör vi då när vi har genomfört en undersökning och vi har bortfall? Utan hjälpinformation finns det inte mycket vi kan göra, men med hjälpinformation finns det många olika metoder att använda sig av. Två grupper av metoder under rubrikerna omviktning och imputering visades.

Ett femtiotal deltagare hade sökt sig till Garnisonen i Stockholm och alla som ville fick med sig ett exemplar av Bortfallssnurren, en skrift om bortfallsberäkning som togs fram av Surveyföreningens bortfallskommitté under ledning av Gösta Forsman och där Jan Hörngren och Åke Wissing ingick. Ett nytt inlägg var avbrottet för lunch efter halva seminarietiden, något som tycktes uppskattas av alla dem som tog tillfället i akt att språka med kollegor.

JOAKIM MALMDIN,
ANDREAS PERSSON,
STATISTISKA CENTRALBYRÅN

Europeisk forskning om urvalsundersökningar

Blir den som svarar på en enkät påverkad av själva utseendet på enkäten? Om man gör en statistisk analys, ska designvikterna tas med i analysen? Spelar det någon roll vilket medium man använder för att kommunicera med respondenten? Dessa och många andra frågor diskuterades i Lausanne.

Den fjärde ESRA-konferensen (European Survey Research Association) lockade runt 700 deltagare från hela världen till Lausanne i Schweiz. Föreningen ESRA syftar till att vara sammanlänkande inom surveyforskningen i Europa. Huvudföreläsare var Roger Tourangeau från University of Maryland, USA, och Chris Skinner från University of Southampton, England. Roger Tourangeau berättade om forskningen han bedrivit inom området webbundersökningar, ofta tillsammans med Mick Couper, University of Michigan. Tourangeaus huvudtema var betydelsen av den visuella presentationen vid enkätundersökningar på webben. I sin forskning har han visat att man kan få olika resultat beroende på vilka bilder som visas i anslutning till frågorna. Det kan till exempel handla om mat- och motionsvanor. I Qvintensen nr 4/2009, sidan 20-21, finns exempel på hur enkätens utseende styr uppgiftslämnare trots identiska frågor.

Chris Skinners föredrag handlade om användandet av stickprovsvikter (designvikter) vid analys av surveydata. Han visade till exempel hur en regressionsanalys påverkas med respektive utan hänsyn till urvalsdessignen. Chris Skinner menade att generellt har argumenten för att inte ta hänsyn till designvikter i analyser minskat i takt med att det finns mer och mer inbyggt i standardprogramvaror.

Konferensen hade tolv parallella sessioner i sammanlagt tretton pass. Det blir i runda tal 150 olika föredrag vi hade att välja emellan, vilket inte är alldeles lätt alla gånger. Ämnena för de

olika sessionerna varierade från metodstinna till mer praktiskt inriktade. Exempel på det förra var "Analyzing social change with repeated cross-sections" och exempel på det senare en session under namnet "Human values". En session med rubriken "Methodological and technical challenges in surveys of immigrant-background and ethnic minority groups" visade sig snarare handla om svårigheten att förhålla sig till den vetenskapliga teorin bakom metodansatsen, vilket också den efterföljande diskussionen bekräftade. Detta var för övrigt något som noterades på flera sessioner – att vetenskapens beståndsdelar hänger ihop och påverkar varandra.

Precis när vi lärt oss (?) skilja på metadata (information som beskriver data) och paradata (information som beskriver datainsamlingsprocessen; se Qvintensen nr 3/2009, sidan 26) införs ett nytt begrepp – pandata. Asaph Young Chun från U.S. Census Bureau menade att ett pandatasystem består av administrativa data, metadata, paradata och surveydata. En annan session som tog upp metadata fokuserade på DDI (Data Documentation Initiative) och de hjälpmedel som utvecklats för att underlätta maskinell datainsamling och användning av surveydata. Arbetet i den så kallade DDI-alliance syftar till att skapa en internationell XML-standard för hur surveydata kan beskrivas i maskinläsbar form. Spridningen av standarden sker i allt snabbare takt. Vi som är aktiva inom surveyområdet kommer att få höra mer om detta framöver. Redan i början av december hölls en europeisk användarkonferens i Göteborg.

Annars var det dominerande inslaget på konferensen metodproblem i surveyer med samhällsvetenskapliga tillämpningar på individer och hushåll. Ämnet för flera sessioner var surveyer med blandade insamlingsmetoder (mixed-mode surveys). Blandade datainsamlingsätt, med till exempel både telefonintervju och enkät i samma survey, har man använt länge men forskningen och studierna på detta har verkligen tagit fart och många observationsstudier av varierande kvalitet görs som bidrar till en ökad samlad kunskaps- och erfarenhetsbank. Bland annat har man inom ramen för undersökningen The British household panel survey genomfört en hel del experiment och

studier vars resultat vi kan förvänta oss få ta del av den närmaste tiden. Två andra teman som återkom på flera sessioner var metoder för att integrera data från olika källor och metodik kring kvalitetsindikatorer.

Det var förhållandevis få presentationer som tog upp estimation. Några som förekom handlade om vägning med svarsbenägenheter (propensity scores). De arbeten som görs inom detta förefaller inriktade på att söka och hitta en tillförlitlig modell för att korrigera bias på grund av svarsbortfall. En omständighet man kan ifrågasätta i detta arbete är att man för att skatta sannolikheten för svar använder modeller med mycket låg förklarad variation. Frågan är om inte dessa lika gärna kan introducera som reducera bias.

Konferensen arrangerades på ett föredömligt sätt. Allt fungerade som det skulle och inte en dator eller projektor strejkade vid något tillfälle. De stora och flertaliga lokalerna med toppmodern teknisk utrustning på universitetet var väl lämpade att husera en så stor konferens. Tacksamt var också att fri trådlös uppkoppling erbjöds alla deltagare, något som mer och mer framstår som ett måste för alla som tänker sig att arrangera en konferens värd namnet.

Utöver ordinarie sessioner hölls det också ett möte för medlemmarna i ESRA. Man kan konstatera att ESRA vuxit i medlemsantal och att konferensen volymmässigt har utvecklats mycket sedan den första hölls i Barcelona 2005. ESRA har också startat internettidsskriften Survey Research Methods. ESRA:s webbplats finns på <http://survey methodology.eu/>. Just storleken på konferensen var en av de frågor som togs upp på mötet. Man har nått en kritisk gräns med den volym som rådde i Lausanne. En ny styrelse valdes på mötet och Willem Saris efterträdde på ordförandeposten av Patrick Sturgis från University of Southampton. Nästa ESRA-konferens blir 2013 men var den kommer att hållas är inte bestämt.

Konferensmiddagen avnjöts på båten Lausanne som förde oss runt den nordöstra delen av Genevesjön.

ANDERS HOLMBERG OCH JOAKIM MALMDIN,
STATISTISKA CENTRALBYRÅN



ORDFÖRANDEN HAR ORDET

”Innehåll: ,risk-benefit’, tragedi och glädje samt rättelse”

Julen närmar sig och ännu ett verksamhetsår går mot sitt slut. När jag skriver detta är det bara några skälvande timmar kvar av sommartiden och mörkret är kompakt utanför mitt fönster.

Måste nu passa på att framföra en rättelse. I min förra spalt skrev jag om Erik Leander och det positiva han gjorde för oss studenter och ber nu om ursäkt för ett klart felaktigt ordval. Jag vill nu förtydliga att Erik har repat sig efter flera allvarliga sjukdomar och vid 86 års ålder lever och bor i sitt hem. Jag ber om ursäkt om jag vilselett er läsare och jag är glad för att Erik mår bra och att jag fått tillfälle att framföra mitt tack.

År 2011 har varit händelserikt. En kraftig jordbävning i Japan fick dramatiska följder. Det fruktansvärda som hände på Utöya i semestertid var förlamande och chockande och våra tankar gick till alla drabbade i Norge. Detta år kommer också att bli ihågkommet för stridigheter och frigörelse i ett flertal av Nordafrikas länder. Föregående år var det svininfluensan som satte skräck i oss. Många vaccinerade sig och nu i år har det pågått en utredning om det finns ett samband mellan vaccinet Pandemrix och risken att drabbas av den neurologiska sjukdomen narkolepsi som ger ett okontrollerbart sömnbehov. Enligt Läkemedelsverkets fallstudie var den relativa risken för de barn och ungdomar som vaccinerats 6,6 gånger högre än för de barn och ungdomar som inte vaccinerats. Hur stor var då den absoluta risken? På Läkemedelsverkets webbplats kan man läsa att detta också kan uttryckas som att det uppkom 3,6 fler fall av narkolepsi med kataplexi (plötslig



På återseende jubileumsåret 2012, då FMS har funnits i 25 år! Det ska vi fira!

kraftlöshet i muskler) per 100 000 vaccinerade barn och ungdomar. Slutsatsen är att det finns ett samband.

Att alla läkemedel och vacciner har både fördelar och nackdelar är ju känt. Fördelarna

»Att alla läkemedel och vacciner har både fördelar och nackdelar är ju känt.»

måste givetvis kraftigt överväga nackdelarna för att patienter och samhälle ska ha nytta av produkterna. Val av dos är en av de viktiga frågeställningarna när fördelar ska balanseras mot risker. Eftersom positiva behandlingseffekter och biverkningar oftast mäts på helt olika skalor och ibland kommer från olika källor så krävs det

särskilda metoder och eventuellt indexeringar för denna typ av analys. ”Number Needed to Treat” (NNT) och ”Number Needed to Harm” (NNH) är exempel på relativt populära index som har kritiserats en del. Vi i FMS planerar att ordna en konferens med tema ”risk-benefit” under våren. Detta kommer antagligen att ske tillsammans med vår danska systerförening och EFSPi (European Federation of Statisticians in the Pharmaceutical Industry). Om möjligt vill vi bredda oss utanför läkemedelsindustrin och täcka även andra områden. Håll utkik efter mer information efter nyår. Vi har också ett årsmöte i Skåne den 22/3 att se fram emot. Då ska vi hålla till på Sveriges lantbruksuniversitet i Alnarp som är mycket sevärt.

Under hösten har vi haft två evenemang. Statistikerträffen i Göteborg (se referat på sidan 6) samt ett höstmöte på AstraZeneca i Södertälje i november som FMS arrangerade tillsammans med Surveyföreningen. Temat var registerbaserad statistik och de inbjudna föreläsarna kom från SCB, Karolinska Institutet och AstraZeneca. Styrelsen har också arbetat med den nya webbplatsen under Statistikfrämjandets webbplats. Besök den gärna - <http://statistikframjandet.se/fms/> - en webbplats som är under utveckling.

Tack för i år. På återseende jubileumsåret 2012 då FMS har funnits i 25 år! Det ska vi fira!



ANNA TORRÅNG



Aten.



Musiker i New Orleans



**Vi ses på
vårkonferensen
i Alnarp
22-23 mars!**

KALENDARIUM

DET HÄNDER 2012

	KONFERENS	PLATS	WEBBPLATS
16-18 feb	Statistical Practice	Orlando, USA	www.amstat.org/meetings/csp/2012/
11-15 mars	Vinterkonferens, Spatial Statistics	Orsa Gönklitt	www.du.se/spatialstatistic
22-23 mars	Svenska statistikfrämjandets vårkonferens	Alnarp	www.statistikframjandet.se
10-12 april	Fifth International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance	Venedig, Italien	maf2012.unive.it/
30 maj – 1 juni	European Conference on Quality in Official Statistics	Aten, Grekland	www.q2012.gr
4-6 juni	Quality and Productivity Research Conference	Long Beach, USA	www.qprc2012.com
10-14 juni	Nordic Conference on Mathematical Statistics	Umeå	www.nordstat2012.se
11 – 14 juni	Fourth International Conference on Establishment Surveys	Montreal, Kanada	www.amstat.org/meetings/ices/2012
18-20 juni	SCB:s och Örebro universitets sommarskola. "Asking and answering survey questions: Cognitive and communicative processes"	Örebro	www.oru.se/hh/summerschool2012
24-27 juni	International Institute of Forecasters	Boston, USA	forecasters.org/isf/index.html
3-6 juli	International Statistical Ecology Conference	Oslo, Norge	www.cees.uio.no/news/2010/isec2012.html
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

Figure 1: Expenditure poverty. Proportion of the population below the lower poverty line in Bangladesh 2003.

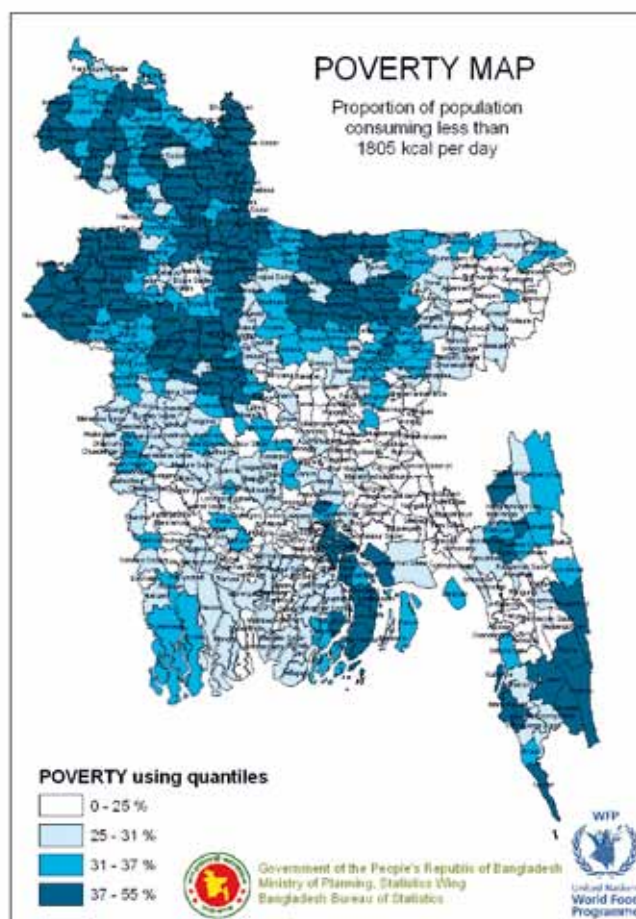


Figure 2: Stunting in children under five years of age in Bangladesh 2003.





Bild 1a. Flygbild med utpekade platser för geografiska företeelser som kan behöva återges i en karta.

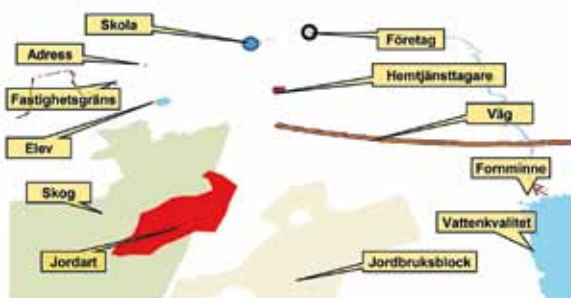


Bild 1b. En grafisk återgivning av de geografiska företeelserna i Bild 1a.

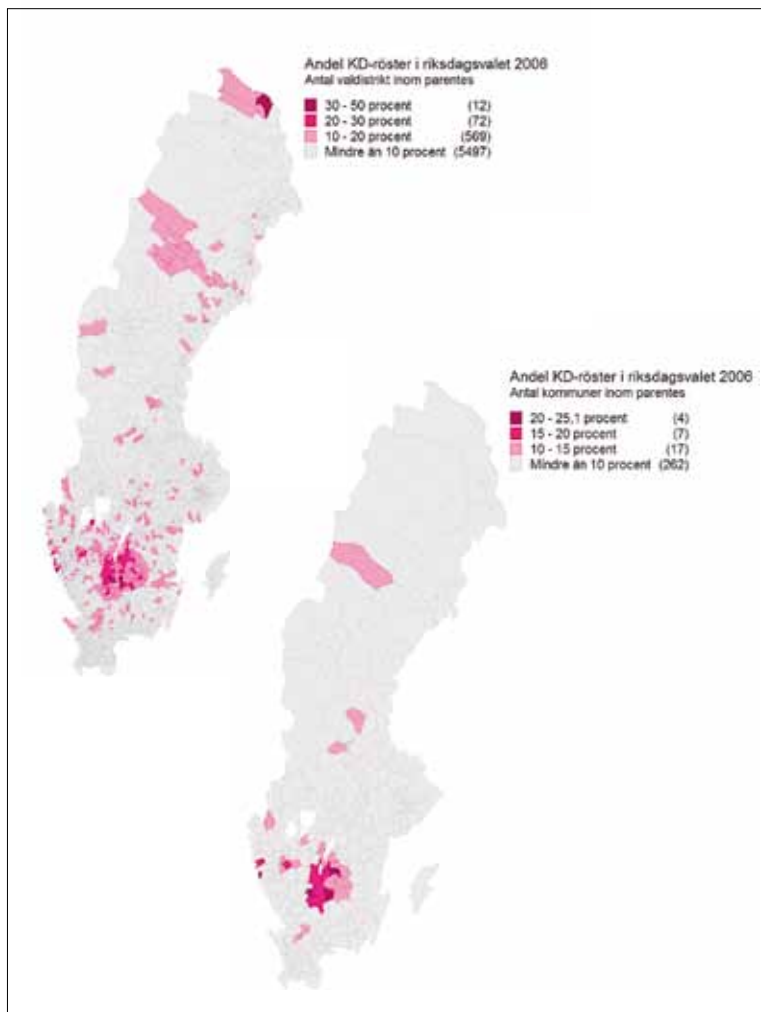


Bild 8. Skalnivåns betydelse i koropletkartor. Andel som röstade på Kristdemokraterna i valet 2006. I den vänstra kartan visas andelarna per valkrets, i den högra per kommun.



Bild 6. Järnvägen flyttas ut från vägen när kartskalen blir mindre och det karterade området större. Källa: Geodata © Lantmäteriet.