

HUR VÄGER MAN  
NYTTA MOT RISK?

SIDAN 21

JON RAOS  
FANTASTISKA RESA

SIDAN 14



# Qvintensen

SVENSKA STATISTIKFRÄMJANDET

NR 3 2012

USA firade  
 $\pi 10^8$  med paj

SIDAN 17

Data-  
beriknings-  
metoden

SIDAN 18

TEMA/

## Koders betydelse

SIDORNA 6-10

**Det känns som arbetsmark-  
naden för statistiker är god,  
men att mixen av inriktning-  
ar håller på att ändras.»**

Anna Ekman  
SIDAN 20

# 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](http://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](http://www.sas.com/sweden)



Leif Petersson, Anton Färnström och Sandra Sandberg.



Datinsamling pågår.



Nytta mot risk



Howard Hogan

## Innehåll

### TEMA: KODERS BETYDELSE/

Tolv procent fel.....	6
Ökningen av anmälningar av olovligt förfogande.....	10

### KAN MAN STRUNTA I ATT SAMLA IN EGNA DATA .....

11

### MÖT JON RAO

– en av världens främsta forskare inom statistik..... 14

USA:S DEMOGRAFER FIRAR NERDVANA..... 17

DATABERIKNINGSMETODEN .....

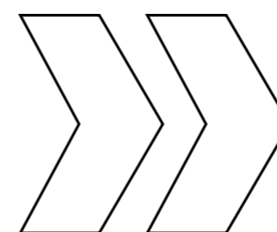
18

KARIN DAHMSTRÖM IN MEMORIAM .....

27

EN DEL MYTER SITTER SOM BERG .....

28



**Surveyteorin ses som tråkig; elementär matematik, inga tekniska utmaningar. Det har aldrig varit sant»**

Jon Rao, sidan 14



## Våra föreningar

### SURVEYFÖRENINGEN

Konsten att dra slumpmässiga urval .....

19

### FÖRENINGEN FÖR MEDICINSK STATISTIK

Svensk läkemedelsindustris framtid.....

20

Nytta mot risk.....

21

### INDUSTRIELL STATISTIK

Att vårda en process, mäta, räkna, rita

– och övertyga.....

25

## Fasta spalter

I HUVUDET PÅ EN REDAKTÖR .....

4

ORDFÖRANDEN HAR ORDET .....

5

KALENDARIUM .....

26





I HUVUDET PÅ EN REDAKTÖR

## Svåra problem kräver välutbildat folk

Läkemedel som har effekt har också risk för biverkningar. Inga risker, ja då finns antagligen inte heller någon effekt.

**Med läkemedel** är det som med det mesta: fördelar ska vägas mot nackdelar. Det är svårt. För- och nackdelar är av olika karaktär och dessutom svåra eller helt omöjliga att sätta siffror på. Ändå måste nytta vägas mot risk. Anna Torrång skriver på sitt medryckande sätt om hur det kan göras för läkemedel i det här numret av Qvintensen.

**Svåra problem kräver** välutbildat folk, som i sin tur har fått sin utbildning av ännu mer välutbildat folk. I hela världen tycks det råda brist på statistiker, inte minst på högre akademiska nivåer. Bo Wallentin sätter fingret på bristen på doktorer som vill gå vidare och bli docenter, som behövs för att utbilda nya doktorer.

**Hur man än gör** med nytta och risk, eller vilka välgrundade beslut som helst, krävs det underlag. Det krävs data. Och data uppstår inte av sig själva. Det finns nästan alltid någon mänsklig handling i bildandet av data, och där det finns människor, där finns det misstag och felgrepp. Anton Färnström, Leif Petersson och Sandra Sandberg berättar om en studie av fel i kodningen av brott. Data kan också avsluta livet på myter, som den att sommaren är inbrottstjuvars högsäsong. Om du nyligen har slutat tro på att det är just på

sommaren tjuven slår till, tänk efter, varför upphörde du att tro på det?

**Det går inte** att helt ersätta data med modeller och matematik även om vi alla är tvungna att fylla i luckorna i vårt vetande med någon form av antaganden. I slutet av det här numret har vi nöjet att återtrycka Tore Dalenius satir där han driver med innehållslöst matematiserande och förespråkar "databerikningsmetoden".

**Det är dyrt** att samla in data själv, om man inte använder nyssnämnda metod. Men det verkar finnas hur mycket data som helst ändå. Det är ju inte riktigt sant, för när man letar efter något speciellt visar det sig ofta att det inte finns. Ändå är det en ständig fråga för stora data-användare hur man kan använda data som redan finns. Även det har vi med i det här numret av Qvintensen.

Mycket nöje!

**»I hela världen tycks det råda brist på statistiker»**

DAN HEDLIN



FOTO: MONICA HOLMBERG

## Qvintensen

Ansvarig utgivare  
Bo Wallentin  
Redaktör  
Dan Hedlin, 070 297 10 27  
Ingegerd Jansson (biträdande)

Redaktion  
Hans Ekholm, Främjandet  
Alf Fyhrlund, Främjandet  
Marie Linder, FMS  
Kadri Meister, Surveyföreningen  
Magnus Arnér, Industriell statistik  
Lars Rönnegård, Cramérsällskapet

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.  
Annonsutskick per e-post bokas med Svenska statistikfrämjandets sekreterare.  
Priset är 6 000 kronor för institutionsmedlem eller motsvarande, och 8 500 kronor för övriga annonsörer. På alla priser tillkommer 25 % moms.

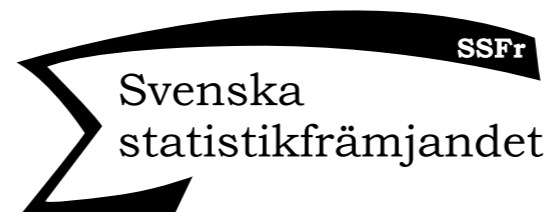
Svenska statistikfrämjandet

### SVENSKA STATISTIKFRÄMJANDETS STYRELSE

Ordförande  
Bo Wallentin, 018-471 51 58,  
Bo.Wallentin@statistik.uu.se  
Vice ordförande  
Jun Yu, 090-78 68 468  
Sekreterare  
Olivia Ståhl  
Skattmästare Jonas Selling  
Surveyföreningen  
Lennart Nordberg, 0709-15 08 02  
FMS

Mats Rudholm, 031-703 73 72  
Industriell statistik  
Ylva Tingström, 08-522 17 016  
Cramérsällskapet  
Maria Karlsson, 090-786 55 96  
Övriga ledamöter  
Mathias Lindholm, Marie Wiberg, Stig Johan Wiklund

Adress  
Svenska statistikfrämjandet  
c/o Olivia Ståhl  
Statistiska institutionen  
Stockholms universitet  
106 91 Stockholm  
E-post sekfram@gmail.com  
Webbplats www.statistikframjandet.se



## ORDFÖRANDEN HAR ORDET VI BEHÖVER FLER ERFARNA DOKTORER

**D**å detta skrivs har de senaste Olympiska spelen just avslutats. Nu är det fyra år till nästa. Det låter långt, medan OS i Peking känns nästan som det var igår. Tittar man framåt blir tiden längre. Även för en nyantagen doktorand är fyra år en lång tid, en tid som fylls med hårt arbete och krav på prestation. Fyra års finansiering är vad en doktorand garanteras. Efter doktorsexamen i statistik väntar som tur är en god arbetsmarknad.

Antal disputerade i ämnet statistik var under de senaste tre kalenderåren cirka 17. Umeå universitet har klart flest och fler än universiteten i Lund, Stockholm och Uppsala tillsammans. En känsla är att antalet lediga lektorat i statistik under samma tid också är av den storleksordningen. Ekvationen kanske synes gå ihop, men betänker man att det finns en annan arbetsmarknad för statistiker med ofta goda lönevillkor (Dagens Industri 2012-05-22), inses att här finns ett stort problem.

**En tillvaro vid** universitet kräver intresse för både forskning och undervisning. Alla har inte det. Av de tio<sup>1</sup> som disputerat vid min institution under de senaste tio kalenderåren var det två som valde att göra karriär vid svenskt universitet. Fyra har valt internationell karriär. Att nydisputerade skaffar sig internationell erfarenhet är något vi bör uppmuntra, men då svenska lektorat har liten garanterad forskningstid kan det bli svårt att få dem att återvända.

Vid utlysning av lektorat i statistik erhålls få sökande och de är ofta relativt nydisputerade. Behöriga sökande kanske inte finns. För andra samhällsvetenskapliga ämnen är situationen oftast den omvända. Ej sällan har man att välja mellan professorskompetenta sökande. Den här ojämlikheten i urval och kompetens försvårar för ämnet statistik att hävda sig inom de samhällsvetenskapliga fakulteterna.

**Det blir då** än viktigare att våra lektorer inte stannar upp i forskarkarriären. Vi behöver inte enbart

<sup>1</sup> Justerade tal



Bo Wallentin, ordförande.

**»Våra masterutbildningar är av strategisk betydelse»**

nydisputerade. Undervisning och handledning på avancerad nivå fordrar en annan kunskap än den också viktiga undervisningen på grundläggande kurser i statistik. På institutioner behövs därför en kompetensförsörjningsplan som gör det möjligt för lektorer att meritera sig även forskningsmässigt. Forskningsanslag kan enbart till del finansiera en sådan så institutionerna måste också vara beredda att ge tid för denna nödvändiga utveckling via grundutbildningsanslag.

**Kompetensförsörjning i ämnet** statistik berör inte enbart statistikinstitutionerna. Det berör också i högsta grad producenter och användare av statistik och naturligtvis också övriga forskarsamhället. Utbildning av och hjälp till kommande användare och producenter av statistik är oerhört viktigt för kvaliteten i den dagliga hanteringen av statistisk information. Inte minst kan vi se i Qvintensen vilken betydelse statistik och statistiska metoder har i olika delar av vårt samhälle.

**Ansvar för att förse** arbetsmarknaden med kvalificerade statistiker åvilar universiteten. Medeltilldelningssystem som baseras enbart på prestation vad gäller publikationer, erhållna externa medel och produktion av doktorander och inte alls tar hänsyn till arbetsmarknadens behov arbetar emot oss. Med den internationaliserade rekryteringen till forskarutbildningen finns inte brist på kvalificerade sökande. Problemet är brist på doktorandplatser. Fler doktorander ställer större krav på tillgång till handledare och aktiva forskare i statistik. Våra masterutbildningar är av strategisk betydelse för att förse arbetsmarknaden med kvalificerade statistiker men också för att generera bra doktorander som med professionell handledning blir klara inom en olympiad.

Bo Wallentin, ordförande

Brottsförebyggande rådet (Brå) har undersökt kvaliteten på de uppgifter som ligger till grund för registret över anmälda brott, uppklarade brott och misstänkta personer. Resultatet visar att ungefär tolv procent av samtliga anmälda brott är felkodade och att problemen framförallt återfinns på statistikens mer detaljerade redovisningsnivåer.

En kvalitetsstudie av brottskodstillämpningen

# TOLV PROCENT FEL

**B**rottsstatistiken har stort nyhetsvärde. Några rubriker som har förekommit under 2012 är: Färre anmälda snatterier (Sveriges Radio), Fler bränder – mindre klotter (Helsingborgs Dagblad), Bedrägerierna ökar lavinartat (Svenska Dagbladet). I artiklarna redovisas ofta det exakta antalet anmälda brott och siffrorna betraktas som en sanning, en grundplåt, som sedan ska kommenteras och analyseras. Detta förhållningssätt till statistiken bygger på en stark tilltro till de svenska registren, och till den svenska officiella statistiken. Men är tilltron för stark? All statistik, även statistik baserad på register, innehåller felkällor som kan vara av sådan storlek att de publicerade uppgifterna bör betraktas som skattningar snarare än sanningar.

Brottsförebyggande rådet har undersökt en av de största felkällorna i brottsstatistiken: klassificeringen av brott.

**Temats författare från vänster Leif Petersson, Anton Färnström och Sandra Sandberg.**

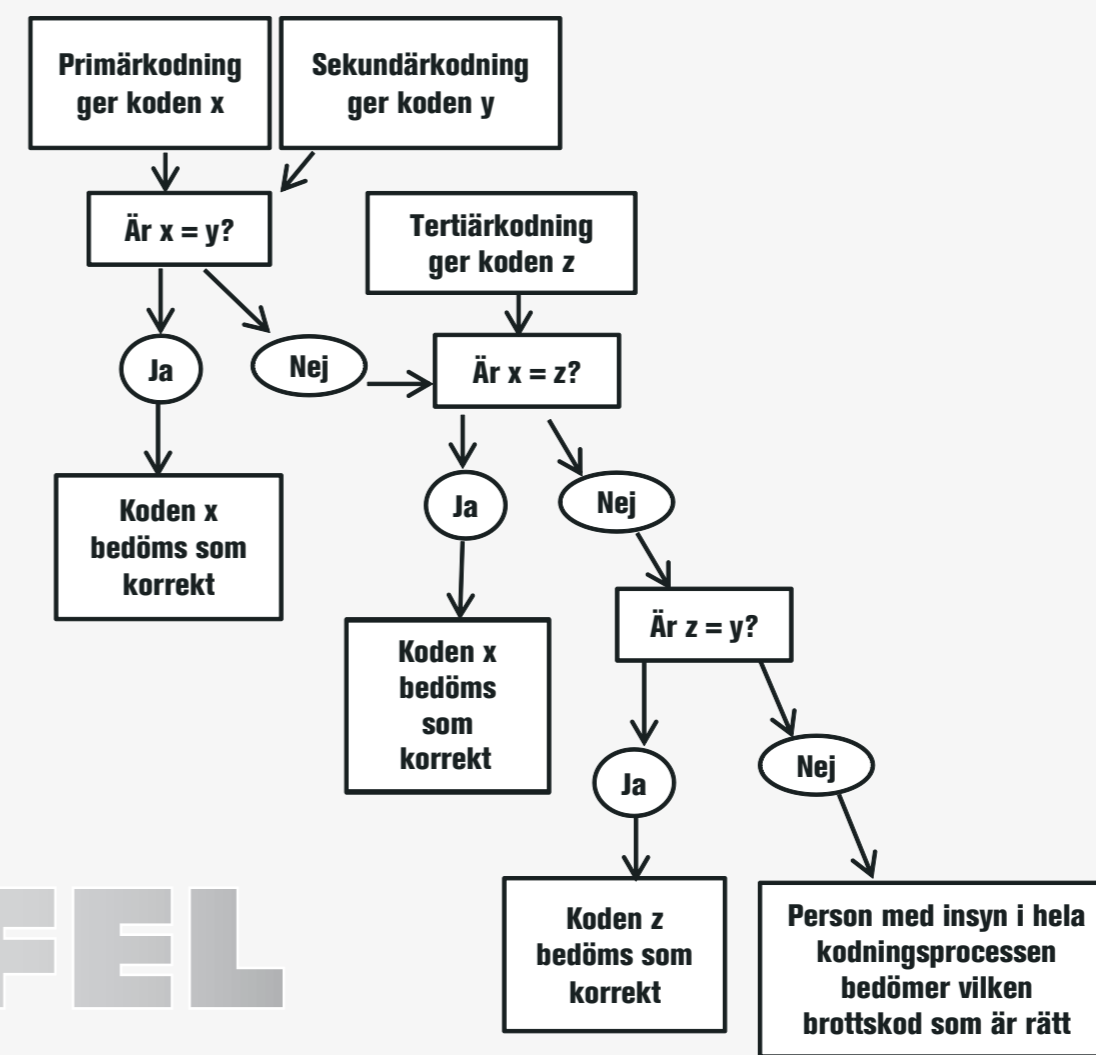


**Brottskodssystemets byggstenar**  
Brottsstatistiken innehåller uppgifter om anmälda brott, uppklarade brott och misstänkta personer. Den framställs av Brottsförebyggande rådet och utgör en central källa till information för beslutsfattare, forskare, media och allmänhet som har intresse av att följa ärendegenomströmningen inom rättsväsendet.

Det så kallade brottskodssystemet (referens [2]) är ett klassificeringssystem som kombinerar information om lagrum som reglerar den brottsliga handlingen med omständigheter kring brottet av allmänt intresse, som relationen mellan offer och gärningsperson eller ålder på den som utsatts för brottet.

Vi har genomfört en studie av hur de som registrerar polisanmälningar

Figur 1. Illustration av oberoende kontrollkodning med slutavgörande.



använder brottskoderna. Studien har fokuserat på koder som fungerar som resttyper, eller uppsamlingstyper, exempelvis Övrigt bedrägeri, vilken är en av flera brottskoder som kan anges för bedrägeribrotten (9 kap 1–3 §§ BrB). Förutom resttyperna studerades tillämpningen av koder för misshandelsbrott (3 kap 5–6 §§ BrB), där skillnad görs på när offer och gärningsman är ”i nära relation” respektive ”ej i nära relation”.

De valda brottskoderna är sådana där problem med brottskodstillämpningen förväntades. Vi drog ett stratifierat oberoende slumpmässigt urval av 1 600 polisanmälningar.

### Metod

För att avgöra om en brottskod hade använts korrekt eller inte för en viss polisanmälan används oberoende kontrollkodning med slutavgörande [1]. I korthet innebär metoden att två personer, i detta fall en jur.kand. och en

juristikstudent, har fått granska fritexterna i polisanmälningarna och därefter kodat varje anmälan med den kod de ansett vara mest korrekt. Granskarna har kodat utan kännedom om varandras bedömningar eller polisens ursprungliga kodtillämpning.

Processen, som illustreras i figur 1, har legat till grund för en förteckning över korrekta koder. Förteckningen har sedan kunnat användas för att ta fram olika mått på felkodningsfrekvenserna inom brottsstatistiken.

Nettofelet är skillnaden i antal mellan de anmälda brotten som de rapporteras i den officiella statistiken och som de skulle ha rapporterats om polisanmälningarna getts korrekta brottskoder. Nettofelet beräknas med hjälp av två komponenter, dels felaktigt medtagna ärenden, dvs. ärenden som registrerats som den studerade brottstypen men borde getts en annan brottskod, dels felaktigt uteslutna ärenden, ärenden som borde kodats som den studerade

brottstypen men som ursprungligen gavs en annan brottskod. Tillsammans kan felet ta ut varandra vilket innebär att det trots stora kodningsfel kan finnas små nettofel. Förutom nettofelet har antalet felaktigt medtagna ärenden redovisats, för att illustrera felkodningsnivån inom olika kategorier av brott.

### Mer än var tionde anmälan var felkodad

Resultatet av studien visar att tolv procent (plus/minus två procent) av samtliga polisanmälningar fick en felaktig brottskod år 2010 och att det för enskilda brottstyper finns betydligt större osäkerheter. För brottskoderna som var i fokus för studien, resttyperna, var felkodningsfrekvensen i de flesta fall 30 procent. De felaktiga tillämpningarna inom resttyperna innebär att de har överskattats i den officiella statistiken medan andra brottstyper underskattats. Resultatet för gruppen bedrägeri illustreras i figur 2. Som framgår av figuren »»»»





Skimmat uttag? Bedrägerier är en typ av brott där felkodning innebär att vissa bedrägeribrott överskattas, medan andra, som datorbedrägerier, underskattas.

**»Antingen höjs kraven på de som registrerar anmälningar om brott, eller så anpassar man redovisningsnivån»**

»»» **TOLV PROCENT FEL**

har övrigt bedrägeri överskattats med 35 procent i den officiella statistiken, samtidigt som datorbedrägeri underskattats med 15 procent och bedrägeri med hjälp av internet underskattats med 19 procent.

Liknande resultat som för bedrägeribrotten finns för samtliga kategorier brott som innehåller en "övrigt"-kod, förutom för skadegörelsebrotten, som uppvisade en låg felkodningsnivå.

Även brottskoderna för misshandel i nära relation visade sig ofta vara felregistrerade. Ungefär vartannat brott som registrerats som misshandel mot man i nära relation var en annan typ av misshandel mot man. För misshandel mot kvinna

var motsvarande siffra ungefär vart fjärde brott.

**Orsakerna till felkodningar**

Det fanns inkonsekvenser och otydligheter i instruktionerna för kodning av brott. De var dock inte systematiska eller av en sådan omfattning att de kan förklara studiens resultat. En annan förklaring som ligger nära till hands är att övrigt-koderna är ett bekvämt alternativ för den som registrerar brott.

Den extensiva användningen av övrigt-koderna hade för en brottskategori en oväntat positiv konsekvens. För kategorin rån fanns ofta bristande underlag

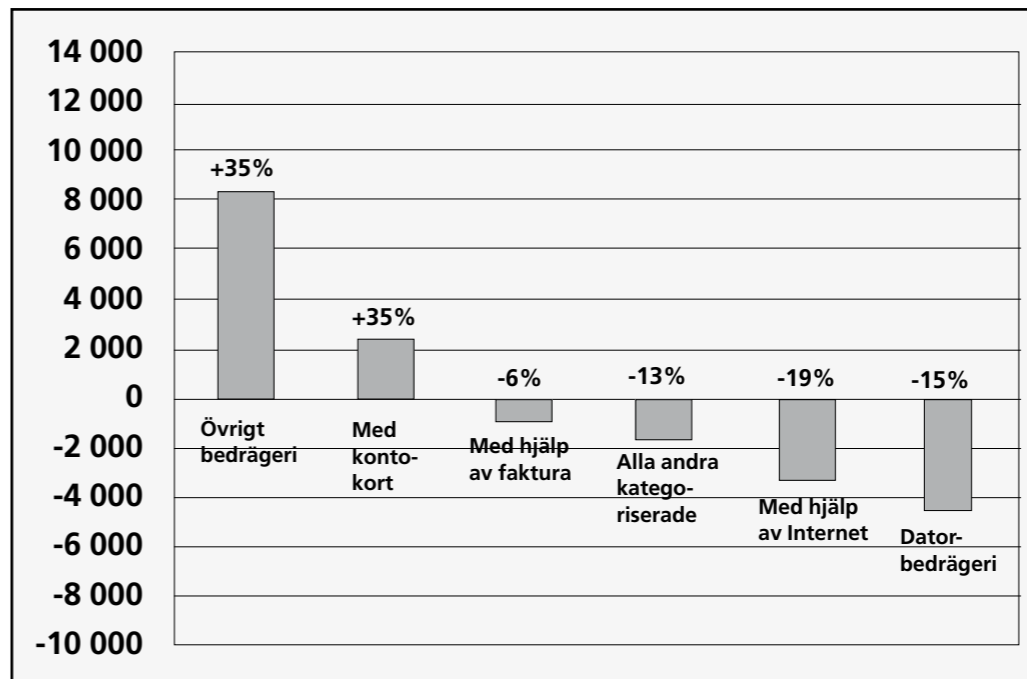
vid registreringarna, vilket enligt instruktionerna skulle ha inneburit att brotten placerats i kategorin rån mot privatperson (funktionsnedsatt) och gett en överrapportering med flera hundra procent för den brottstypen. Dessa anmälningar var dock kodade som övrigt rån, vilket innebär att de som registrerat anmälningarna har bortsett från en otillräcklig instruktion och istället registrerat det som kan anses som rimligt.

För de brott där brottskoderna innehåller alternativen "i nära relation" och "ej i nära relation" kan orsaken till felrapporteringen gå att hitta i definitionen av "i nära relation" i brottskodsystemet. Studien visade att polisen använde begreppet i en vidare bemärkelse än vad som avses i instruktionerna för kodning av brott, där det enbart avser de fall då offret och gärningspersonen är eller har varit gifta eller sammanboende under äktenskapsliknande former eller har gemensamma barn även om de inte bor tillsammans. Definitionen är därmed snävare än vad som avses i vardagligt språk och skiljer sig även från hur begreppet används av övriga myndigheter i rättskedjan, där exempelvis särbo kan anses vara en nära relation.

De flesta kodningsfelen var begränsade

**OM BROTTSKODSSYSTEMET**

- Brottskodssystemet infördes 1965 i samband med att den nya brottsbalken trädde i kraft.
- Brottskoden består av fyra siffror.
- Brottskodssystemet möjliggjorde automatisk databehandling.
- I början infördes 115 olika brottskoder, men ytterligare har tillförts och vid senaste revideringen räknades till 451 aktiva brottskoder, se Figur 3.
- Brottskoderna används i statistiken över anmälda brott, uppklarade brott och misstänkta personer.
- Brottsförebyggande rådet, i samråd med polisen, åklagarmyndigheten, Tullverket och ekobrottsmyndigheten, beslutar om införande av nya koder.



Figur 2. Nettofelet är skillnaden i antal mellan de anmälda brotten som de rapporterats i den officiella statistiken och som de skulle ha rapporterats om polisanmälningarna getts korrekta brottskoder. Olika sällsynt förekommande bedrägerityper är samlade i gruppen "alla andra kategoriserade".

till artbestämelse inom en viss kategori. Till exempel skulle de felkodade fallen av bedrägeri oftast getts en annan bedrägerikod, men inte i något fall placerats utanför bedrägerikategorin. Enskilda brottstyper uppvisade också hög kvalitet på brottskodstillämpningen, som till exempel butikstölder och stöld ur eller från motordrivna fordon. Även de anmälda skadegörelsebrotten hade en mycket låg felkodningsfrekvens.

Det skulle föra för långt att i denna artikel diskutera detaljer om bland annat statistiska test. Vår rapport [3] finns tillgänglig på internet. I den finns också förslag på alternativa brottskoder.

**Generella slutsatser**

Senast en studie av brottskodstillämpningen genomfördes var 1978 [4]. Då producerades kriminalstatistiken av SCB och studien omfattade 15 000 polisanmälningar. SCB rapporterade en generell felkodning på tio procent och även i deras studie fanns betydande problem med att ange relationen mellan offer och gärningsperson i misshandelsärenden samt en hög andel felkodade anmälningar i övrigt-koderna.

Sedan 1978 har nya datasystem införts, ett flertal polisdistrikt har upphört och flera hundra nya brottskoder har tillkommit. Figur 3 visar utveckling i antal brottskoder. Mot bakgrund av att resultatet i Brås studie är samstämmigt med resultatet i SCB:s studie går det att anta att de identifierade kvalitetsbristerna beror på brottskodsystemets utformning.

En gemensam nämnare för de stora kvalitetsbristerna är att de återfinns inom områden, och på redovisningsnivåer, där det inte på ett enkelt sätt går att se att de identifierade kvalitetsbristerna beror på brottskodsystemets utformning. Fallet med kodningar av relationer inom misshandelsärenden är ett bra exempel på detta. Den efterfrågade definitionen på "i nära relation" återfinns inte inom verksamheterna och är inte heller intuitiv, det enda syftet med definitionen är att sammanställa statistik.

Man kan föreställa sig två alternativ för att komma tillrätta med kvalitetsbristerna. Antingen höjs kraven på de som registrerar anmälningar om brott, eller så anpassar man redovisningsnivån

så att den bättre stämmer överens med hur de aktuella myndigheterna behöver sortera och klassificera anmälningarna i sitt dagliga arbete.

Utgångspunkten för det andra alternativet är att allt registreringsarbete som inte krävs i den egna tjänsteutövningen innebär en risk för kvalitetsbrister i statistiken och därför bör undvikas. Det är inte heller rimligt att lyfta statistik användarnas behov som enda argument för att ha en hög detaljeringsgrad, med tanke på att statistiken enbart kan användas om den uppnår en viss kvalitet.

Brå saknar i stort möjligheter att komma tillrätta med bristerna i dagens brottskodssystem, men det pågår ett omfattande utvecklingsarbete tillsammans med polisen, åklagarmyndigheten och Domstolsverket, vilket bland annat kommer att leda till att brottskoderna fasas ut.

Resultatet av studien, och analysen av kvalitetsbristernas orsaker, kommer att föras fram inom ramen för det arbetet, då nya strukturer för det statistiska systemet utformas.

ANTON FÄRNSTRÖM,  
SANDRA SANDBERG,  
BROTTSFÖREBYGGANDE RÅDET

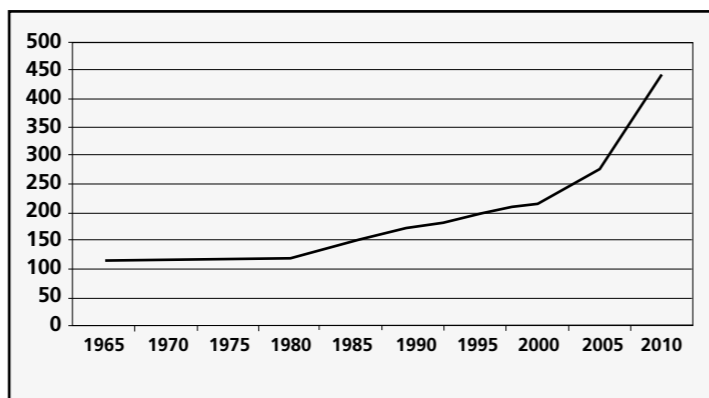
**Referenser**

- [1] Biemer P.P. and Lyberg, L.E. (2003). Introduction to Survey Quality. New York: Wiley.
- [2] Brottsförebyggande rådet (2011). Kodning av brott. Anvisningar och regler. Version 10.0, reviderad december 2011.
- [3] Brottsförebyggande rådet (2012). Användningen av brottskoder. En kvalitetsstudie inom kriminalstatistiken.
- [4] Statistiska centralbyrån, SCB (1978). Fel vid kodning av brott – Kvalitetskontroll av brottstatistiken. Promemoria 1978:1.

Leif Peterson berättar på nästa sida om olovligt förfogande, hur förändrade registreringsrutiner hos polisen kan påverka statistiken. »»»

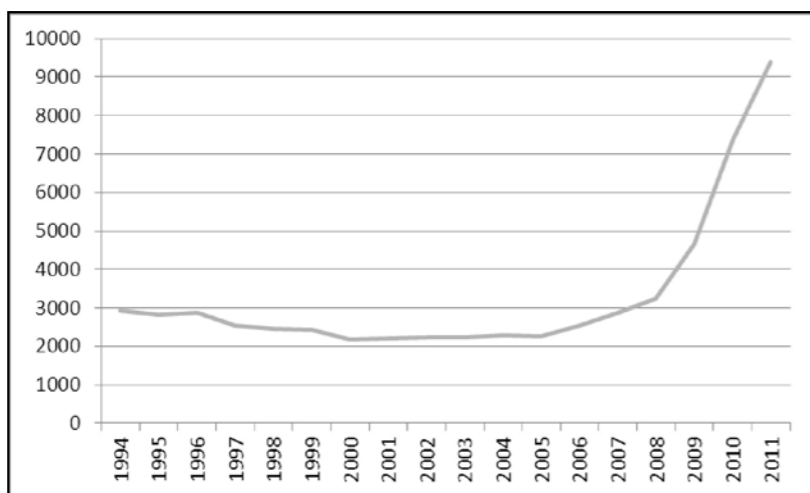
»»»» TOLV PROCENT FEL

Figur 3. Även för registerstatistik är det viktigt att tänka på uppgiftslämnarbördan, något som inte minst blir tydligt när man ser på utvecklingen av antalet brottskoder.



# Ökningen av anmälningar av olovligt förfogande

Det sägs ibland att systematiska över- eller underskattningar har liten betydelse om man främst intresserar sig för förändringar. Man antar att de systematiska felen är ungefär konstanta över tid och tar ut varandra när man skattar förändringen. Men stämmer det alltid?



Figur 4. Brottsypen olovligt förfogande ökade dramatiskt, dock inte beroende på brottsligheten...

**N**är vi framställer statistik, granskar och analyserar vi den framtagna statistiken noga inför publicering. I statistiken över anmälda brott 2010 upptäckte vi att brottsypen olovligt förfogande hade ökat anmärkningsvärt mycket, med närmare 60 procent, under redovisningsåret, se figur 4.

I vanlig ordning sökte vi efter enskilda fall eller polismyndigheter som kunde förklara ökningen (det finns 21 polismyndigheter). Där kunde vi inte hitta någon förklaring, utan ökningen visade sig vara generell.

Vid kontakt med företrädare för polisens kontaktcenter fick vi veta att man hade infört nya instruktioner för hur vissa typer av stölbrott där förövaren hade nyttjat det stulna godset skulle rubriceras. Fler brott av den typen blev kodade som olovligt förfogande.

Det var alltså regeländringen, inte brottsligheten, som ledde till att brottsypen ökade i redovisningen.

LEIF PETERSSON,  
BROTTSFÖREBYGGANDE RÅDET



FOTO: US CENSUS BUREAU, PUBLIC INFORMATION OFFICE

Även i USA intresserar man sig på US Census Bureau för möjligheten att i framtida folkräkningar utnyttja administrativa data.

Intresset för användningen av administrativa register som datakälla för officiell statistik har stadigt ökat under senare år.

I de nordiska länderna har man använt sig av administrativa data länge, men intresset ökar nu även i andra länder, främst kanske för att man ser en möjlighet att få tillförlitliga och aktuella data till lägre kostnad.

## Kan man strunta i att samla in egna data?

**I Storbritannien** genomfördes folkräkningen 2011 på traditionellt vis genom utskick av ett frågeformulär till hela befolkningen. Det är en väldigt kostsam och tidsödande historia som bara kan genomföras med långa mellanrum. Därför har man nu startat ett projekt för att undersöka hur man i framtiden kan utnyttja administrativa datakällor i kombination med surveydata. Även i USA intresserar man sig på US Census Bureau för möjligheten att i framtida folkräkningar utnyttja administrativa data. Det är exempel på länder där det bara för några år sen ansågs i princip omöjligt att använda administrativa data till statistik.

Det är lätt att hitta argument för att

använda administrativa data. Varför inte använda data som redan finns tillgängliga och som kostar lite. Just kostnaden är vanligen det starkaste argumentet. Efter initiala kostnader då data från det administrativa systemet ska anpassas till fortsatt användning i ett statistiskt system blir kostnaderna i allmänhet låga jämfört med urvalsundersökningar. I många fall håller administrativa data god kvalitet och har god täckning. Man sitter på en guldgruva av data som går långt tillbaka i tiden och som kan kombineras med andra administrativa data. Britt och Anders Wallgren skrev en artikel om detta i Qvintensen nr 1/12.

Man ska inte blunda för att det finns svårigheter också. Det finns många kvalitetsproblem med administrativa data, precis som det gör för surveydata, men de är av en delvis annan art. För många surveystatistiker krävs det nytänkande och ny inläring. För urvalsundersökningar har vi den klassiska designbaserade urvals- och inferensteorin som är populär på statistiska centralbyråer,

men något motsvarande finns inte för administrativa data. Det är ett av flera problem som börjat uppmärksammas på flera håll på senare tid, se till exempel Zhang (2011, 2012), Bakker (2012) och

**»Varför inte använda data som redan finns tillgängliga och som kostar lite?»**

»»»»



## »»» KAN MAN STRUNTA I ATT SAMLA IN EGNA DATA?

Hedlin et al. (2006). Li-Chun Zhang har också skrivit i Qvintensen nr 4/10 om sin forskning på området. Inom survey och sampling talar man om totalfelet (total survey error), dvs. det totala felet i en skattning som består av både varians och bias, där bias kan uppstå genom bortfall, dålig frågekonstruktion, intervjuarpåverkan, osv. Något motsvarande är önskvärdt för den totala felkattningen i en undersökning baserad på administrativa data.

Vad är det då för speciella förutsätt-

ningar som man skulle behöva ta hänsyn till för att kunna konstruera ett relevant mått för det totala felet? Här följer några exempel på sådant som kan vara viktiga felkällor.

- Statistikproducenten har inte kontroll över datainsamlingen utan är beroende av data som samlats in i ett annat syfte än att göra statistik. Hur vet vi att data mäter det vi vill mäta? Ibland vet vi att data inte mäter det vi som statistikproducenter vill. En utmaning är då att konstruera en modell för skillnaden mellan den in-

formation vi har och det vi vill skatta – och sedan försvara den modellen gentemot kritiker.

- Det uppstår olika typer av mätfel även i registerbaserade undersökningar. Man kan till exempel misstänka att kodningsfel utgör en viktig komponent vid användningen av en del administrativa register (se artikel i detta nummer på sidan 6).
- Hur vet vi att data avser den tidpunkt vi tror? Eftersläpningar i rapportering eller korrigeringar i efterhand gör att ett administrativt register i många fall inte är en fryst bild av tillståndet vid ett givet tillfälle.
- Det finns också exempel på regeländringar som leder till vad en statistikproducent skulle kalla korrelerade mätfel, alltså samma sorts ”fel” i data som uppkommer hos många objekt ungefär samtidigt.

I syfte att diskutera dessa frågor och hur man gemensamt skulle kunna driva utvecklingen ett litet steg framåt hölls nyligen en workshop i samarbete mellan SCB, den norska motsvarigheten SSB och Örebro universitet. Deltagarna kom från universitet och statistikbyråer

i länder med olika erfarenhet av användningen av administrativa data men med ett gemensamt intresse att utveckla användningen vidare.

Man kan konstatera att förutsättningarna för att använda administrativa data som källa till statistik ser väldigt olika ut i olika länder. Det beror till viss del på lagstiftning men också på kulturella skillnader. I länder som Sverige, Norge, Finland och Holland finns en mängd administrativa register, de har använts länge i statistikframställningen och den allmänna uppfattningen är att de ska användas. I länder som USA, England, Australien och Tyskland har man inte samma tradition, administrativa register finns i mer begränsad omfattning och allmänheten uppfattar lätt användningen av det som

**»Förutsättningarna för att använda administrativa data som källa till statistik ser väldigt olika ut i olika länder»**

Insamling av data från hästryggen. Folkräkning i USA 1930.



FOTO: US CENSUS BUREAU, PUBLIC INFORMATION OFFICE



US Census Bureau: Folkräkningen 1950.



FOTO: JOHAN STENEGÅRD

”Vi älskar fel”. Sexton registerentusiaster i ett ljusst atriumrum på Örebro universitet.

finns som ett intrång i den personliga integriteten. Trots att vi bara var 16 deltagare var alla dessa länder representerade på workshopen, och det gav en intressant inblick i olika sätt att se på saken.

Liksom för urvalsdata så är det de möjliga felkällorna som är det största intresset för en surveystatistiker som vill vidareutveckla användningen av administrativa data. ”We love errors”, som någon glatt proklamerade under middagens arbetspass. Men hur mäter vi dessa felkällor, och hur gör vi giltig inferens om våra registerbaserade skattningar?

Li-Chun Zhang från Statistisk centralbyrå i Norge och University of Southampton utgår från en utvidgning av den ursprungliga synen på totalfelet. Hans bild är att data genomlever två faser (Zhang 2012). I den första samlas primärdata in, det kan vara i en urvalsundersökning eller via administrativa register. Då kan olika typer av fel uppstå under mätning eller insamling, eller när data editeras och processas. I den andra fasen återanvänds redan insamlade data och ska i princip alltid integreras med andra befintliga data. Då uppstår nya felkällor, eller så kan felkällor från fas ett fortplantas till fas två. Det blir tydligt att det generella problemet handlar om att integrera olika datakällor, och genom att fokusera på det blir metodproblemen väldigt likartade oavsett den kulturella bakgrunden. Li-Chun pekade också på behovet av gemensamma koncept och begrepp, ett

universellt problem inom de flesta discipliner. Just systemtänkandet har betonats av Britt och Anders Wallgren (Qvintensen nr 1/12).

Även John Eltinge, US Bureau of Labor Statistics, förespråkar en utvidgad syn på totalfelet. Han föredrar att tala om risk, i meningen allt som kan orsaka osäkerhet och fel i data. De traditionella felkällorna i datainsamling kallar han ”aggregate risk”, men han talar även om ”systemic risk” som ett samlingsnamn på större händelser som ger stora effekter på datainsamlingen och datas kvalitet, t ex plötsliga ändringar i administrativa system.

När data ska integreras är det en fördel om det går att unikt identifiera objekt, t ex genom personnummer eller organisationsnummer. Om det inte finns någon unik identifierare kan andra identifierare användas för att matcha objekt, t ex namn, adress och diverse karakteristika, och matchningen kan utföras enligt någon sannolikhetsmodell. Då finns alltid risken att matchningen blir fel. Ray Chambers från University of Wollongong i Australien pratade om felaktiga länknings mellan objekt som ett mätfel och presenterade sin forskning kring hur man kan korrigera för denna typ av mätfel, oavsett om det handlar om att matcha registerdata eller data från en urvalsundersökning och ett register.

Anders Holmberg, SCB, som var initiativtagare och huvudsaklig organisatör av workshopen, inledde den avslutande diskussionen med

ett illustrativt exempel.

För att ta fram en skenbart enkel tabell som visar genomsnittlig energianvändning efter hushållstyp och inkomstkategori skulle det krävas ett flertal register och många kopplingar mellan dem. Hushåll är komplicerat i sig och skapas genom befolkningsregistret och adressregister med lägenhetsnummer. Beskattningsbar inkomst fås via Skatteverket. Dessa register finns redan och används naturligtvis flitigt för statistikframställning. Elanvändning skulle kunna fås genom uppgifter om hushållens abonnemang och abonnemangens koppling till elmätare, som är unik för lägenhet. På så vis kan alla uppgifter kopplas samman och ge en bild av energianvändningen. Det är fullt möjligt, men kräver förstas en initial kraftanstängning.

För att statistikproducenten sedan ska kunna beskriva och kommunicera datas kvalitet och visa att statistiken är tillförlitlig krävs detaljerad kunskap om datainsamlingens underliggande processer, det krävs möjligheter att validera data och det krävs modeller för att beskriva osäkerheten i data.

Den avslutande diskussionen fokuserade på hur samarbete mellan alla intressenter kan fortsätta och vidareutvecklas. Det råder ingen brist på bra idéer och engagemang, så förhoppningarna inför framtiden är stora.

INGEGERD JANSSON

### Referenser

- Bakker, B. F. M., (2012). Estimating the validity of administrative variables. *Statistica Neerlandica* 66, 8-17.
- Hedlin, D., Fenton, T., McDonald, J.W., Pont, M. och Wang, S (2006). Estimating the Undercoverage of a Sampling Frame Due to Reporting Delays. *Journal of Official Statistics* 22, 53-70.
- Zhang, L.-C., (2011). A Unit-Error Theory for Register-Based Household Statistics. *Journal of Official Statistics* 27, 415-432.
- Zhang, L.-C., (2012). Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica* 66, 41-63.





Möt Jon Rao, en av världens främsta forskare inom statistik

# EN FANTASTISK RESA

**J**on Rao är en av de forskare inom surveyteorin som betytt mest för hur vi idag arbetar med urvalsundersökningar i praktiken. Han föddes 1937 men är fortfarande intensivt verksam, idag som professor vid Carleton University i Kanada. Vi träffades i Trier i Tyskland på en konferens som handlade om skattning baserat på små urval (Small Area Estimation). Jon var inbjuden föredragshållare. Hans bok med titeln Small Area Estimation kom ut 2003 och blev snabbt standardboken på det området.

Jon läste astronomi i Indien på 50-talet, det var väldigt gammalmodigt framställt på den tiden. Han hade ingen aning om vad statistik var när läraren i algebra sa att statistik är framtiden, det ska du läsa.

Efter studier i matematik som tonåring i Bombay ville han in på en masterkurs i statistik. Utbudet av kurser var mycket mindre än i matematik och därför var konkurrensen så mycket hårdare. Det krävdes absoluta toppresultat och det hade inte Jon. Han knackade på professor Chakrabarty's dörr nästan varje

dag och undrade om det fanns en plats för honom. En dag blev han kallad till professorn. Stående – man satte sig ju inte ner när man talade med en professor – förväntade han sig en utskällning.

## Gick ut med toppresultat

Men professor Chakrabarty meddelade att det nu uppstått en vakans på kursen, ville Jonnagadda ha den platsen? Men, fortsatte professorn, jag har bara elitstudenter och du kommer in på ett sent stadium, hur ska du kunna klara av det? Jon ”worked his butt off” och tog sig igenom de två åren. Som slutexamen hade man sex tentor på sex dagar med frågor konstruerade av bland andra Sukhatme och C.R. Rao (båda enormt skickliga statistiker). Det hårda arbetet lönade sig: han fick toppresultat.

1953 kom boken Sample Survey Methods and Theory av Hansen, Hurwitz och Madow ut. Den inspirerade Jon. Vid

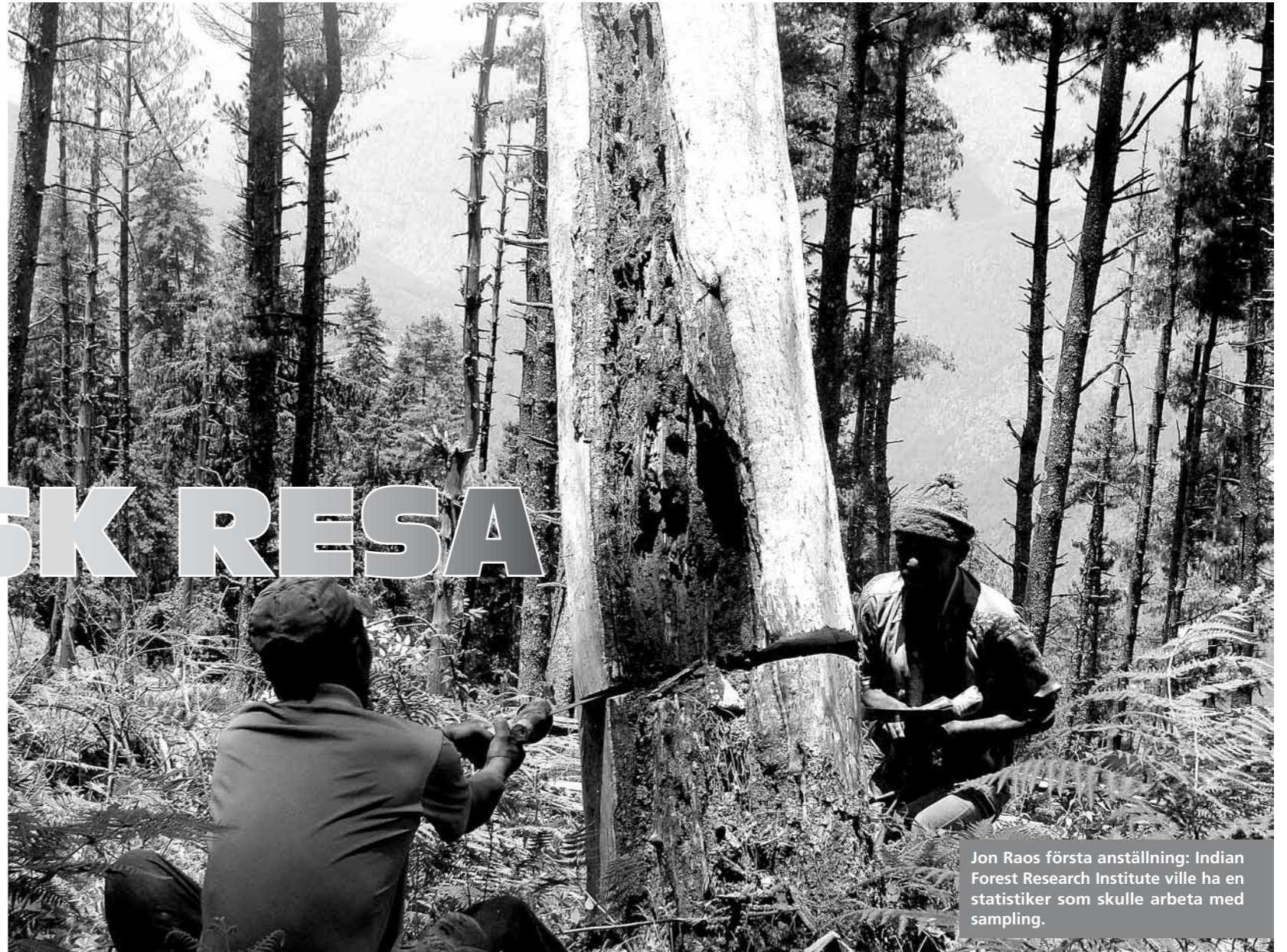
den tiden utvecklades metoder och teori för urvalsundersökningar snabbt.

Jon:  
– Jag ville komma ut i världen och ta en doktorexamen utomlands, men först måste jag få något eget arbete publicerat för att ha en chans.  
– Jag blev erbjuden ett stipendium i experimentdesign men det var sampling jag ville hålla på med. Så jag tackade nej till stipendiet. Istället fick jag en anställning

på Indian Forest Research Institute som ville ha någon som skulle arbeta med sampling. I skogsundersökningar användes då systematiska urval och ett av problemen var att skatta variansen för sådana urval. Jag läste Bertil Matérns böcker. Min handledare på institutet hade haft H.O. Hartley som sin handledare, och det ville jag också ha. Ännu en gång hade jag tur med ett sent återbud: den här gången i Iowa. Så det var två slumpmässiga händelser som fick stor betydelse för mitt fortsatta liv.

## Produktivt samarbete

– Jag påbörjade forskarutbildning i Iowa 1958 med Hartley som handledare. Hartley var min guru, min hjälte som jag såg



Jon Raos första anställning: Indian Forest Research Institute ville ha en statistiker som skulle arbeta med sampling.

**»Jag tror att man ska ha en bred utblick. Det gör det roligare också.»**

**»Man måste ha en hälsosam diskussion om problem.»**

upp till. Vi indier blir lärda att respektera våra lärare. Hartley var alltid inspirerande, vi hade så många intressanta samtal. Idag samtalar inte studenter tillräckligt med sina professorer.

## Fellow i ASA

Samarbetet med Hartley blev synnerligen produktivt. Det skulle föra för långt här att gå in på all den forskning Jon varit med om. Det finns en artikel av David Bellhouse med titeln JNK Rao: an appreciation of his work som hittas enklast med Google. Men den är gammal, den är tryckt 2001. Jon har hunnit med mycket sedan dess. Det finns idag många metoder, teorem eller resultat som heter något med ”rao”, åtskilliga av dem uppkallade efter Jon eller C.R. Rao. Men åter till historien. Jon återvände

till Indien för att ta hand om sin mamma. Det hade han lovat henne. Hans pappa dog när han var sex år. I Indien fick han en av sitt livs överraskningar: ett brev från American Statistical Association där han blev inbjuden som fellow i ASA, vid 25 års ålder. Det blir man normalt bara mycket senare i karriären. Jon blev professor i statistik fyra år senare, bara 29 år gammal.

Han växlade en period mellan Indien, för att ta hand om sin mor, och Nordamerika, där han var på flera universitet. 1973 hamnade han i Ottawa, först med en tjänst som delades mellan Statistics Canada (SCB:s kanadensiska motsvarighet) och universitetet. Det var uselt, han var ”varken här eller där”. Men kontakten

med Statistics Canada har han behållit sedan dess. Han är fortfarande där flera dagar varje månad. ”Jag har så många av mina före detta studenter där, alltid roligt att återses”.

Jon:  
– De flesta universitet ignorerar surveyteorin. Det finns undantag, till exempel JPSM (Joint Program in Survey Methodology, ett amerikanskt konsortium med ett par universitet och en konsultfirma, i nära samarbete med US Census Bureau); JPSM ger en hel utbildning på området.  
– På 90-talet fanns en period då studenterna i statistik på Carleton University var svaga, men nu är tillströmningen



Indian Forest Research Institute.



## »»» MÖT JON RAO

av ambitiösa studenter bra. I Indien går många av de bästa studenterna till IT eftersom de kan få en mycket hög lön med en IT-utbildning i ryggen. Idag verkar inte doktorsexamen uppskattas i statligt arbete lika mycket som förr. Doktorsexamen är inte längre lika attraktiv bland studenter.

### Visst kan man bli rik

**Statistikerycket är ett bra yrke, men rik blir man ju inte, säger jag.**

– Jodå, man kan visst bli rik som statistiker, om det är något som man eftersträvar. Det hjälper till om man flyttar till USA. Där kan man bli rik på konsultarbete, svarar Jon. Eller om man skriver en populär lärobok som används på många universitet.

Jon räknar upp några namn på statistiker som kan förmodas ha tjänat ihop en förmögenhet.

### Vilka råd ger du dagens studenter?

– Det här kanske är gammalmodigt men studenter ska studera bevis och lära sig förstå hur man bygger upp ett bevis. Kunskaperna verkar ha blivit smalare. Jag tror att man ska

ha en bred utblick. Det gör det roligare också. Man ska läsa mycket. Inte bara på internet, man ska läsa riktiga böcker och även tidskrifter som Annals of Statistics.

### Du kan vara en hård kritiker.

– Man måste ha en hälsosam diskussion om problem. Det är bra med kontroverser. En del bayesianer tror att jag är anti-bayes bara för att jag pekar ut problematiska frågor. Kring 1970 hade vi intressanta debatter om grundläggande frågor, till exempel surveyteorins fundament. På konferenser hade

man sessioner med debatter om grunderna men den typen av diskussioner dog ut.

– Surveyteorin ses som tråkig; elementär matematik, inga tekniska utmaningar. Det har aldrig varit sant. Jaroslav Hájek och Bengt Rosén till exempel var väldigt sofistikerade. Small area estimation har gjort att folk med mer allmän statistisk bakgrund har börjat intressera sig för surveyteorin. Saknade data (missing values) är ett annat område som drar skickliga forskare in i surveyteorin. Det är den "amerikanska stilen", säger Jon, man flyttar dit möjligheterna finns.

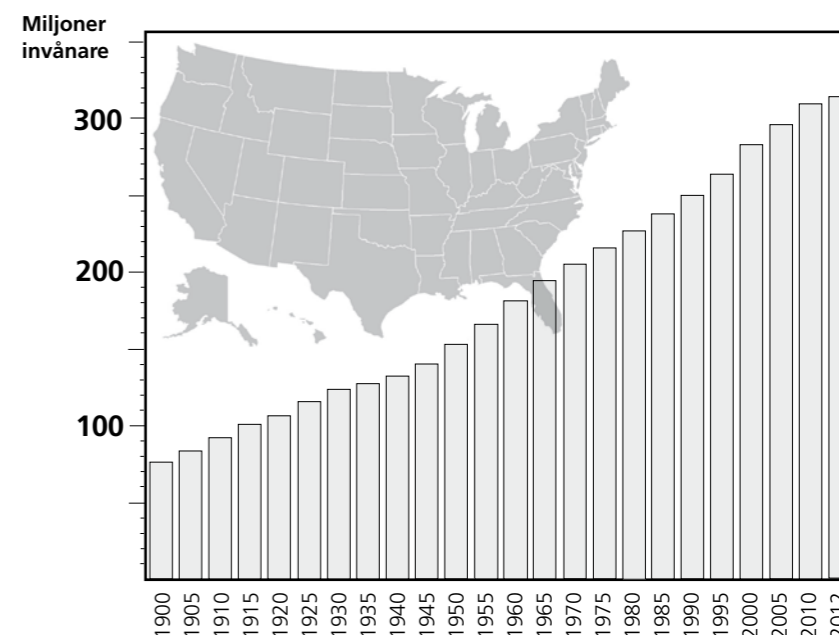
### Framtiden för surveyteorin

– Framtiden för surveyteorin och dess tillämpningar ser mycket bra ut. Den används mer och mer, i allt fler områden. En del säger att surveyteorin inte hänger med den tekniska och den sociala utvecklingen och det kan stämna. Att använda sociala medier för surveyer kräver utveckling av nya teoretiska grunder för nya tillämpningar. Det finns många utmaningar för forskningen inom surveyteorin, bland annat svårigheten med stigande bortfall. Multipel imputering har blivit populärt, "people love it", men där återstår en del svåra problem.

– Andra områden som vi behöver forska mer på är till exempel metodproblem i hälsoundersökningar och surveyer bland människor som är svåra att hitta och att kontakta, säger Jon.

"Data fusion" är något Jon forskar på just nu. Det handlar om att kombinera data från urvalsundersökningar. Det kom ut en artikel i Biometrikas första nummer för i år om data fusion som Jon skrivit tillsammans med Jae Kwang Kim. Jon pratar länge och passionerat om intressanta områden, såväl inom surveyteorin som i allmän statistik i stort.

DAN HEDLIN



# USA:s demografer firar Nerdvana

**Dagens Nyheter** den 16 augusti kunde man läsa att USA:s befolkning två dagar tidigare hade uppnått det imponerande antalet av 314 159 265 personer, vilket är lika med  $\pi \cdot 100$  miljoner. Dagens Nyheter citerar Howard Hogan på US Census Bureau som uppmanar folket att gå ut och fira. Lite googlande avslöjade raskt att händelsen uppmärksammades på flera håll i världen, främst förstås på hemmaplan.

Jag skickade Dagens Nyheter notis till Howard Hogan, som talar lite svenska och vars korrekta titel är Chief Demographer of the US Census Bureau, och passade på att intervjua honom om den märkvärdiga dagen:

**It is obviously a good idea to keep track of mathematical constants in the population figures since it has attracted a lot of attention, even as far away as in Sweden, but how did the idea appear in the first place?**

– We came up with the idea last March 14 (3-14) when someone brought in a pie for "π day." It was simple to project that our population would soon reach 314,159,265 and the demographers began to joke that this was the real day to celebrate.

**Are you looking for other constants too?**

– We had big media events when the US



**Howard Hogan, Chief Demographer of the US Census Bureau**

population reached 200 million in 1967 and 300 million in 2006. I don't think anyone even noticed when our population reached 271,828,182, and 602,214,129 seems a long way off.

**How did you celebrate this particular day at the Census Bureau?**

– With apple pies and cherry pies, of course.

**What has been the reaction among people in general? Will it be a lasting effect of good will or a growing interest in population statistics? Or could it even be the case that some people think that important things like population figures shouldn't be treated in such a frivolous way?**

– We were surprised about how much fun people had with this, with reporters talking about "Nerdvana" and "math nerds rejoice."

Of course, the real "math nerds" criticized us for not being more precise. We should have written population = integer( $\pi 10^8$ ). For me, seeing the clip in Dagens Nyheter was absolutely the best!

### How does your population clock work?

– We started with the most recent census number, which was 308,745,538 as of April 1, 2010. We then used data on births, deaths and migration to bring it forward to our most recent population estimate of 311,591,917, as of July 1, 2011. We then made short term projections and worked out the implied components per second. The current settings are:

- One birth every 8 seconds
- One death every 14 seconds
- One international migrant (net) every 46 seconds
- Net gain of one person every 13 seconds

INGEGERD JANSSON

### Noter:

- Man hittar Censusbyråns Population Clock enklast genom att googla på "population clock census".
- Talet 2,71828182 är "e", basen för den naturliga logaritmen, med åtta decimaler. Talet 6,02214129 · 10<sup>23</sup> är ett skattat värde av Avogadros tal.



## Så blir du 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 Olivia

Ståhl på [sekrfram@gmail.com](mailto: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.



## Har du flyttat?

■ Meddela din nya adress genom att skicka ett e-brev till Olivia Ståhl på [sekrfram@gmail.com](mailto:sekrfram@gmail.com).



## SYNPUNKTER PÅ ARTIKELN OM "DATA-MASKINELL OPTIMERING M. M."

AV TF. PROFESSOR TORE DALENIUS, ENEBYBERG

Jag har med stort nöje och mycken behållning tagit del av artikeln "Datamaskinell optimering, analys och integration vid kliniska laboriemetoder" av herrar Samuelson och Lundberg (denna tidning, 1963, sid. 2728 o. f.). Jag vill lyckönska läsekretsen, som fått förmånen att berika sitt vetande genom ett studium av nämnda opus.

Jag vill också lyckönska författarna, som på ett föredömligt sätt löst den svåra uppgiften att levandegöra en viktig, för att icke säga livsviktig, tillämpning av matematik och ADB utan att ge avkall på kravet på vetenskaplig stringens och kvalitet. Det är kanske förmåtet av mig att taga upp Sv. Läkartidningens värdefulla spaltutrymme med att redovisa vissa matematiskt-medicinska synpunkter, som författarna – av en mera klädsam än befogad anspråkslöshet – underlåt att beröra; därmed har de, som jag strax skall antyda, avstått från en möjlig vidgning av framställningens redan nu betydande vetenskapliga djup. Jag dristar mig likväl att göra några enkla och följdriktiga påpekanden. Eftersom integralen av

kurvan – som så riktigt uttalas på sid. 2728 – "blir ett resultat av ett biologiskt skeende", hade det varit naturligt att utnyttja tekniken med Foole's "Path coefficients" och integrera längs den därav genererade konturen. Som omedelbart inses, får man då direkt

$$B^+ = \int_F Sdt \rightarrow D$$

Denna integral är som bekant alltid konvergent; den är därför ur matematisk-medicinsk synpunkt att föredraga framför den av författarna använda integralen

$$B = \int Sdt$$

vilken ju är divergent för variabler, vilka kan uppvisa språngvist uppträdande diskontinuiteter; den medicinska statistiken erbjuder många vackra exempel härpå. Parentetiskt är det anledning påpeka, att graden av divergens hos B är ett förväntningsriktigt estimat av proportionen patienter (i materialet) i behov av en diagnos (D).

**Alternativt** – och detta ter sig särskilt naturligt i en situation karakteriserad av förekomsten av de ovan nämnda språngvisa diskontinuiteterna – kan man överväga att tillgripa en modifierad Stieltjesintegral. Denna är erfarenhetsmässigt väl lämpad för användning vid konstruktionen av matematiska modeller av fenomen av det slag, som eljest brukar studeras medelst automatisk "retrieval" (sid. 2728).

**Författarnas ovan** nämnda anspråkslöshet har lett till att diskussionen (sid. 2732) av kriterieproblematiken blivit mindre generell än genial. I stället för den triviala och matematiskt-medicinskt omotiverade minimeringen av  $A_n$  i förening med maximering av  $D_n$  – maximin-principen är enligt min mening knappast att rekommendera i detta fall – kunde författarna lämpligen gjort sup  $A_n D_n$  till föremål för en gränsövergång. Som lätt inses (infinitesimalkalkylens fundamentalsats) ger nämligen lim sup  $A_n D_n$  entydigt det av författarna redovisade resultatet; men detta sätt att bestämma optimum har den fördelen, att något särskilt existensbevis icke tarvas eller ens är möjligt.

Även här är en alternativ ansats icke endast tänkbar utan rent av fördelaktig. Operationen sup  $A_n D_n$  kan (för ändliga, diskreta värden i vivisektionen av  $A_n$  och  $D_n$ , dvs.  $A_n \vee D_n$ ) lätt verkställas medelst simulering i en (binärt arbetande) datamaskin, förutsatt att denna arbetar med likström. Denna förutsättning är – såvitt jag har mig bekant – uppfyllt i de anläggningar, som användes i vårt land för avancerad medicinsk databehandling.<sup>1</sup>

**Till slut vill jag** utnyttja tillfället att rikta kritik mot författarna för deras styvfaderliga behandling av databerikningsmetoden till förmån för den medicinska integralkalkylen.

Databerikningsmetoden – som framför allt utvecklats av den i USA verksamme italiensk-födde forskaren Gigo, vars namn numer är förevigat för lång tid i computerterminen "gigo" = "garbage in, garbage out" – har dock mycket stora fördelar framför konventionella metoder, generellt sett. För empirisk forskning i medicinsk databehandlingsmiljö är databerikningsmetoden – med utnyttjande av Gigo's selektiva, självkorrigering MD-program med "feed back and forth" – så lovande, att det måste betecknas som en skönhetsfläck, svår att bortoperera, att den icke ens består ett omnämnande.

### Fotnot:

<sup>1</sup> Som lätt inses är det vissa risker förenade med användningen av växelström. Som en konsekvens av s. k. synkroperiodisering skulle en godtycklig binär sifferserie i princip kunna transformeras till idel nollor. Om detta inträffar vid medicinsk databehandling, skulle resultatet kunna bli en snedvridning av diagnosen (D), även om detta dock icke är särskilt sannolikt eller ovanligt.

*Vi tackar Eskil Dalenius för hjälp med artikeln och tillstånd att återtrycka den. Artikeln publicerades ursprungligen i Svenska Läkartidningen 1964: 61: 1124 (nr 14).*

Red

## TIPS!

Passa på och gå på Frimis-seminariet i Linköping den 10 oktober om du har tid och chans att ta dig dit med kort varsel. Temat för heldagsseminariet är:

**Vilka trender kan vi idag se när det gäller undersökningar på webben?**



## ORDFÖRÄNDEN HAR ORDET

# Konsten att dra slumpmässiga urval

**E**n av de största framstegen någonsin i officiell statistik är tekniken att dra slumpmässiga urval. Egentligen är det märkligt att man utifrån ett litet urval kan dra slutsatser om en stor population. Ännu märkligare är det att man drar slutsatser om tillförlitligheten i slutsatserna om populationen.

Det är enormt mycket billigare att hämta in uppgifter från ett urval jämfört med att göra det från en hel population. Men det är ändå dyrt. Vi kan inte heller begära att personer, företag, eller vilka objekten i undersökningen kan vara, ska ta tid till att lämna uppgifter om det inte är nödvändigt. I västvärlden blir det år för år dyrare och mer tidskrävande att få in uppgifter. Bortfallet ökar. Det stör slutsatsernas tillförlitlighet.

Därför är stickprovsteori och praktiskt genomförande av undersökningar ett ämne som hela tiden diskuteras bland surveystatistiker. Hur ska vi göra undersökningar billigare? Hur ska vi underlätta för uppgiftslämnarna? Hur ska bortfallet minskas eller göras mindre snedvridande? Är mindre resurskrävande sätt att genomföra surveyundersökningar, som webbpaneler, tillförlitliga?

**Surveyföreningen arrangerade** i början av juni ett seminarium under rubriken Stickprov i teori och praktik. Ett drygt trettital personer slöt upp för att lyssna till de tre föredragshållarna Jan Selén, SCB, Anders Haraldsson, Sinitor, och Michael Carlson, då SCB men numera Stockholms universitet.

Arbetskraftsundersökningen (AKU) fick 2010

mer medel för att öka urvalet med hela 37 procent. Huvudsyftet var, berättade Jan Selén, att med ett tilläggsurval förbättra redovisningen av utsatta grupper. Men medför ett större urval med nödvändighet en bättre precision av skattningarna? Jan förklarade pedagogiskt vilka faktorer som är viktiga att beakta vid sammansatt estimation.

**Urvalsförfarandet** vid webbpanelundersökningar var temat för Anders Haraldssons presentation. Anders tog upp de uppenbara problem som uppstår vid självselektering och pekade på övervikten av "teknikintresserade" samtidigt som det råder brist på män, unga och låginkomsttagare i självselektade webbpaneler. Frågan om kvoturval versus slumpmässigt urval togs upp och diskuterades i sammanhanget. Avslutningsvis pekade Anders på vikten av att engagera respondenterna, att möta dem där de finns (sociala medier) och att ge dem något i utbyte för sitt engagemang. Webbpaneler har varit uppe förut i Qvintensen vid olika tillfällen, senast i nr 2/11.

**Michael Carlson avslutade** dagen med att diskutera och problematisera tänkandet som ligger bakom, eller före, själva urvalet. Hur ser vi på modell och population? Är vi inriktade på designbaserad eller modellbaserad inferens? Vill vi presentera en ögonblicksbild av en population eller vill vi ta fram en förklaringsmodell genom att belysa sambanden mellan olika fenomen?

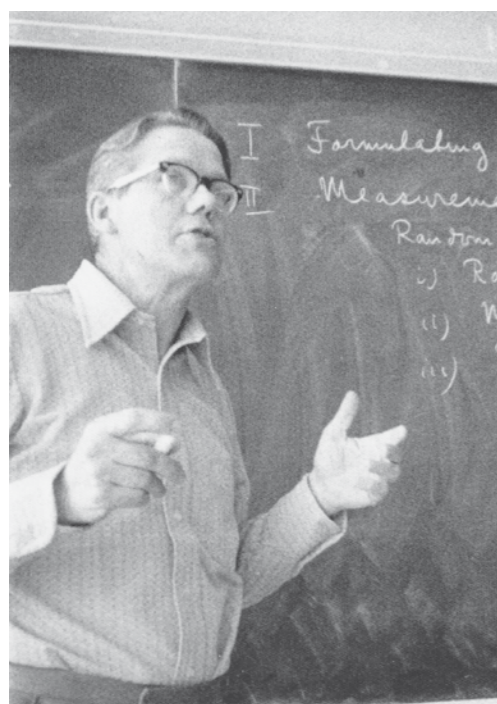
En annan fråga som vi surveystatistiker ofta diskuterar är om man skulle strunta helt i att

samla in data specifikt för en viss undersökning? Istället skulle man förlita sig på de data som redan finns, framför allt i administrativa register. SCB har nyligen ordnat en workshop i ämnet och på sidan 11 i det här numret av Qvintensen kan du hitta några tankar om vad som diskuterades där.

**Till sist vill jag** passa på att flagga för Frimis-seminariet i Linköping den 10 oktober om du har chans att ta dig dit med kort varsel. Temat för heldagsseminariet är Vilka trender kan vi idag se när det gäller undersökningar på webben? Från Novus kommer Torbjörn Sjöström och Mikael Ohlsson för att berätta om studier som gjorts för att jämföra resultat av webbpanelundersökningar och telefonundersökningar. Jörgen Svensson från SCB talar om metodproblem som uppstår vid användandet av webbpaneler. Därtill presenteras en remissversion av Surveyföreningens webbpanelkommittéarbete där fokus legat på att hitta lämpliga mått för kvalitetsbedömning av webbpaneler och webbpanelundersökningar. Henrik Kronberg, Åke Wissing och Jan Wretman från webbpanelkommittén håller i presentationen.



JOAKIM MALMDIN  
ordfsurveyat@gmail.com



Tore Dalenius undervisar.



FMS

## ORDFÖRANDE HAR ORDET

Svensk läkemedels-  
industris framtid

**N**är jag skriver denna spalt håller sommaren fortfarande på att försöka bryta igenom de molnskyar som inte vill ge sig av och de regnskurar som inte inser att de kan komma nattetid.

När ni läser detta har hösten kommit och jag hoppas på att den är så skön som den faktiskt kan vara med frisk, hög luft och vackra höstlöv.

Som FMS ordförande har jag den senaste tiden funderat på den turbulenta tid som råder för statistiker inom läkemedelsindustrin i Sverige idag. Jag har själv aldrig jobbat inom läkemedelsindustrin. Jag har ändå på visst avstånd följt utvecklingen inom detta statistiska arbetsfält då jag har en man och flera före detta kursare som jobbar eller har jobbat inom Astra och sedan AstraZeneca. Vad jag förstått på deras berättelser har denna arbetsplats ändrats mycket genom åren.

Under hösten 2011 och våren 2012 har jag också varit lite mer uppmärksam än vanligt på vilka platsannonser som distribueras till Svenska statistikfrämjandets medlemmar via dess e-postlista. Det har varit en ansevärd mängd annonser som till en stor del har kommit från olika statliga myndigheter. Förutom statliga myndigheter har nog de flesta andra jobb erbjudits inom akademien i form av lektorat eller tidsbegränsade postdoktorstjänster. Det känns som arbetsmarknaden för statistiker är god, men att mixen av inriktningar håller på att ändras.

Det finns flera funderingar som dyker upp i mitt huvud kring denna utveckling. Kommer detta att påverka oss statistiker som grupp i Sverige och i så fall hur? Hur kommer läkemedelsindustrin, i form av en av de stora arbetsgivarna för statistiker, att förändras? Kommer den statistiska metodforskningen som har inriktning mot kliniska prövningar och RCT (randomized controlled trials) att minska och förändras?

Det finns flera funderingar som dyker upp i mitt huvud kring denna utveckling. Kommer detta att påverka oss statistiker som grupp i Sverige och i så fall hur? Hur kommer läkemedelsindustrin, i form av en av de stora arbetsgivarna för statistiker, att förändras? Kommer den statistiska metodforskningen som har inriktning mot kliniska prövningar och RCT (randomized controlled trials) att minska och förändras?



Anna Ekman,  
ordförande i FMS.

Som den moderna människa jag är så tänkte jag lite på dessa frågor, men satte mig sedan vid min dator och googlade. Jag fick ganska snabbt upp två intressanta artiklar i ämnet "svenska läkemedelsindustrins framtid".

Båda artiklarna var publicerade i Biotech Sweden, biotech.idg.se, under februari och maj i år. Båda artiklarna understryker vikten av nya samarbetslösningar mellan privata och akademiska aktörer för att Sverige ska kunna behålla en god läkemedels- och biotechutveckling. Båda

framhåller vikten av att sjukvården och regeringen beställer forskning och ger andra incitament till forskning. Nya mindre läkemedelsbolag med gott samarbete med sjukvården och akademien anses vara framtiden.

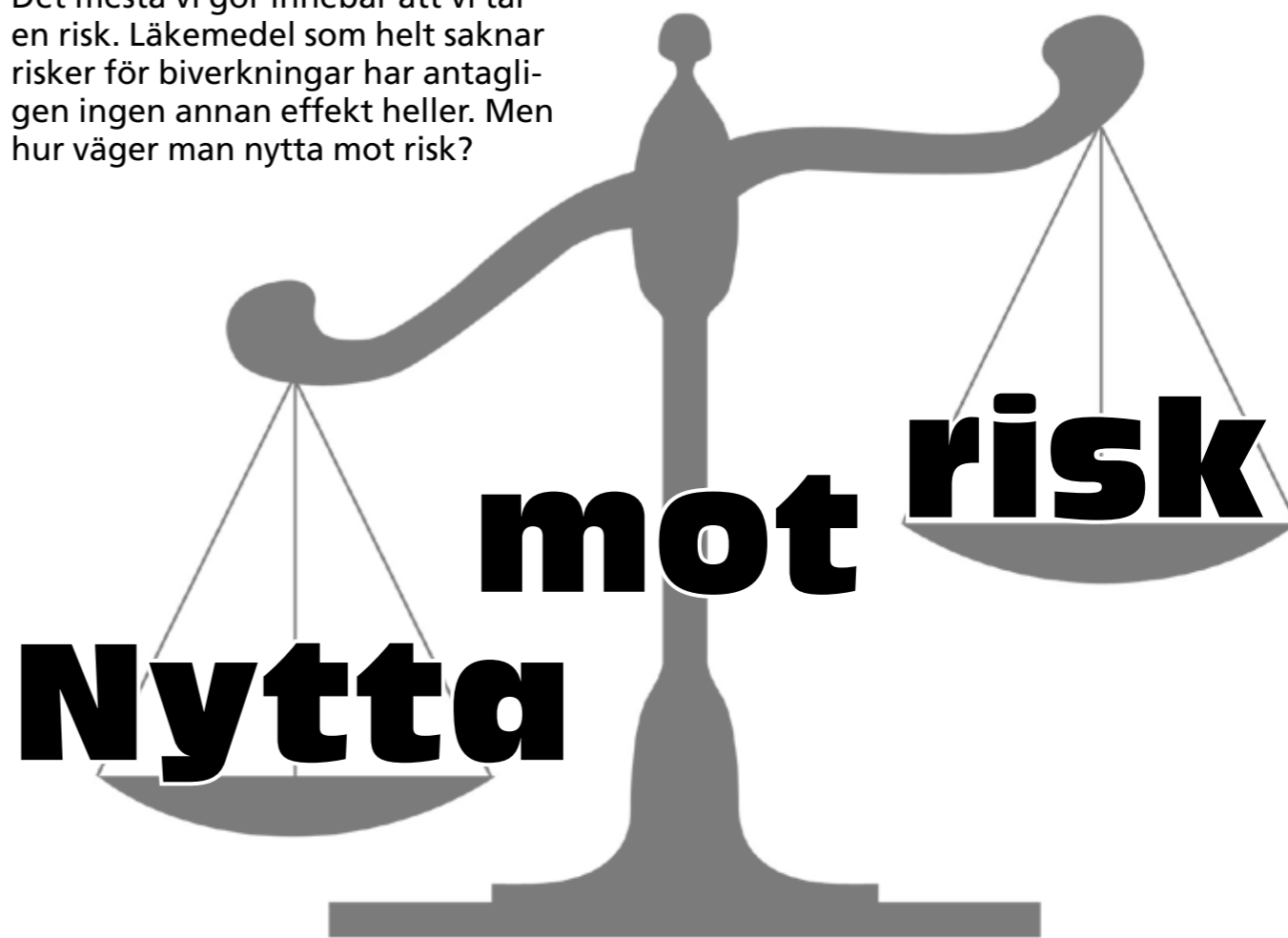
Vad detta betyder för statistiker inom dessa branscher är svårt att förutsäga. Jag kan se flera möjliga konsekvenser. Till exempel att den samlade kompetensen, som grupper av läkemedelsstatistiker utgör på företagen idag, till stor del splittras och sprids ut och därmed går synergieffekter till spillo. "Läkemedelsstatistiken" kan till stor del komma att utföras av akademiskt anställda statistiker. Detta är generellt sett sämre betalda statistiker, vilket skulle innebära en vinst för läkemedelsbolagen. Det skulle i och för sig istället kunna leda till att de akademiska statistikerna får högre löner om deras arbetsuppgifter inkluderar läkemedelsprövningar.

Ett bättre samarbete mellan industri, akademi och sjukvård kan ge statistiker ett bredare arbetsfält och skulle kunna motverka att man som statistiker blir alltför nischad. Andra sidan på myntet skulle kunna vara att statistiker i allmänhet får mindre expertkunskap och tvingas till att vara breda i sina statistikfärdigheter.

Jag slutar mina spekulationer här, men hoppas kunna diskutera dessa frågor vidare under hösten på FMS jubileumsmöte. Då FMS i år fyller 25 år hoppas jag att många av er kan delta på vårt jubileumsmöte i oktober.

ANNA EKMAN

Det mesta vi gör innebär att vi tar en risk. Läkemedel som helt saknar risker för biverkningar har antagligen ingen annan effekt heller. Men hur väger man nytta mot risk?



**E**n aning stressad, mitt i den underbara skolavslutnings- och studentskivettiden kastade jag mig iväg på utflykt söderut. Resans mål var en workshop på Universitetssjukhuset i Malmö. FMS styrelse hade sedan över ett halvt år tillbaka arbetat med förberedelserna och spänd av förväntan klev jag in i lokalen för att avnjuta resultatet.

Workshopens tema var värdering av nytta mot risk och mötet var ett samarrangemang mellan FMS, vår danska motsvarighet DSBS samt EFSPi (European Federation of Statisticians in the Pharmaceutical Industry). Mötet hade lockat 48 deltagare från ett stort antal läkemedelsföretag och även myndigheter, institut och konsultföretag. De flesta deltagarna kom från Danmark och Sverige men även övriga Europa var representerade.

Läkemedel som har effekt har i allmänhet också risk för bieffekter. Nyttan ska klart överväga riskerna för att läkemedlet ska godkännas. Ett ökat användande av kvantitativa metoder för nytta mot riskanalyser vid utvärdering av nya medicinska produkter

har föreslagits som en hjälp för läkemedelsverk och vårdbetalare över hela världen. Syftet med denna workshop var att få en lägesrapport och erbjuda ett forum för diskussion av den senaste utvecklingen i tillämpningen av dessa metoder med inbjudna talare som representerade europeiska myndigheter, industrin och akademien.

Larry Phillips, London School of Economics, blev som nybliven ingenjör intresserad av samspelet mellan människa och maskiner och därför började han studera beteendevetenskap. Han är numera väldigt insatt i läkemedelsindustrins behov av produktutvärdering och granskning. Larry tycker att reglerare på myndigheter borde övergå från implicita till mer explicita metoder och från kvalitativa till kvantitativa metoder. Han introducerade oss i det nuvarande läget inom utvärdering av kvantitativa metoder för utvärdering av balansen mellan nytta och risk för medicinska produkter. Myndigheterna använder i regel inte kvantitativa metoder. Den europeiska läkemedelsmyndigheten genomförde ett utredningsprojekt



Lawrence (Larry) Phillips talade om kvantitativa metoder för utvärdering.

2009-2011 för att utveckla och testa olika tillvägagångssätt och metoder. Ett antal europeiska myndigheter fick svara på frågor. Bland annat fick de frågan: Vad är nytta och vad är risk? En sammanfattande definition kan sägas vara fördelaktiga och ofördelaktiga effekter – alla mer eller mindre osäkra.

Två kvalitativa fackverk utgör riktlinjer för analys av balans mellan nytta och risk. I det ena föreslås stegvisa arbetsätt för den grupp som ska göra utvärderingen. Beslutskontexten ska först definieras, utfallsvariablerna ska väljas, källdata ska identifieras och allt ska skraddarsys för den specifika utredningen. Det ska sedan göras en uppskattning av hur viktiga olika utfall är och tillslut ska de olika måtten för nytta och risk redovisas i tabeller och grafer för att sedan tolkas.

I det andra fackverket ingår följande steg som guide för modellering: beskriv problemställning, mål (och kriterier) som kan illustreras i träd, behandlingsalternativ (t.ex. placebo, alternativa läkemedel, olika doser), beskriv konsekvenser (t.ex. omräkning till preferensvärden - linjärt eller icke-linjärt), diskutera avvägningar (bestäm vikter för när effekter pendlar och väger över till något som är kliniskt relevant – fördelaktigt eller ofördelaktigt), osäkerhet (utför känslighetsana-



## »»» NYTTA MOT RISK

lys genom att t.ex. variera vikten för ett kriterium), utred riskbenägenhet och ytterligare konsekvenser som kan uppstå till följd av ett beslut.

**Nyckelkraven för en sammanhängande utvärdering av nytta mot risk är omfattande. Bara ett fåtal kvantitativa metoder håller måttet. Ytterst få läkemedelsföretag använder dessa metoder och hittills har inga myndigheter accepterat dess användning även om metoderna används brett utanför medicinssektorn. För beslutsanalys med flera kriterier är "Multi-Criteria Decision Analysis" (MCDA) en av de mest använda metoderna. MCDA är en utveckling av beslutsteorin som innefattar beslut**

**»Modeller skrämmer myndigheter, för vem ska egentligen ha makten att sätta värden på ett medicinskt beslut?»**

med många delfrågor och motstridiga kriterier som ska kombineras till en samlad utvärdering. Larry menade att MCDA i kombination med den sociala process där en eller flera arbetsmöten med experter ingår är nyckeln till smarta beslut.

Vi fick också ta del av vad "Phillips lag" innebär – "Lita aldrig enbart på en expert".

Larry ställde en fråga till sig själv och publiken: Varför har den medicinska professionen varit så tveksam till att tillämpa balanseringsmetoder för nytta mot risk? Han själv föreslog fem olika anledningar:

- 1) Läkare föredrar att arbeta med ord istället för siffror och modeller
- 2) Nyttan mot risk vägs olika i olika kulturer
- 3) Det finns en förkärlek för beslut baserade på beprövad erfarenhet
- 4) Modeller skrämmer myndigheter, för vem ska egentligen ha makten att sätta värden på ett medicinskt beslut?
- 5) Det finns en stor tillit till den medicinska bedömningen

Larrys förhoppning är att medicinsk praxis och myndigheters tillsyn i framtiden bättre utnyttjar möjligheten att kombinera moderna statistiska metoder med beslutsteori.

**Andrew Thomson**, statistiker på det brittiska läkemedelsverket, gav sin syn på kvantitativa metoder för beslut om medicinska produkter har en positiv

nytta vägt mot risk eller inte. Besluten som fattas i dessa sammanhang är verkligen betydelsefulla för många människor. Viktiga frågor som bör beaktas:

- Vem konstruerar modellerna? Industri eller myndighet?
- Vem ska välja lämpliga vikter och hur och när är de valda?
- Kan vi definiera vikter a priori? För positiva effekter går det bra men svårare är det att i förväg definiera vikter för biverkningar.

I det här sammanhanget saknar modeller utan vikter vetenskaplig grund samtidigt som modeller med vikter kan anses alltför subjektiva. Det är ju tyvärr möjligt att välja vikter så att du får den önskvärda balansen. Det viktigaste är att modellen får utgöra ett diskussionsunderlag och att de inblandade experterna får redogöra för hur de har tänkt.

Ytterligare övervägande måste göras efter det att ett positivt beslut fattats och läkemedlet är godkänt och registrerat eftersom data fortsätter att strömma in från nya experiment men även från observationsdata från olika sjukvårdsregister. Att bygga in denna information innebär ännu fler utmaningar.

**Johan Bring från Statisticon** har för AstraZenecas räkning bidragit till en utredning med syfte att förbättra övervakning av läkemedels nytta mot risk inom Europa. Johan har jämfört olika metoder för att bedöma nytta mot risk.

En bra beslutsmodell kan hjälpa oss i den komplicerade process där vi ska bestämma om ett nytt läkemedel ska godkännas eller inte. Det skulle vara bra att även ha en strukturerad process som hanterar all relevant information för att kunna dra tillbaka en produkt från marknaden. Det finns många olika metoder. Dessa kan delas in i tre klasser: 1) rent deskriptiva mått som "Benefit-Risk Ratio" 2) deskriptiva och delvis normativa som "Stochastic Multicriteria Acceptability Analysis" 3) deskriptiva och normativa som "Multi-Criteria Decision Analysis".

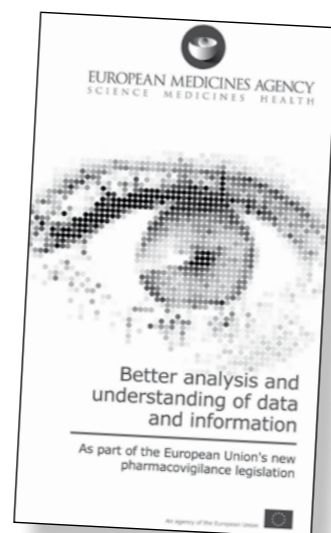
Johan diskuterade de utmärkande egenskaperna för de olika modellerna och belyste de huvudsakliga svårigheterna med respektive metod samt visade ett antal varianter av illustrativa grafer som är lämpliga som diskussionsunderlag. Beslut bör fattas på basis av en

kombination av deskriptiva fakta och en värderad bedömning.

Johan framförde en filosofisk fundering. Är det kanske trots allt bättre att göra en bedömning utifrån ett holistiskt synsätt istället för att splittra allt i små bitar och värdera nyttor och sannolikheter för alla separata delar? Denna viktiga fråga diskuterades men blev sedan hängande i luften när det var dags för en god lunch med mingel och trevligt umgänge.

Alla har sina personliga åsikter om ett läkemedel. Helst ska både myndigheter som Food and Drug Administration (FDA) i USA och den europeiska läkemedelsmyndigheten (EMA, European Medicines Agency) och sjukvårdsbetalare i olika länder samt naturligtvis även patienterna vara överens om hur bra ett nytt läkemedel är.

Nytta mot riskanalyser av behandlingar drar enorma resurser och resultatet verkar ofta bero på personliga erfarenheter och vilken expertis som blandas in. Sinan Sarac, "Medical officer", från det svenska läkemedelsverket demonstrerade hur kliniska data kan viktas, analyseras och presenteras vid användning av en datadriven analys som är enkel att tillämpa och som möjliggör direktjämförelse av olika risker och nytta, kvantifierar klinisk signifikans och som är anpassad för att jämföra olika alternativ, behandlingar etc. i en stegvis



**Den europeiska läkemedelsmyndigheten (EMA, European Medicines Agency) ska bli bättre på att koordinera läkemedelsinformationen i Europa.**

process liknande det som beskrivits tidigare under dagen. Han demonstrerade metodologin med en analys av en kohort med drygt 300 patienter med tarmcancer som behandlats på Köpenhamns Universitetssjukhus. Sinans slutord och rekommendationer:

- Genomsynlighet i beslutsfattandet ökar trovärdigheten av analysen och kan säkras genom att följa en specificerad struktur, motivera de val som gjorts i de kritiska analysstegen och genom att vara konsekvent mot tidigare beslut.
- Diskussion om klinisk signifikans är viktig och stöder beslutsfattandet i ett större perspektiv.
- Använd visualiseringsverktyg för att samtidigt kunna beskriva så mycket data som möjligt.

**Talaren Doris Stenver** var även hon från det svenska läkemedelsverket. Den europeiska läkemedelsmyndigheten och de nationella tillsynsmyndigheterna inom Europa arbetar tillsammans i ett nätverk för godkännande av nya läkemedel och ständigt pågående säkerhetsövervakning av redan registrerade läkemedel (s.k. "pharmacovigilance") där upprepade analyser av nytta mot risk utförs. Doris beskrev hur nätverket arbetar och vad de har för uppgift och ansvar. De viktigaste verktygen som finns idag är s.k. "Periodic Safety Update Reports" och riskhanteringsplaner.

Hon redogjorde även för de huvudsakliga målsättningarna med den nya europeiska lagstiftade läkemedelsövervakning som gäller från juli 2012. Precis som tidigare ingår upptäckt av signaler, analys och minimering av risker i arbetet med säkerhetsövervakningen. Det som är nytt är att myndigheterna ska bli bättre på att kommunicera risker via den medicinska litteraturen, massmedia och webbplatser vilket är nödvändigt i den internet-era som råder och i takt med att patienterna själva vill hålla sig informerade utan att gå via sjukvårdspersonal.

En internationalisering och harmonisering kommer att ske och be-

slutsfattande ska baseras på internationell erfarenhet. EU-kommittén för säkerhetsövervakning av läkemedel består av medlemmar från varje medlemsland samt representanter från patientorganisationer och sjukvård. Kommitténs agendor och protokoll från möten, analyser, beslut och rekommendationer ska vara tillgängliga för allmänheten via webbportaler, offentliga möten (s.k. hearings) och en publicerad lista över produkter som är föremål för extra övervakning.

Myndigheternas tidigare brist på koordinering inom Europa, då information inte kommit ut samtidigt överallt, har resulterat i allmänhetens misstankar om att information döljs i vissa länder. Detta ska nu åtgärdas genom att den europeiska läkemedelsmyndigheten ska koordinera kommunikationen.

**George Quartey, Genentech**, höll ett föredrag som var en riktig energiinjektion.

Nytta mot riskanalys är en viktig del av läkemedelsutvecklingen med syfte att stärka beslutsfattandet för folkhälsans bästa. På Genentech/Roche fanns ett växande intresse för att utveckla ett strukturerat system för att utvärdera nytta mot riskprofilen

av produkterna med hjälp av kvantitativa modeller i ett enkelt format. De ville integrera effekt och biverkningar i resultatrapporten istället för att visa dessa i separata avsnitt.

George visade hur han och kollegorna har utvecklat och implementerat ett arbetssätt, författat dokument med riktlinjer och metodbeskrivningar, och byggt ett analysverktyg.

Publiken blev mycket imponerad av att höra Georges berättelse om hur han fick företagets beslutsfattare och den högsta ledningen att börja använda detta strukturerade verktyg för utvärdering av nytta mot risk. Om jag förstod det hela rätt filmades ledningen och filmen användes sedan i utbildnings-syfte i övriga delar av organisationen.

Allt som behövdes var en stilig, lång,

**Varför har den medicinska professionen varit så tveksam till att tillämpa balanseringsmetoder för nytta mot risk, undrade professor Larry Phillips.**

svart statistiker med stark övertygelse som han själv uttryckte det, och syftade alltså på sig själv.

Företaget har nu börjat bifoga kvantitativ analys med bakomliggande resonemang och överväganden i den dokumentation som skickas till myndigheter, nu senast till den amerikanska läkemedelsmyndigheten FDA, inför registreringsbeslut. Larry Phillips kommenterade detta: "Väldigt modigt och lycka till".

Stämningen under den efterföljande fikarasten var uppsluppen och kaffet gott.

**När vi återsamlades** tog Ulf Persson från svenska Institutet för Hälso- och Sjukvårdsekonomi till orda.

Utvärdering av kostnad mot nytta har använts inom den svenska väg- och trafiksektorn sedan 1970-talet. Inom sjukvårdssektorn har sådan utvärdering inte använts av någon myndighet förrän år 2002 då det s.k. "Value Based Pricing"-systemet togs i bruk. »»»

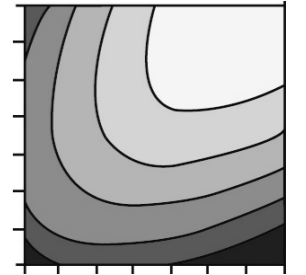




FOTO: JONAS ARDEFORS

Inom transport används analyser av kostnad mot nytta. Prioritering mellan olika insatser är baserat på diverse trafiksäkerhetsmått.

## industriell statistik



### »»» NYTTA MOT RISK

Inom transport används analyser av kostnad mot nytta för beslut om hur resurser bäst ska användas till byggande av nya säkrare vägar. Prioritering mellan olika insatser är baserat på diverse trafiksäkerhetsmått. I sjukvårdssektorn används analyser av kostnad mot nytta mest för att utvärdera behandlingar men också för att utveckla riktlinjer för behandling.

Värdering av nyttan för trafiksäkerhet, d.v.s. värdet av riskreduktion i sig, värdet av ett statistiskt liv (VSL), har många likheter med värderingen av hälsa som oftast uttrycks i värdet av ett kvalitetsjusterat levnadsår (QALY). Faktum är att QALY är baserat på VSL och båda hänger samman med individens villighet att betala för säkerhet och hälsa i Sverige.

#### »...magkänsla och modeller måste gå hand i hand»

Ett räkneexempel: Om den initiala risken (p) är 10/100 000 och vi vill reducera risken med 50% till 5/100 000 ( $\Delta p$ ). Hur mycket är var och en beredd att betala för denna riskreduktion? Det genomsnittliga WTP ("Willingness To Pay") är 1 200 kr per individ i detta fall. Värdet av ett statistiskt liv, VSL, kan då beräknas med formeln  $WTP/\Delta p$ , d.v.s.  $1200/(5/100\,000) = 24$  miljoner kr. Värdet av en QALY beräknas som  $VSL/\text{antal förlorade kvalitetsjusterade levnadsår}$ .

**Dagen avslutades** med en paneldiskussion som leddes av Carl-Fredrik Burman under rubriken: Kvantitativa

modeller – leder dessa till att vi fattar bättre beslut?

Vi förvånades över att panelen var så pass överens om att kvantitativa metoder är önskvärda.

Andrew Thomson tyckte dock att det skulle vara svårt för myndigheter att specificera vikter. Vi kom tillbaka till Johan Brings fråga om helhets- kontra atomistiskt perspektiv, som även kan syfta på andra tillämpningsområden (borde man t.ex. göra formella sannolikhetsberäkningar för olika detaljer i brottmåll?). Även om det hängde i luften så ville ingen i panelen riktigt säga att vi ska gå på magkänsla istället för (bara) siffror.

Larry Phillips menade att beslutsanalys är vida använt, och att forskningen pekar mot att formaliserade beslutsprocesser är bra.

**Vidare diskuterades varför** regulatorisk prövning (av läkemedelseffekt och säkerhet) och förhandling med vårdbetalare om prissättning (kostnad mot nytta) görs separat (och ofta sektoriellt). Andrew nämnde att det finns visst samarbete mellan den engelska sjukvårdsförsäkringsmyndigheten och den engelska läkemedelsmyndigheten. Ulf påminde om att betalningsviljan är olika i olika länder (p.g.a. olika välbefinnanden) och att det komplicerar processen.

Det framkom dock inte några starka skäl mot att olika utvärderare skulle kunna titta parallellt på data och (delvis) använda samma nytta mot riskmodeller. Med ett sådant förfarande skulle sedan betalare i olika länder kunna plugga

in sina värderingar (t.ex. värdet för en QALY).

**Det var en** mycket givande dag som satte igång många funderingar. Den slutsats som jag hade kommit fram till när jag satte mig på tåget och åkte norrut på det krängande X2000-tåget var att magkänsla och modeller måste gå hand i hand när det gäller nytta mot risk och nytta mot kostnad.

Om inte modellerna beskriver det vi känner och leder till det utfall som känns rimligt, ja då justeras antagligen vikterna och modellen anpassas till den uppkomna situationen. Det låter kanske oprofessionellt och ovetenskapligt. Jag tror ändå att alla är medvetna om att det är oerhört svårt att i förväg specificera modeller kring risker och stå fast vid dessa innan vi vet vilka ofördelaktiga effekter som finns och i vilken omfattning.

Att jag sedan har en egen ofördelaktig känsla i magen av illamående när jag åker dessa hemskt krängande tåg har väl egentligen inte med saken att göra. Eller? Tyvärr måste jag konstatera att åksjuepiller gör mig trött så det är nu definitivt dags att avsluta skrivandet.

Jag vill avsluta med att tacka organisationskommittén Carl-Fredrik Burman (EFSPi+FMS), Sören Lohaven (DSBS), James Matcham (EFSPi), Hanna Svensson och Linda Werner Hartman (båda FMS) för ett väl utfört uppdrag och lyckat arrangemang.

ANNA TORRÅNG  
ASTRAZENECA

### ORDFÖRANDEN HAR ORDET

## Tema för höstens aktiviteter: Att vårda en process, mäta, räkna, rita – och övertyga

**D**en här gången ägnas större delen av min spalt åt höstens aktiviteter i vår förening och jag tänkte börja med vår serie webinarier.

Vi har nu under ett års tid träffats virtuellt en gång i månaden och lyssnat på ett föredrag. Därmed lyckas vi undvika ett av de stora problem som föreningar av vår typ har, nämligen att medlemmarna är spridda över hela landet och att vi egentligen inte någonstans är tillräckligt stora på det lokala planet för att bedriva någon lokal verksamhet. Att samlas rent fysiskt är besvärligt och tidsödande och sker i stort sett bara två gånger om året, vid höst- och vårseminarierna. Webinarier råder till viss del bot på detta och har varit mycket lyckade. Ingemar Sjöström delade med sig av våra erfarenheter i nr 1/12 av Qvintensen.

**När ni läser** detta har höstens bägge första webinarier, som handlade om mätosäkerhet respektive robust konstruktion, redan gått av stapeln.

Men det kommer fler och ni hittar mer infor-

FOTO: LINA LANDE  
Göran Lande.

mation på vår webbplats, <http://www.indstat.se>. Vi har ett webinarium i månaden och vad de kommer att handla om i fortsättningen låter jag vara osagt. Du kanske har något eget önskemål eller till och med något du vill presentera? Hör då gärna av dig till mig.

Höstens stora aktivitet är vårt höstseminarium. Det kommer att äga rum i Göteborg den 7 november, dagen innan arrangemanget World Quality Day äger rum i samma stad.

Temat är "Att vårda en process, mäta, räkna, rita – och övertyga". Det kommer att bli fyra föredrag inom detta breda område som omfattar allt från processen att generera data till att presentera analysresultaten på ett övertygande sätt, exempelvis med väl valda grafer.

Många statistikers favoritprocess, Markovprocessen, kommer att beröras i form av ärendehantering och hur länge ärenden befinner sig i de olika tillstånden i ärendehanteringsprocessen. Jag ser fram emot att träffa er i Göteborg!

**Förutom våra höstaktiviteter** skulle jag också vilja reflektera kring vår medlemsskara

som glädjande nog har vuxit sista året. En titt på medlemslistan visar att de nyrekryterade medlemmarna nästan uteslutande kommer från industrin. Det är både glädjande och alarmerande.

Det är fantastiskt roligt att ha så många medlemmar från Volvo, Tetra Pak, Siemens och Sony Mobile Communications. Man slås också av hur omfattande och hur avancerat arbetet med statistiska metoder är bland dessa nya medlemmar.

Men medlemmarna från universiteten är ganska få. Jag vet faktiskt inte varför. Föreningen skulle för dem kunna vara en kanal för att fånga in statistiska problem från industrin, problem som kan användas i allt från problemsamlingar på den mest elementära nivån till avancerad forskning. Framför allt verksamma vid de tekniska högskolorna borde ha ett stort utbyte av detta. Föreningen skulle kunna bygga nätverk och broar mellan industri och akademi.

Vi i Föreningen industriell statistik får helt enkelt försöka vidta åtgärder för att väcka intresse för vår förening hos akademiker på universitet och högskolor. Har just du något tips på hur vi ska bära oss åt får du gärna kontakta undertecknad.

GÖRAN LANDE  
goranlande at yahoo.se





Utö i oktober ...



...och Australien i februari.

## IN MEMORIAM

Universitetslektor Karin Dahmström avled efter en tids sjukdom den 27 juni i en ålder av 68 år. Närmast sörjande är maken Per.



Universitetslektor Karin Dahmström.

# Karin Dahmström, en av våra mest kända statistiker

**S**om de flesta svenska statistiker vet används Karin Dahmströms bok "Från datainsamling till rapport" på grundläggande statistikkurser vid nästan alla svenska högskolor. Boken används även vid andra lärosäten än de som traditionellt ger statistikkurser, t ex vid Musikhögskolan. Det betyder att hon förmodligen var en av våra mest kända statistiker. Tiotusentals svenska studenter känner till henne. Boken har nu kommit ut i fem upplagor med ständigt utökad och uppdaterad innehåll. Den sista upplagan var hon osäker på om hon skulle hinna bli färdig med före sin bortgång. Hon satt till och med och skrev samtidigt som hon var intagen och behandlades på Karolinska sjukhuset. Men hon var inte bara en bra läroboksförfattare utan också en omttyckt lärare. I kursutvärderingarna sade studenterna ofta att de var överraskade över hur roligt, intressant och nyttigt det är med kunskaper i statistik. Det är ord som vi andra lärare i statistik inte är så vana vid.

**Karin föddes** den 4 december 1943. Under Karins första år flyttade familjen mellan många platser men när hon var sexton kom de till Spånga. Hon har sedan dess varit en inbiten Spångabo, även om hon några år efter giftermålet med Per bodde i Växjö. Hon hade till sin död fortfarande kvar huset som hon växte upp i. Hon var full av omtanke inte bara om sin man, Per, och sin mamma fram till hennes bortgång, utan också om sin katt, Magnus, som under cirka 15 år spelade en mycket viktig roll i hennes liv.

**Karin började** läsa statistik i Stockholm. Då expanderade universitetet och behovet av lärare ökade snabbt. När hon läst "två betyg" (~60 hp) behövde statistiska institutionen 20 nya "handledare" (~timlärare) och Karin kom med bland dem. Hon blev sedan kvar vid akademien, där hon 1980 disputerade på avhandlingen "Some Binary Response Situations - Theory and Applications" om bl a logistisk regression

och effektivitetsjämförelser med andra binära modeller för små stickprov. Efter disputationen blev Karin lektor, först på Förvaltningshögskolan inom Stockholms universitet, där hon var med och utvecklade en helt ny linje, och sedan vid Statistiska institutionen, där även hennes man var verksam. Sedan dess var hennes främsta intresse studenterna, undervisning och alla delar av produktionsprocessen för statistik, från behovsanalys, problemformulering och enkätkonstruktion över datainsamling och analys till databaser och annan resultatpresentation. I Stockholm gick kursen Undersökningsmetodik och statistisk dataanalys under namnet "Karins kurs".

**Som lärare var** Karin noggrann, ambitiös och krävande. Men hon krävde mer av sig själv än av studenterna. Hon moderniserade till exempel övningsexempel och hemuppgifter årligen till verkliga och aktuella förhållanden istället för att behålla dem och slippa räkna om mellan åren. Det blev ibland för jobbigt för henne själv men studenterna uppskattade henne. Om vi skall lyckas entusiasmera studenter och ge statistik kurserna ett gott rykte behövs sådana lärare.

Karin var intresserad av mycket annat än undervisning. Närmare 70 kvinnliga statistiker minns med glädje den "Första svenska konferensen för kvinnliga statistiker" som hon var en av initiativtagarna till och organisatör för 1995. Bland annat bjöd hon in dåvarande vice statsminister Mona Sahlin som inledningstalare. Det är på tiden att någon annan tar initiativ till en andra konferens.

Vid universitetet valdes hon in i fakultetsnämnden, där hon blev en uppskattad medlem och man lyssnade mycket noga på hennes synpunkter. På äldre dagar medverkade Karin i styrelserna för Svenska statistikfrämjandet och Surveyföreningen. Hon deltog också i en kommitté som tog fram regler för hur bortfall skall redovisas, och valdes också in i en kommitté som skall arbeta fram kvalitetsmått för webbpanelundersökningar. Karin var i kommittéarbetet

synnerligen uppskattad för sin noggrannhet, klarsyn och stilistiska förmåga. Dessutom förde hon ut kommittéernas resultat till omvärlden via sin lärobok och sina statistikkurser. In i det sista var hon aktiv och deltog mindre än en månad före sin bortgång i webbpanelkommitténs senaste möte trots hennes svåra sjukdom.

**Att Karin var** en bra skribent visade hon inte bara i sin bok utan också när hon deltog i Statistikersamfundets uppsatstävling och vann med "Allting varierar - Därför statistik".

Vid ett seminarium med tillbakablick över sin tid som statistiker slutade hon med råden: "Som i alla mänskliga sammanhang: Se varandra, bry er om varandra. Positivt att få beröm. I undervisningen: Aktivera studenterna under hela kursen. Sänk inte kraven. Och till sist: Protestera mot dålig statistik".

**Karin var politiskt** aktiv speciellt inom kyrkopolitiken. Hon var en av initiativtagarna till den opolitiska listan "Spånga-Kista Församlings Vål". Hon blev sedan en respekterad medlem av kyrkofullmäktige och deltog i kyrkorådets arbete. Hon visade ett stort engagemang för församlingens utsatta. Spånga-Tensta är ett område med många invandrare och olika religioner. Karin strävade efter att kyrkan skulle samarbeta ekumeniskt med alla trossamfund. Det är intressant att nämna att hon förmedlade flera intressanta uppsatssämnen från sitt kyrkliga liv. Bland annat om inställningen i kyrkliga frågor och varför så många av medlemmarna lämnade Svenska kyrkan.

**Även i det privata** var hon ambitiös och värdfast. Hon höll kontakt med sina gamla vänner och ställde gärna upp för dem. Efter pensioneringen ringde hon t ex ofta för att prata och höra nyheter från institutionen. Hon kommer att saknas.

DANIEL THORBURN  
GÖSTA FORSMAN  
GEBRENEGUS GHILAGABER

## KALENDARARIUM

### DET HÄNDER 2012

NÄR?	VAD?	VÄR?	WEBBPLATS
24-25 okt	FMS 25 år jubileumsmöte	Utö	<a href="http://statistikframjandet.se/fms/">http://statistikframjandet.se/fms/</a>
31 okt-2 nov	Seminar on New Frontiers for Statistical Data Collection	Genève, Schweiz	<a href="http://www.unece.org/stats/documents/2012.10.coll.html">www.unece.org/stats/documents/2012.10.coll.html</a>
31 okt - 3 nov	International Conference on Methods for Surveying and Enumerating Hard-to-Reach Populations	New Orleans, USA	<a href="http://www.amstat.org/meetings/h2r/2012">www.amstat.org/meetings/h2r/2012</a>
... OCH 2013			
2-5 jan	Statistics, Science, and Society: New Challenges and Opportunities	Chennai, Indien	<a href="http://www.iisaconference.info/">www.iisaconference.info/</a>
21-26 jan	Swiss Epidemiology Winter School	Wengen, Schweiz	<a href="http://www.epi-winterschool.org/">www.epi-winterschool.org/</a>
7-8 feb	Young Statisticians Conference	Melbourne, Australien	<a href="http://www.ysc2013.com/">www.ysc2013.com/</a>
5-7 mars	New Techniques and Technologies for Statistics	Bryssel, Belgien	<a href="http://www.ntts2013.eu/">http://www.ntts2013.eu/</a>
18-23 mars	3rd Joint Statistical Meeting DAGStat 2013	Freiburg, Tyskland	<a href="http://www.dagstat2013.uni-freiburg.de/">www.dagstat2013.uni-freiburg.de/</a>
22-25 apr	7th meeting of the Eastern Mediterranean Region of the International Biometric Society (EMR-IBS)	Tel-Aviv, Israel	<a href="https://event.pwizard.com/ims/index.py?p=620">https://event.pwizard.com/ims/index.py?p=620</a>
10-12 juni	The Stockholm Criminology Symposium	Stockholm	<a href="http://www.criminologysymposium.com/symposium/event-information.html">http://www.criminologysymposium.com/symposium/event-information.html</a>
10-12 juni	Nordic-Baltic Biometric Conference	Stockholm	<a href="http://nbbc13.org/">http://nbbc13.org/</a>
3-8 aug	Joint Statistical Meetings	Montréal, Kanada	<a href="http://www.amstat.org/meetings/">www.amstat.org/meetings/</a>
18-24 aug	The Eighth International Research Forum on Statistical Reasoning, Thinking, and Literacy	Minneapolis, USA	<a href="http://isi-web.org/news/2012-feb-iase-news">http://isi-web.org/news/2012-feb-iase-news</a>
9-11 sep	3rd European Establishment Statistics Workshop (EESW13)	Nürnberg, Tyskland	



### »SOMMARTID ÄR HÖGSÄSONG FÖR INBROTT»

# En del myter sitter som berg

**S**ommartid är högsäsong för inbrott.

– Högsäsong för inbrott, säger Erik Thulin, Polisen Östergötland, 10 juni 2011.

Ja, det vet ju alla. Men vad är det här?

– Färre inbrott på sommaren, SR P4 Sjuhärad, 18 juli 2011.

– Det är en myt att påstå att de flesta bostadsinbrotten sker under sommarmånaderna, säger Monica Berntsson Drake, tillförordnad spaningschef vid polisen i Alingsås, till Göteborgs-Posten, 14 november 2011.

”Inbrottstjuvarna slår till på hösten” lyder en rubrik i Dagens Nyheter den 5 september 2012.

**En del myter** sitter som berg. ”Kinesiska muren är det enda byggnadsverk på jorden som syns från månen”. ”USA har 52 stater”. Kan man ta död på en myt med statistik eller med att redovisa fakta?

Sommaren är inte inbrottstjuvarnas högsäsong. Det är under hösten och vintern som det sker flest inbrott, enligt statistik från Brottsförebyggande rådet, Brå. Figuren ovan visar ett kraftigt säsongmönster för

anmälda inbrott i villa och radhus i vissa delar av landet. Särskilt i Västra Götalands, Hallands och Skåne län sker många av inbrotten under fjärde kvartalet. Det är statistik från Brottsförebyggande rådet som uppmärksammats av bland andra Monica Berntsson Drake och Kristoffer Örstadius, som skrev artikeln i DN den 5 september.

**Figuren finns publicerad** i ”Vil-lainbrott – en statistisk kortanalys”, Brottsförebyggande rådet, 2011. Där finns även statistik från år 2000 som visar säsongvariation i de tre mest

utsatta länen: Västra Götalands, Hallands och Skåne län. Rapporten är skriven av Anton Färnström och Olle Westlund.

Vi tackar också Peter Olausson som samlat citat från olika källor och vars text vi lånat ovan. Peter Olausson skriver om ”faktoider”, som finns utgivna på <http://faktoider.blogspot.se/> och bland annat i boken ”Blindspår: vetenskap på villovägar” utgiven i år på Leopard förlag.

