ÅRETS CRAMÉR-PRISTAGARE

SIDAN 20

BÄSTA SURVEY-UPPSATS

SIDAN 27

SSFr
Svenska statistikfrämjandet

Qvintensen

SVENSKA STATISTIKFRÄMJANDET NR 2 2014

Att väga ihop opinionsundersökningar

SIDAN 10

Transparens är ett nyckelord när det gäller redovisning av undersökningsresultat.” Rekommendationer för undersökningar med webbpaneler.

SIDAN 25

Hur mäter man gång?

– Jo Røislien, årets statistikfrämjare, om statistik och medicin

SIDAN 15

Hantera multipla test: att justera eller inte

SIDAN 11

Hack for Sweden

SIDAN 6

PRICKSÄKRA KUNSKAPER OM IMORGON. BARA SÅ DU VET.

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

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

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

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

Välkommen in.

sas | THE POWER TO KNOW.

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*

Qvintensen

SVENSKA STATISTIKFRÄMJANDET

NR 2 2014

Jo Røislien.

Andrea Gustafsen och Bergdís Sigurjónsdóttir – prisade.

Hack for Sweden.

Nestor Arne Modig.

Innehåll

HACK FOR SWEDEN 6-7

TEMA OPINIONSUNDERSÖKNINGAR/
Hur utförs opinionsundersökningar idag?
Intervju med Arne Modig 8-9
Att väga ihop olika opinionsundersökningar 10
Feministiskt intiativ får 18 % av rösterna 10

DEALING WITH MULTIPLE COMPARISONS: TO ADJUST OR NOT TO ADJUST 11-14

ÅRETS STATISTIKFRÄMJARE: STATISTIK OCH MEDICIN 15-16

Vad ska vi med statistik till? Vi ska rädda liv. »

Jo Røislien, sidan 15-16

Fasta spalter

I HUVUDET PÅ EN REDAKTÖR 4

ORDFÖRANDEN HAR ORDET 5

HÄNT & HÖRT 14

Våra föreningar

INDUSTRIELL STATISTIK
Sex Sigmas historia 17-18

CRAMÉRSÄLLSKAPET
Ny ordförande i Cramérsällskapet – hedrande och spännande 19
Årets Cramérpristagare: Complex Models for Complex Data via the Bayesian Approach 20-22

SURVEYFÖRENINGEN
Vårkonferens i vårvintrigt Stockholm 23
Rapport från webbpanelkommittén..... 24-26
Bästa uppsats 27

FÖRENINGEN FÖR MEDICINSK STATISTIK
Nu är det dags för dig att ta över stafettpinnen!..... 28-29
Statistics Alps 2014 30

I HUVUDET PÅ EN REDAKTÖR

Vad ska vi med statistik till?

Hur vet vi att rökning höjer risken för cancer?
Statistik är en bred disciplin med många förgreningar. En av dem är biostatistik som är livets statistik.

Anna Torrång skrev en gång i en krönika att en del tror att biostatistiker räknar biobesökare. Vetenskapen statistik spelar en central roll vid Sveriges internationellt högst rankade högskola: Karolinska Institutet. Svenska statistikfrämjandets nye ordförande John Öhrvik verkar där som professionell statistiker. Han presenterar sig på sidan 5 i det här numret av Qvintensen.

I statistiska tillämpningar används ett stort maskineri av principiella resonemang, matematik, statistiska metoder och kunskaper. Statistiska test som principiellt resonemang väcker ständig debatt. Två andra KI-statistiker, Yudi Pawitan och Arvid Sjölander, skriver om hur man kan resonera när man testar.

Statistik är svårt. Somliga har luftat sina förhoppningar om att det förhatliga ämnet statistik snart har försvunnit. Nu när vi har så mycket data ("big data") kommer statistikämnet att tyna bort! De förstår inte att det krävs ett vederhäftigt resonemang för att ta sig från råa data till övertygande och väl kommunicerade slutsatser.

De statistiska analysernas tyngd fick till slut de stora tobaksbolagen på knä.

»Folk tycker om siffror och grafik.»

Opinionsundersökningar är inte konstigare än några andra undersökningar. Det som gör dem till särlingar är deras plats i samhället. Resultaten kommuniceras via massmedia där varje kanal har "sin egen" opinionsundersökning. De principiella resonemangen och de statistiska metoderna är ganska snålt kommunicerade, men personerna på opinionsinstituten svarar gärna på frågor när jag ringer. Men varför görs det så många opinionsundersökningar? Jag tror att svaret är att en opinionsundersökning är nära nog den perfekta nyheten. De ger alltid något att skriva om, de kan till skillnad från de flesta andra nyheter planeras i förväg - och läsarna är ju intresserade! Folk tycker om siffror och grafik.

DAN HEDLIN

FOTO: LIESELOTTE VAN DER MEIJS

Qvintensen

Ansvarig utgivare
John Öhrvik
Redaktör
Dan Hedlin, 070 297 10 27
Ingegerd Jansson (biträdande)
Redaktion
Hans Ekholm, Främjandet
Alf Fyhrlund, Främjandet
Marie Linder, FMS
Kadri Meister, Surveyföreningen
Maria Blixt, Industriell statistik
Lars Rönnegård, Cramérsällskapet
Anna Torrång
Adress
Qvintensen, c/o Dan Hedlin
Dammtorps allé 6, 170 62 Solna
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
John Öhrvik, 070-666 08 08,
John.Ohrvik@ki.se
Vice ordförande
Inger Eklund, 08-506 947 13
Sekreterare
Thorbjörn Gudmundsson
Klubbmästare Anders Hallberg
Skattmästare Jonas Selling
Surveyföreningen
Maria Öman
FMS
Frank Miller, 08-16 29 76
Industriell statistik
Jenny Jonasson
Cramérsällskapet
Maria Karlsson, 090-786 55 96
Övriga ledamöter
Marie Wiberg, Stig Johan Wiklund
Adress
Svenska statistikfrämjandet
c/o: T Gudmundsson
Institutionen för matematik
KTH, 100 44 Stockholm
E-post sekrfram@gmail.com
Webbplats www.statistikframjandet.se

ISSN 2000-1819

Författare och andra som utan ersättning bidrar med material för publicering behåller upphovsrätten till sina verk. Svenska statistikfrämjandet äger rätten att publicera dessa verk i valfri form och upplaga samt att behålla alla eventuella intäkter från sådan publicering, om inget annat avtalas. Författare och andra som utan ersättning bidrar med material för publicering garanteras rätten att själva få sprida sitt verk under förutsättning att Qvintensen anges som originalkälla med angivande av nummer och år. Svenska statistikfrämjandet äger upphovsrätten till publicerat material då det är framställt av redaktionsmedlem, styrelsemedlem, någon till Främjandet knuten förenings styrelse eller person som erhållit arvode för sitt bidrag. © Svenska statistikfrämjandet 2014.

Svenska statistikfrämjandet

ORDFÖRANDEN HAR ORDET

BIOSTATISTIK – gränssnitt mellan biomedicinsk och statistisk vetenskap

Jag önskar alla gamla och nya statistikfrämjare välkomna till ett nytt spännande statistikår. Från och med i år är vi som förening medlem av The Federation of European National Statistical Societies. Vad som händer inom Fenstats vill jag återkomma till, men besök gärna webbplatsen fenstats.eu. Själv är jag ny som ordförande i Svenska statistikfrämjandet och jag tycker därför att det kan vara på sin plats att jag ger en presentation av mitt forskningsområde som professor i biostatistik vid Karolinska Institutet.

Biostatistik kan beskrivas som ett gränssnitt mellan biomedicinsk och statistisk vetenskap – tillämpning av modern statistisk teori för att studera variation och korrelation i levande organismer. En central del i min forskning omfattar utveckling och tillämpning av robusta statistiska metoder i studier av genetiska sekvensvariationer som förklaring av olika hjärt/kärlsjukdomar och metabola störningar.

»Från i år är vi som förening medlem av The Federation of European National Statistical Societies.»

Kartläggningen av det humana genomet har inneburit enorma möjligheter att studera samband mellan genomets struktur och egenskaper i hela människan, men har ställt forskningen inför stora utmaningar när det gäller att tolka och förstå denna information. Det mänskliga genomet har cirka tre miljarder baspar, nukleotider. En singel nukleotid polymorfism (SNP) är den vanligaste formen av genetisk variation och förekommer i genomsnitt i tre promille av basparen. Det är bara en liten andel av dessa SNP:ar som är funktionella och därmed kan resultera i en skillnad mellan två människor. Dessa SNP:ar är mycket betydelsefulla för att förstå samspelet mellan vårt genom och vår hälsa. Robusta biostatistiska metoder för att identifiera funktionella SNP:ar är därför helt avgörande för vår förståelse av detta samspel. Förstår vi det, kan vi förbättra prevention och behandling inom olika medicinska områden såsom hjärta/kärl och diabetes.

Idag analyseras sambanden mellan olika sjukdomar och SNP:ar oftast univariat, vilket innebär att man gör ett mycket stort antal samtidiga test, ett per SNP. Men då blir antalet så kallade falska positiva fynd väldigt stort; vi tror oss ha funnit ett samband mellan en SNP och en sjukdom men fyndet återspeglar bara att vi tolkat våra data på fel sätt. Detta innebär en risk för framtida patienter och poängterar ytterligare behovet av effektiva statistiska metoder för att öka förståelsen av samspelet mellan våra gener och sjukdomar. Traditionella korrigeringar av signifikansnivån för multipla test i syfte att minska risken för falska positiva fynd, till exempel Bonferroni-korrigering, leder till mycket konservativa testkriterier och därför har metoder som ”false discovery rate” använts, men även de har sina begränsningar.

För närvarande studerar jag möjligheten att förbättra dessa korrigeringar genom att utnyttja att SNP:ar funna genom sekvensering av hela genomet är korrelerade, och därmed kan representeras av ett mindre antal okorrelerade variabler.

Jag utvecklar också metoder för intern validering, det vill säga validering i samma studiepopulation som det ursprungliga fyndet gjorts i. Hur det går till vill jag återkomma till.

I en större europeisk studie, IMPROVE, mättes tjockleken på avståndet mellan ytterväggen och endotelet i halspulsådrornas (karotis) kärlväggar. Dessa kan förtjockas genom plackbildning, vilket innebär en risk för proppbildning. Vi har kunnat identifiera och validera ett antal nya SNP:ar som signifikant ökar risken för plackbildning med hjälp av bootstrap-baserad intern validering. Vi hoppas att dessa fynd i framtiden kan komma att förbättra diagnostiken för patienter inom hjärt/kärlområdet.

Avslutningsvis önskar jag alla en skön och avkopplande sommar.

John Öhrvik, ordförande

Giftspridnings-simulator – Hack for Sweden Award

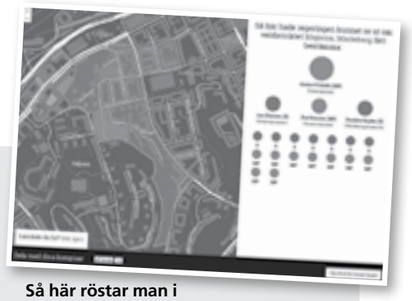
En webb-app för analys av hur giftiga utsläpp sprider sig i naturen baserat på topografi, moränlager och vattenföring. Appen ger också en indikation på vilka vattenskyddsområden, sjöar och brunnar som löper risk att påverkas vid ett sådant utsläpp. Tävlingsbidraget bygger på data från SLU (skog), SGU (jordlager och brunnar), MSB (krisdata), Naturvårdsverket (vattenskyddsområden), SMHI (prognos för nederbörd) och Google Maps.

http://prezi.com/umqrrzkwawmg/rsas-risk-simulation-and-notification-system/

App för Artportalen – bästa allmännytta

En iphone-app som stöder www.artportalen.se som är ett webbaserat system för fynd av Sveriges växter, djur och svampar. Appen kan inspirera människor att komma ut i naturen genom att det blir roligare och enklare att använda Artportalen. Artportalen drivs av ArtDatabanken vid SLU på uppdrag av Naturvårdsverket och i samarbete med myndigheter och ideella organisationer.

www.youtube.com/watch?v=Ya3-dRRJ_f4&feature=youtu.bef4&feature=youtu.be



Så här röstar man i Majorna i Göteborg.

Kvartersregeringen – bästa visualisering

En app som visar hur Sveriges regering skulle se ut om just ditt kvarter fått bestämma. Man skriver in en adress och utifrån det senaste valresultatet för valdistriktet räknar appen ut en tänkbar regeringssammansättning. Utan procentsatser, paj- eller stapeldiagram ökar förståelsen för användarens närområde. Appen bygger på valdata från SCB och OpenStreetMap.

www.kvartersregeringen.se/

Riskminimering vid odling – bästa affärsnytta

Bidraget Virtuella mätinstrument för en hållbar framtid kombinerar egna sensordata med myndighetsdata, vilket ger bättre möjligheter för jordbrukaren att spara pengar och minimera miljöbelastningen. Tävlingsbidraget bygger på öppna väderdata från SMHI, som kompletterar ett befintligt system med fysiska sensorer i fält.

https://graph.sensefarm.com/livePanel.php?period=last_week

Hitta nästa hamn – juryns specialpris

Juryn delade också ut ett specialpris till tävlingens yngsta deltagare (två tolvåringar) som presenterade sin idé till en app som hjälper sjöfarare att hitta nästa natts lugna hamn. Appen bygger på data om väder (SMHI), algblomning (HaV), brunnar (SGU), livsmedelsbutiker och drivsmedelsstationer (Tillväxtanalys), fornminnen (RAÄ) och naturreservat (Naturvårdsverket).

http://prezi.com/w4fxblh_g7jx/skargardshackarna/?utm_campaign=share&utm_medium=copy

Hitta skyddsrum – bästa krisapp

Krisinformation.se som drivs av Myndigheten för samhällsskydd och beredskap prisade tävlingsbidraget Hitta skyddsrum. Tjänsten ger snabb vägledning till hur man vid en katastrof hittar till närmaste skyddsrum. Syftet är att uppmärksamma och aktualisera frågan om hur man ska bete sig i en krissituation, och att i skarpt läge hjälpa människor att hitta skydd och vård. Tjänsten bygger på data om skyddsrum från Myndigheten för samhällsskydd och beredskap.

www.hittaskyddsrum.se/

De vinnande tävlingsbidragen

■ Länkarna finns även på http://statistikframjandet.se/qvintensen/

Vinnarna av Hack for Sweden Award får åka på en konferens om mobil app-utveckling i Berlin. Dessutom fick man presentera sitt bidrag på Kartdagarna i mars.

Hack for Sweden

Media skriver spaltmeter om myndighetens evenemang, unga IT-utvecklare pratar initierat och positivt om en myndighetsdatabas i sociala media, ett stort antal myndigheter samarbetar för att de själva vill det och bedömer att det är bästa sättet att nå framgång. Låter det som en utopi? Under Hack for Sweden blev det verklighet!

Tidigt på våren 2013 började tankarna ta form om att flera myndigheter gemensamt skulle arrangera ett hackaton. Det var visserligen flera av oss som hade rätt dimmiga begrepp om hur ett hackaton går till, men vi visste alla att vi ville skapa ett arrangemang för att öka användningen av våra data. Framåt sommaren hade det blivit klart att vi var ca tio myndigheter som ville arrangera ett hackaton i mars 2014 – Hack for Sweden. När tävlingen gick av stapeln var vi tretton arrangerande myndigheter (se listan här bredvid).

Projektgruppen formades och vi började ta itu med alla de saker som måste finnas på plats innan Hack for Sweden kunde bli verklighet. Det första var att ta fram vilka data som skulle ingå i tävlingen. Utgångspunkten var data som har geografisk anknytning. Vi hade många diskussioner om format och tillgänglighet eftersom alla myndigheter har sina egna databaser och sätt att tillgängliggöra och tillåta användningen av data. Vi försökte göra en samlad beskrivning av alla data som ingick. Deltagarna ser gärna att myndigheterna ytterligare förbättrar tillgängligheten, och ska det bli ett nytt Hack for Sweden så vill man ha data från myndigheterna på ett mer enhetligt sätt.

I slutet av hösten godkände Vinnova vår ansökan om 200 000 kr för att täcka merparten av de externa kostnaderna. Flera myndigheter bidrog också med totalt ungefär lika mycket. Alla myndigheter svarade förstås för arbetstiden för de personer som arbetade med att planera och genomföra Hack for Sweden.

Vi hoppades att IT-utvecklare, studenter och datajournalister skulle anmäla sig till tävlingen, men detta är inte någon målgrupp myndigheterna normalt har täta kontakter med, så hur kommunikationen skulle läggas upp var viktigt. Vi fokuserade på en webbplats, Facebook och Twitter. Dessutom gjorde vi riktade e-postutskick till universitet och högskolor och använde oss av myndigheternas egna webbplatser och nyhetsbrev. Rent juridiskt är det komplicerat att använda sig av sociala medier när det är många myndigheter tillsammans som är ansvariga, så den externa kommunikationen kom inte igång förrän i slutet av januari, mindre än två månader innan själva eventet skulle gå av stapeln. De två veckorna runt tävlingshelgen hade vi en räckvidd på närmare 4 000 personer per vecka på Facebook.

Vi fick snabb spridning i de sociala medierna och anmälningarna tog fart i februari. Till slut landade antalet deltagare på ca 75 personer. Det var en bra blandning i ålder, kanske kunde man ha önskat sig något fler kvinnor. Det var också blandat med studenter och företag, med fokus på IT. Datajournalisterna var ganska få. Vi var inte beredda på att så många helst kommunicerade på något annat språk än svenska – mer information på engelska hade varit bra.

Att lokalen är viktig för ett hackaton var det många som poängterade. Att Naturvårdsverket upplät sina lokaler till Hack for Sweden var ett lyckokast – Naturvårdsverket huserar i fina ljusa lokaler på Vallhallavägen i Stockholm där det lätt går att kombinera både gemensamma samlingar och arbete i grupprum. Dessutom finns en restaurang i direkt anslutning och vi lyckades ordna ”god mat och mycket mat och mat i rättan tid” till deltagarnas belåtenhet. Vår yngsta deltagare, Leo 12 år, sa att ”…det här var den bästa dagen någonsin. Ingenting var dåligt. Inte ens maten var äcklig.”

Det var viktigt för oss att knyta till oss kompetens om målgruppen och om hur man arrangerar ett hackaton. Vi hade diskussioner med många olika personer, men de som hade störst betydelse var utan tvekan vår jury; Pernilla Rydmark, Kristofer Sjöholm, Peter Krantz och Jonas Lejon. Deras primära uppgift var förstås att välja ut de tävlingsbidrag som skulle tilldelas pris, men deras erfarenhet från att arrangera och delta i hackaton var viktig. Vi hade flera möten med dem som på ett avgörande sätt bidrog till att vi la upp Hack for Sweden på ett framgångsrikt sätt.

Vinnarna av Hack for Sweden Award 2014, från vänster Marcus Kempe (SP Sveriges Tekniska Forskningsinstitut), Heiti Ernits (SP Sveriges Tekniska Forskningsinstitut), Per-Olov Jernberg (Spotify) och Mats Törnberg (Cavagent)

De 21 lagen som kom till start hade ganska olika förutsättningar. Några hade redan jobbat med idéer som skulle utvecklas med nya myndighetsdata, andra började med ett blankt papper och hade ganska vaga planer för vad det närmaste dygnet skulle innehålla. Många tog hjälp av ämnesexperter från myndigheterna som fanns på plats under tävlingens första dag. Resultatet blev tjänster inom väldigt olika områden och även om man hann olika långt så höll tävlingsbidragen mycket hög klass.

CECILIA WESTSTRÖM,
STATISTISKA CENTRALBYRÅN

FOTO: LENA BLÅSJÖ-JANSSON

Vad är ett hackaton?

■ Ett hackaton är ett evenemang där programmerare träffas för att programmera och inspireras av varandras färdigheter i att skriva kod. Ett hackaton genomförs vanligtvis under fria former i 24 timmar eller över en helg. Ordet myntades 1999.

Myndigheterna som arrangerade Hack for Sweden

- Naturvårdsverket
- Statistiska centralbyrån
- SMHI
- Trafikverket
- Myndigheten för samhällsskydd och beredskap
- Sveriges geologiska undersökning
- Sveriges lantbruksuniversitet
- Socialstyrelsen
- Lantmäteriet
- Myndigheten för tillväxtpolitiska utvärderingar och analyser
- Riksantikvarieämbetet
- Havs- och vattenmyndigheten
- Rådet för främjandet av kommunala analyser

Nu är det ”supervalår” och undersökningarna om politiska opinioner ges ännu större uppmärksamhet än vanligt. Inför förra riksdagsvalet publicerades fler än 40 opinionsundersökningar av de olika företagen i branschen. Vi träffade en person som är mycket välkänd i branschen: Arne Modig.

Hur utförs opinionsundersökningar?

ARNE MODIG, BRANSCHENS NESTOR, OM UNDERSÖKNINGARNA SOM FRAMKALLAR BRASKANDE RUBRIKER

Gör opinionsinstituten särskilda undersökningar med partisympatifrågor?

– Nej, i de flesta fall ingår frågor om partisympatier i en omnibus som opinionsinstituten gör varje vecka. Det är en undersökning med en mängd olika frågor. Vilka frågorna är beror på vad kunderna har begärt och betalat för. Frågorna om vilket politiskt parti man föredrar kommer alltid först i omnibussen för att svaret på den frågan inte ska påverkas av andra frågor.

– Det tar cirka fyra-fem dagar för ett opinionsinstitut att genomföra 1 000 eller fler intervjuer i en omnibus.

– Det är i själva verket tre frågor som ställs. Först frågar alla institut om vilket parti man skulle rösta på om det vore val idag. Om intervjupersonen inte vet det, frågar man vilket parti det lutar åt. Kan intervjupersonen inte uppge det heller, frågar man vilket parti som är minst dåligt. Det som redovisas är svar på någon av de tre frågorna. Ungefär 65% av dem som man fått kontakt med besvarar första frågan med ett politiskt parti. Ungefär 12% svarar på andra frågan. Ungefär 80-85% av dem man fått kontakt med svarar med ett politiskt parti på någon av de tre frågorna.

– Ipsos ställer inte den tredje frågan. De får därför en något större andel osäkra än de andra instituten.

Hur drar instituten sina urval av personer som ska tillfrågas?

– Det är olika. Under 1980- och 90-talen dominerade random digit dialling. Man utgick från ett befintligt telefonnummer, manipulerade det så att ett slumpmässigt telefonnummer bildades och det ringdes upp av en maskin. När någon svarade fanns en intervjuare beredd.

– Något institut drar ett befolkningsurval ur befolkningsregistret och intervjuar per telefon.

– De flesta opinionsinstitut har databaser med webbadresser och epostadresser. Där har man uppgifter om kön, ålder, om man äger en bil, vilken sorts bil det är i så fall, uppgifter om shoppingvanor, vilket eller vilka politiska partier man kan tänka sig att rösta på – den typen av uppgifter. Ett opinionsinstitut kan ha 200 000 personer i sin databas. Man viktar de som svarar på så sätt att man får ett representativt urval.

– Dåliga svarare tas bort ur databasen. De som svarar får någon typ av gåva. Det kan vara poäng, och när man samlat tillräckligt många poäng kan man byta dem mot någon produkt i opinionsinstitutets webbutik för respondenter.

Hur görs estimationen?

– Alla gör någon form av vägning av stickprovsandelarna. Man väger med bland annat kön, ålder och region. Dessutom har man det parti intervjupersonen uppger att han eller hon röstade på i förgående val med som vikt. Den vikten justeras ofta på delvis subjektiva grunder. Demoskop slutade att väga med föregående val ett tag men återinförde den vikten eftersom skattningarna blev instabila.

Vad kostar det att få ställa en fråga till svenska folket?

Det kostar kunden cirka 8 000-10 000 kronor att få med en fråga i en omnibus med 1 000 intervjuer, men priset beror av hur många andra kunder man har med i omnibussen.

Så idag använder man telefon och webbenkäter. Ställer man frågor med pappersblanketter fortfarande?

– Sifo gör fortfarande en undersökning där man skickar ut en tjock pappersblankett. Den kallas Orvesto och är till största delen en marknadsundersökning. Den är ett undantag; pappersblanketter används annars nästan inte alls inom den kommersiella delen av surveybranschen.

Arne Modig.

Skulle du säga att opinionsinstituten är metodologiskt seriösa?

– Jadå. De flesta opinionsinstituten är tekniskt bra. De tycker att metodfrågor är viktiga.

Hur är det med bortfallet?

– Idag har man över 50 procent bortfall i telefonundersökningar. Undersökningarna görs under tidspress med begränsad period att följa upp dem som är svåra att nå. En del av instituten ”sparar” på respondenter som man inte har fått kontakt med till nästa omnibus. På så sätt får man med en del av dem som är svåra att nå i varje omnibus. Men det gör bortfallsredovisningen besvärlig.

– Det är vanligt att man försöker få kontakt i 14 dagar. Man gör sina försök enligt utprovade scheman. Om det till exempel är upptaget ska man ringa igen inom 30 minuter.

Har opinionsinstituten sina egna intervjuare?

– Flera av dem har sina egna intervjuarkårer. En del anlitar någon extern firma.

Diskuterar opinionsinstituten med varandra? Eller håller var och en sig på sin egen kant?

– Jag ordnade branschdiskussioner. De pågår fortfarande. Dessutom känner många av dem som jobbar med opinionsundersökningar varandra.

Varför gör de större politiska partierna sina egna opinionsundersökningar?

– De vill ha förklaringar, inte bara siffror på hur många som skulle rösta på det egna partiet eller på motståndarna. De vill alltså ha rikare information.

Hur hamnade du i opinionsundersökarbranschen?

– Efter universitetet började jag på SCB, på deras Utredningsinstitut. Jag ville göra något annat, söka något äventyr. Jag hamnade på IMU. Det var ett företag som gjorde marknadsundersökningar. Efter en tid började vi med opinionsundersökningar. Via uppköp och företagsombildningar blev min arbetsplats till slut Temo. Det sista opinionsinstitutet jag jobbade på var Novus.

Vad gör du numera?

– Nu driver jag och en kollega ett litet företag som heter Calmando. Vi ger kurser i opinionsbildning. Många av dem som går kurserna jobbar som informatörer.

DAN HEDLIN

Novus sammanställer

Novus uppdaterar och lägger ut excelfiler med resultaten av opinionsinstitutens väljarbarometrar. Det var Arne Modig som började med dem när han jobbade på Novus. Länken är http://www.novus.se/vaeljaropinionen/samtliga-svenska-vaeljarbarometrar.aspx

Kommersiella företag som utför opinionsundersökningar i Sverige. Endast de som publicerar resultaten är med i listan.

Företag	*När de började med opinionsundersökningar*
Sifo	1954
Ipsos	1970-talet (dåvarande IMU)
Demoskop	1990
Novus	2006
Skop	1985
United Minds	1999
YouGov	2011
Forskningsgruppen för samhälls- och informationsstudier	1971

Att väga ihop olika opinions-undersökningar

Opinionsinstitut använder telefonintervjuer eller webbenkäter, olika långa insamlingsperioder och olika grad av bearbetning, t.ex. efterstratifiering. Alla har dock som mål att spegla partisympatierna i valmanskåren vid en viss tidpunkt. Olika mediaföretag samarbetar med olika opinionsinstitut och därför presenteras vanligen olika mätningar i olika media. Två frågor som man då ställer sig är: spelar det någon roll vilken mätning man tar del av? Kan man väga ihop olika opinionsundersökningar till ett sammanfattande resultat, en s.k. poll of polls?

En viktig sak att observera när det gäller undersökningar av partisympatier är att de anger andelen väljare (d.v.s. den relativa frekvensen väljare) och inte antalet väljare (den absoluta frekvensen). Andelen som sympatiserar med ett parti är beroende av de andelar som sympatiserar med andra partier. Om ett parti ökar måste ökningen tas från andra partier, som alltså minskar. Detta fenomen påpekades redan av Pearson (1897), som kallade det ”spurious correlation”, men det har dessvärre inte alltid observerats av sentida efterföljare.

Jakob Bergman och Björn Holmquist har utvecklat en metod som utgår från att andelarna är ”compositions”; en composition är en summa av slumpvariabler som alltid summerar sig till en och samma konstant (Bergman och Holmquist 2013). I det här fallet är den konstanten 1, eftersom andelarna alltid summerar sig till 1. ”Sammansättning” är kanske en lämplig svensk term för composition.

I Figur 1 visas resultatet av 509 partisympatiundersökningar mellan september 2006 och juni 2013 utförda av åtta opinionsinstitut samt SCB. Data kommer från Novus (2014). De sammanvägda skattningarna för de olika partierna visas som heldragna linjer.

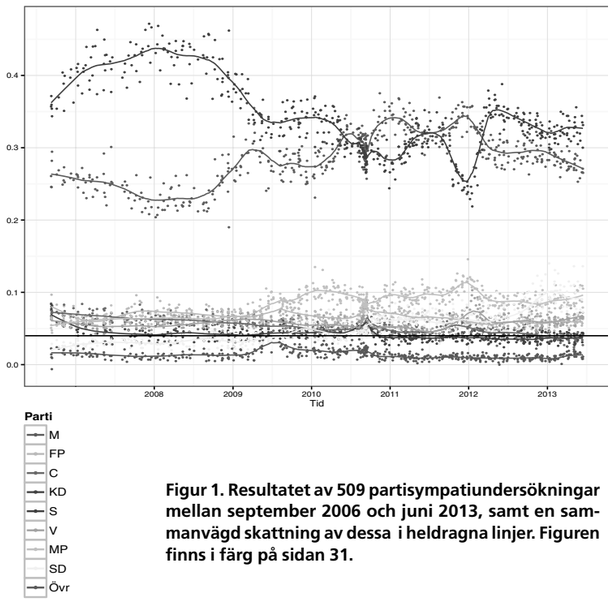
I nästa nummer skriver Jakob Bergman mer om vägningsmetoden och om det spelar någon roll vilken mätning man tar del av.

Referenser

Bergman, J. & Holmquist, B. (2014). Poll of polls: A compositional loess model. Scandinavian Journal of Statistics, 41, 301-310.

Novus (2014). Samtliga svenska väljar barometrar. Webbadress: http://www.novus.se/vaeljaropinionen/samtliga-svenska-vaeljarbarometrar.aspx (Excelfil hämtad i april 2014).

Pearson, K. (1897). Mathematical contributions to the theory of evolution.— On a form of spurious correlation which may arise when indices are used in the measurement of organs. Proceedings of the Royal Society of London, LX, 489-498.



Figur 1. Resultatet av 509 partisympatiundersökningar mellan september 2006 och juni 2013, samt en sammanvägd skattning av dessa i heldragna linjer. Figuren finns i färg på sidan 31.

Hur gick det till när Gudrun Schymans Fi fick 18 % i en opinionsundersökning?

Feministiskt initiativ får 18% av rösterna

■ Richard Öhrvall skriver på sin blogg i Svenska Dagbladet hur det gick till när Feministiskt initiativ fick 18% i en opinionsundersökning utförd av YouGov. ”Att fela är mänskligt. Men det är viktigt att ha kontroller för att upptäcka de fel som oundvikligen begås. Här verkar det ha brustit hos YouGov.”

Richard Öhrvall fortsätter: ”Samtidigt sätter det som har hänt fingret på ett av problemen med opinionsmätningar: instituten som genomför dem redovisar ytterst begränsad information om hur själva mätningarna går till. För oss som konsumerar opinionsmätningar vore det välkommet med redovisning av hur urvalet har gått till, hur stort bortfallet var, hur resultaten har viktats, med mera. Mitt förslag är att tidningar bör kräva sådana grundläggande uppgifter innan de förmedlar institutens resultat.”

Webbadress till bloggen: http://blog.svd.se/trehundratolv/2014/05/17/galna-opinionssiffror/

Dealing with multiple comparisons:

To adjust or not to adjust

Let's start by saying that dealing with multiple comparisons is tricky, perhaps not mathematically, but logically for sure. The wrapping of this problem should be labeled with at least two warnings, one from Oscar Wilde: 'the truth is rarely simple and never pure' and the other from Einstein: 'as far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality'. Einstein's statement was made in a lecture on geometry in relation to experience, partly highlighting his insight that Euclidean geometry fails to explain reality, so any branch of mathematics that is less intuitively obvious in its relation to reality than geometry has an even larger gap to close.

»In assessment of evidence our intention matters»

Multiple comparisons arise naturally in most scientific studies because of the need to capture or convey a lot of information with many variables. But there are many steps taken during a scientific process that also have multiplicity implications, such as variables to be included in the analysis, subgroup analysis, etc. Even a simple decision as deciding on a proper transformation has a multiplicity implication as we can search for a transformation that produces a more significant result.

First, the easy part: mathematical certainties can be derived assuming the null hypotheses are true, such that, for example, 'at 5% significance level, for every 100 tests we shall get 5 spuriously significant tests on average.' Alternatively, assuming independence and still under the null, it is virtually certain (probability =1-(1-0.05)100=0.994) that the test among the 100 tests with the smallest p-value will have a raw (i.e. unadjusted) p-value <0.05. This means that if we are naïve we would be easily misled by false positives, hence the usual warning about data fishing.

In practice, to account for multiplicity, one can simply adjust the p-value by multiplying it by the number of tests (denoted by M), or dividing the nominal significance level by M. So instead of using a level of 0.05, we would use the level 0.05/M, or, alternatively, instead of using the raw p-value, we would use the adjusted p-value obtained by multiplying the raw p-value with M. This makes it harder to declare significance. Using this so-called Bonferroni correction, we can guarantee that if we only reject hypotheses with adjusted p-value 0.05, say, then the probability of making any false rejection among all tests is ≤ 0.05. To highlight the mathematical certainty, suppose we have M null hypotheses, where a subset containing M0 hypotheses are true; let A be the event that we obtain at least one false rejection, then

Prob(A) = Prob(adjusted p-value ≤ 0.05 for at least one hypothesis)
= Prob(p-value ≤ 0.05/M for at least one hypothesis)
≤ M0*0.05/M
≤ 0.05.

Due to its simplicity, this procedure is commonly used in practice. If we want to protect against false positives, it seems reasonable to apply this procedure. The literature on multiple testing is enormous and there are numerous alternative methods. Many are improvements of the Bonferroni correction, indicating that »»»

DEALING WITH MULTIPLE COMPARISONS: TO ADJUST OR NOT TO ADJUST

multiple-testing correction is taken seriously, and when applied, it is mathematically straightforward.

Example 1: Cancer drug study. Table 1 shows the p-values in a study of a metastatic cancer drug vs placebo for ten patient characteristics. The most significant raw p-value is 0.007 for the Karnofsky index (let's keep it mysterious for now), but when corrected using the Bonferroni method, nothing is significant at the 5% level. (The last column actually gives an estimated number of false positives when we consider the variable and those above it as significant, so we do not truncate it at 1.00. We will use this later.)▫

Yet the statistical concerns are not shared universally among scientists. The clearest objection was formulated by Rothman (1990), who declared 'No adjustments are needed for multiple comparisons'. Being the first editor of the journal Epidemiology, Rothman is one of the most influential (and highly cited) epidemiologists, so it is worth understanding his arguments. Briefly, two non-mathematical presumptions are needed for application of multiplicity adjustments: (i) the 'universal null hypothesis', covering all hypotheses under consideration, is a reasonable state of nature, so that chance does cause many unexpected findings, and (ii) no one would want to investigate further something caused by chance. He argued that in a scientific process these presumptions are not true, that 'chance' is not a scientific explanation, so scientists should 'grasp at every opportunity' to understand unusual findings, and that the possibility of being misled is part of the trial-and-error process of science.

It is still the case today that most statistical results in epidemiology and medical literature are rarely adjusted for multiple comparisons, with notable exceptions in clinical trials and high-throughput molecular studies; see more below. In clinical trials, multiplicity issues arise in, for example, the choice of hypotheses to be tested, sidedness of the test, interim analyses during the trial, main analysis plan, subgroup analyses after the trial, etc. The US Food and Drug Administration's (FDA) "Guidance for Industry: E9 Statistical Principles for Clinical Trials" stipulates that these should be addressed explicitly in advance. In particular, for multiple comparisons, "adjustment should always be considered and the details of any adjustment procedure or an explanation of why adjustment is not thought to be necessary should be set out in the analysis plan." Furthermore, since September 2007, in the so-called FDAAA 801 Requirement, any clinical trial of drugs or medical interventions that will seek FDA approval must be registered when the trial begins (see: www.clinicaltrials.gov), hence limiting or avoiding completely the reporting of unplanned analyses. Why are the strict guidelines not universally adopted in science in general? Imagine asking scientists to register all the hypotheses and analysis plans – including interim analyses in advance.

Example 1: Cancer drug study (continued). In the cancer drug study, suppose we know nothing about the variables prior to the study, so for us all these variables assume equal status. Then it would be naive to take the raw p-values seriously, thus forcing us to accept the adjustment and the overall null result. But suppose prior to collecting the data, because this is a study on end-stage metastatic patients, we declared that the Karnofsky index, a generalized measure of functional performance, was of primary interest, while the other variables were of secondary interest. Then, no adjustment would be necessary. Thus, as accepted in clinical trials, in assessment of evidence our intention matters. This does not feel controversial. We note that, for justifying the decision not to adjust, the primary interest in the index does not require prior knowledge or data or anything sensible; in principle we only need to declare it in advance.

Now suppose we accept we know nothing prior to the study, hence the multiplicity adjustment, but another research group who is working on the Karnofsky index contacts us to share the data. We agree to give them that variable only. Now, for them, it seems reasonable that no adjustment is needed and to conclude that the Karnofsky index is statistically significant. So, here we have a seeming paradox that, with the same data, two research groups can claim different evidence: one group cannot claim significance, but the other can. But what if we contact them? Now it seems the adjustment must be used, since we could have contacted any group working with the most significant variable. This means the mechanism of contact becomes an issue, but in practice the appearance of the 'second group' in the scene can of course be completely haphazard, e.g. via a chance encounter at a party, a friend of a friend, etc. In this social contact, how do we keep track of who brings up the topic first? It would be impossible to formalize such a process. ▫

Example 2: Single test. Surprisingly, the logical issue associated with the application of multiplicity adjustment arises even when we only perform a single test (Berger and Berry, 1987; Pawitan, 2001, p.202). Suppose a scientist comes to a statistician with a study (say with sample n=100), and the statistician performs a single test and obtains z= 2.1 (p-value=0.036). It seems uncontroversial to claim significance at α=0.05. Yes, but wait … what did the scientist plan to do were the result not significant? Suppose he planned to collect more data. So, his actual procedure is as follows:

- Collect n=100 and test if $|z_1| > c$, if significant stop.
- Otherwise, collect 100 more, and test if the overall $|z_2| > c$.
- To get an overall significance level α=0.05, we must use c=2.18, such that

$$\text{Prob}(|z_1| > c) + \text{Prob}(|z_1| \le c \text{ and } |z_2| > c) = 0.05.$$

So, the observed z=2.1 is actually not significant! As before, the statistical significance is affected not just by the data, but also by the intention of the scientist, but in this case it feels disturbing because the second stage of the study is still only a thought. (Before the reader accuses this example as perverse, adjusting for the intention here follows the FDA guideline on interim analysis of clinical trials.) ▫

Table 1. p-values in a study of a metastatic cancer drug vs placebo for ten patient characteristics.

Variables	p-value	10*p-value
1 Karnofsky index	0.007	0.07
2 Body weight	0.013	0.13
3 Tricep skin-fold	0.091	0.91
4 Hemoglobin concentration	0.236	2.36
5 Erythr sedimentation rate	0.350	3.50
6 Albumin in serum	0.525	5.25
7 Creatinine in serum	0.535	5.35
8 Bilirubin in serum	0.662	6.62
9 S-alkaline phosphatase	0.823	8.23
10 Alanine aminotransferase	0.908	9.08

These examples highlight a generic logical question: to what collection of tests do we need to apply the adjustment? To cover all tests relating to the same biological process? All tests in a single paper? If the latter, then theoretically we can avoid multiplicity correction by splitting the results into separate papers. The primary problem is that we are using the same statistic (p-value) both as a measure of evidence in a specific dataset (statistical distance between the hypothesis and the data) and as a measure of uncertainty (decision-making error rates) over hypothetical repetitions of the study. In the former adjustment for multiplicity is not an issue, but in the latter it is. However, the latter also requires a precise setup of how the study repetitions are to be done, and here intention matters, for example, in deciding which tests are to be included or prioritized. The examples show that any test can legitimately belong to distinct collections with distinct repetition studies that depend on the perspective of the experimenters. (This is not strange, e.g. a person may belong to distinct clubs with conflicting rules.) In Example 2, on seeing the observed data, the statistician's immediate reaction is to imagine hypothetical repetitions involving one test, but the scientist's intention involves hypothetical repetitions with two tests. Whose perspective is correct?

In view of this logical difficulty, how are we to react to potentially conflicting conclusions from unadjusted and adjusted analyses? We are put into this corner by an implicit demand that we make a decision. When a study is only performed once, i.e. the hypothetical repetitions remain hypothetical, we are in a never-ending unsolvable logical puzzle on how to decide. How do we break this puzzle? In clinical trials, for the key hypothesis, we break it by decree (e.g. FDA Guidelines), in essence limiting or avoiding the issue by stating the hypothesis and analysis methods in advance. Science does not follow such strict rules, but we need replication or validation studies to confirm interesting discoveries. Before further confirmation, discoveries are considered provisional and, in contrast to clinical trials, it is not necessary to make a decision about the true state of nature. However, we also note that, to confirm a discovery, it is not necessary to perform exactly the same experiment as before (i.e. the hypothetical repetitions). For example, a discovery in an observational study in humans will be substantially more credible if validated biologically in mice, and vice versa.

Eventually, in treating a study as a screening tool to identify interesting discoveries, we have to go back to the basic trade-off between type-I (false positive) and type-II (false negative) errors: protecting against one will increase the other. Strict adherence to multiple-testing adjustment protects against inflation of type-I errors and increases type-II errors, but who decides which error is more important? Different areas of science may evolve differently depending on their experience with these two errors. For example, molecular epidemiology went through the lamented candidate-gene approach to complex diseases, roughly from 1980s up to early 2000s (Hirshhorn et al, 2002; Chabris et al, 2012), where few findings were replicated.

Example 3: Publication bias and the Winner's curse. The human genome is a rich source of variables/genes. For many complex phenotypes, suppose there is no real effect or, more likely, the individual-gene effects are so tiny that the power of fundable studies is too small to detect anything. Say there are 100 research groups investigating different genes and phenotypes; each group is essentially generating random numbers. At 5% level, there will be 5 lucky groups with significant results: these are much more likely to get published, and then fail to replicate. If these are really 100 distinct groups, what can we do about this problem? No system of publication now can communicate so many negative findings to balance the false positives, so the problem seems to be an inevitable price of the scientific process.▫

Since high-throughput molecular studies, particularly the genome-wide association studies

References

Berger, JO and Berry, D (1987). The relevance of stopping rules in statistical inference. In Statistical Decision Theory and Related Topics IV (S Gupta and J Berger, Eds). New York: Springer-Verlag.

Chabris CF, et al. (2012). Most reported genetic associations with general intelligence are probably false positives. Psychological Science, 23(11), 1314-1323.

Goeman J and Solari A (2011). Multiple testing for exploratory research, with discussion. Statistical Science, 26 (4), 584-597.

Hirschhorn JN, et al. (2002). A comprehensive review of genetic association studies. Genetics in Medicine 4(2), 45-61.

Lee W, et al. (2012). Estimating the number of true discoveries in genome-wide association studies. Statistics in Medicine, 31(11-12), 1177-1189

Pawitan Y (2001). In All Likelihood. Oxford: Oxford University Press.

Rothman, KJ (1990). No Adjustments are Needed for Multiple Comparisons. Epidemiology, 1 (1), 43-46.

(GWAS), came into the scene, a single study/paper routinely performs millions of tests of single nucleotide polymorphisms (SNPs), and the genome-wide significance-level based on the Bonferroni correction is the accepted method of dealing with the huge multiplicity problem. One may argue correctly that we would be missing a lot of signals (type-II errors), but since the field had seen a lot of wasted effort at replicating false leads during the candidate-gene era, the consensus on the use of Bonferroni correction seems unchallenged. The problem in molecular epidemiology in the genomic era is that there are simply too many potential leads, so one needs a method to limit them. Note, however, that the p-values for each phenotype are usually adjusted separately, e.g. if there are one million SNPs, then the 5% significance level is divided by one million for each phenotype, regardless of the number of phenotypes, so the multiplicity adjustment is not followed consistently.

Multiple testing ideas are useful during an exploratory phase of a study. Goeman and Solari (2011) identified three sensible requirements for a multiple correction procedure: (i) not too strict, i.e. should allow possibilities of false positives; (ii) post-hoc, i.e. should allow choice after seeing the data; (iii) flexible, i.e. should allow whatever results to pursue, not just the significant ones. Instead of focusing on the probability of making any false positives, which is too strict, we can instead estimate and provide confidence lower-bounds for the number of true discoveries (=true positives). To emphasize its post-hoc feature, they called their procedure 'cherry picking'. Recent developments, for example in false discovery rate (FDR) estimation, are also in line with these requirements (e.g. Lee et al, 2012). The purpose is more to set realistic expectations for further studies rather than making final decisions about the true state of nature.

Example 1: Cancer drug study (continued). Suppose this study was performed on end-stage cachexic patients, which are characterized by severe wasting/loss of body mass and functional impairment. The top three variables in Table 1 are then of special interest, but this interest can be decided post-hoc (after seeing the data). From the third column in Table 1(10*p-value), the estimated number of true discoveries is 3-0.91≈2 (Lee et al, 2012). An application of the cherry-picking procedure (Goeman and Solari, 2011) gives 95% guarantee of at least 1 true difference from the top three findings, and 75% guarantee of at least one true difference from the second and third tests. So, even if we do not want to specify any hypothesis in advance, the analysis indicates the drug is worth studying further. ▫

To conclude, we do not mean to sound skeptical of the use of multiplicity adjustment, but only to emphasize the nuances when it is applied to real studies. We highlight some logical problems with its formal use when assessing scientific evidence, thus partly explain the lack of universal acceptance. There are areas that must pay close attention to the multiplicity adjustment, for example clinical trials, where type-I errors can have enormous human or financial costs, or high-throughput molecular studies with a flood of false positives when traditional significance levels are used. On the other hand, in areas of science where potential leads for discoveries are not abundant, formal multiplicity adjustment is not the norm; findings are considered provisional until further confirmation, and the inherently skeptical scientists pragmatically accept the trial-and-error aspect of the scientific process.

YUDI PAWITAN AND ARVID SJÖLANDER
KAROLINSKA INSTITUTET

HÄNT OCH HÖRT

Disputationer

2014

14 februari **SULTANA NASRIN** *Statistical aspects of modeling panel data originating from smallholder household surveys in Sub-Saharan Africa and Vietnam*, statistik, Lunds universitet. Opponent: Anders Nordgaard, Linköpings universitet.

2013

5 december **ANNIKA TILLANDER** *Classification models for high-dimensional data with sparsity patterns*, statistik, Stockholms universitet. Opponent: Patrik Rydén, Umeå universitet.

29 november **RASHID MANSOOR** *Assessing Distributional Properties of High Dimensional Data*, statistik, Internationella Handelshögskolan i Jönköping. Opponent: Helgi Tomason, University of Iceland.

29 november **VERA LISOVSKAJA** *A Contribution to the Design and Analysis of Phase III Clinical Trials*, matematisk statistik, Chalmers. Opponent: Martin Posch, Medical University of Vienna.

8 november **MARTIN SOLBERGER** *Likelihood-Based Tests for Common and Idiosyncratic Unit Roots in the Exact Factor Model*, statistik, Uppsala universitet. Opponent: Jean-Pierre Urbain, Universiteit Maastricht.

Licentiatseminarier 2013

19 december **BJÖRN LÖFDAHL** *Stochastic modelling in disability insurance*, tillämpad matematik och beräkningsmatematik, KTH. Diskutant: Mogens Steffensen, Københavns Universitet.

17 december **THORBJÖRN GUDMUNDSSON** *Markov chain Monte Carlo for rare-event simulation in heavy-tailed settings*, tillämpad matematik och beräkningsmatematik, KTH. Diskutant: Ola Hössjer, Stockholms universitet.

12 september **JOHN KARLSSON** *A class of infinite dimensional stochastic processes with unbounded diffusion*, matematisk statistik, Linköpings universitet. Diskutant: Christian Bayer, Weierstrass Institute.

Årets statistikfrämjare: Jo Røislien

■ Jo Røislien utnäms till Årets statistikfrämjare för att i det populärvetenskapliga programmet Siffror (SVT1/Utbildningsradion), genom ett underhållande format och en genomtänkt kommunikationsstrategi, ha gjort statistik intressant och förståeligt för nya målgrupper.

Statistik och medicin

De sista åren har jag lagt väldigt mycket tid på kommunikation i det offentliga rummet. TV är ett fantastiskt starkt medium. Syns man på tv blir man väldigt snabbt ”den där killen på tv”. Jag kan inte gå utanför dörren hemma utan att någon ber om autograf. När småflickor ser mig kommer de fram och berättar hur duktiga de är på sannolikheter.

Jag har numera hemligt telefonnummer – och jag är statistiker!

Jag utbildade mig till civilingenjör. Vill man bli civilingenjör i Norge ska man ta sig till Trondheim. Vädret i Trondheim var en katastrof. I Lillehammer skiner solen jämnt, i Trondheim regnar det och är kallt och eländigt. Så jag pratade med min morsa varje helg om hur miserabelt livet var.

En söndag fick jag det där telefonsamtalet man är rädd för att få. Min morsa ringde, men inte för att diskutera vädret, men för att berättade att min onkel fått hjärnblödning och låg inne på Oslo Universitetssjukhus. Han hade tagits dit med helikopter och efter flera timmars operation hade man fått stopp på blödningen.

En medicinsk succé – på nåt sätt. Men efter en stor hjärnblödning måste man, om man överlever, ofta lära sig saker på nytt. Lära sig att äta, lära sig att prata, lära sig att gå. Men när är träningen av gången färdig? Den är klar när man går normalt.

Men vad betyder det, att gå ”normalt”? Alla går ju på olika sätt.

För att besvara den frågan måste man först mäta ”gång”.

Hur mäter man det?

Man sätter markörer på benen; höfter, lår, knä, anklar, och personen får gå fram och tillbaka, och man mäter markörernas läge och beräknar de olika vinklarna och placeringarna av lederna under ett steg. Man ritar en kurva över hur en markör i bäckenet rör sig över tid och så vidare. Sjukhuset som min onkel låg på när han skulle lära sig att gå behövde hjälp med att räkna på den här typen av kurvor. De hade ingen aning.

Sådana här gångkurvor är korta kurvor som upprepar sig varje steg. Matematikern Fourier studerade den här typen av matematik på början av 1800-talet. Han var besatt av forskning på värmeöverföring. Det var han som först kom med hypotesen om jordens växthuseffekt, utan vilken vi inte skulle kunna leva på jorden. Det skulle bli för kallt.

Man hade redan på Fouriers tid förstått hur de matematiska ekvationerna för värmeöverföring ser ut. Men man kunde inte lösa dem, annat än för väldigt enkla fall – utan nån praktisk nytta. Fourier kom på en av de smartaste idéerna i den matematiska historien; den att en repeterande funktion, vilken som helst, kan approximeras med en summa av viktade sinus- och kosinusfunktioner. Den approximationen kan göras i princip hur noggrann som helst.

Om man går in i en musikstudio idag, då ser man inte en enorm massa paneler och maskiner. Där finns en dator och ett tangentbord. På grund av den matematiska utveckling som Fourier inledde kan de ljudvågor som vi uppfattar som musik fångas i ekvationer och lagras i filer så små att de går att ladda ner till ens telefon.

Den här typen matematik kunde vi använda för att testa om detta sätt att mäta gång på fak- »»»

tiskt fungerade. Jag tyckte det var jättespännande. Men läkarna, som inte var lika vana vid matematik, fick lite mer av matematiska andningsproblem. Men jag hade Fourier i ryggen. Sju mer eller mindre frivilliga studenter fick gå i timtal fram och tillbaka med markörer på benen i ett rum och bli filmade. Om markörerna ska fungera bra ska deras rörelse inte bero på annat än hur man går; inte vem som placerade markörerna. Med det gjorde de. Det blev skillnad om olika personer satte på dem på samma frivilliga. Vi skickade tillbaka dem till den firma i Oxford som gjort systemet, som fick förbättra sitt system och vi kunde använda markörerna med bättre och bättre reliabilitet.

King Kong ser väldigt levande ut i Hollywoodfilmen som kom för några år sen. Det beror på att han är väldigt levande. Man har filmat markörer på olika primater och sedan i animeringen av King Kong utnyttjat data från levande djur.

Det lilla sjukhuset ute i Oslofjorden blev en världsledare på rehabilitering. Och jag ville använda de metoder jag nu lärt mig till mer.

Jag och min hustru flytte till San Francisco. Hon skulle vara vid Berkeley-universitet ett tag, så jag tog med papprena om funktionalanalys och gångmarkörerna. Så blev hon gravid. Det speciella med en graviditet är att man bara ser utsidan, men inne i kvinnan pågår en otroligt komplicerad process när ett foster växer och blir ett barn. En graviditet är en väldigt konstig upplevelse att observera. Vi statistiker älskar att diskutera hur viktig del av ett fenomen variation är. Men när det är ens eget barn som ska bli till, då vill man inte se extrema observationer, utan då är det medelvärdet man vill ha. ”Normalt”.

Men hur kan man veta att allt är normalt? Man kan ju inte ta ultraljud flera gånger i veckan. Man behöver ett enkelt sätt att följa hur graviditeten fortlöper. Ett sånt är att varje dag räkna och bokföra hur många minuter det tar att känna tio sparkar. Om det tar längre och längre tid att känna tio sparkar kan det bero på att barnets aktivitet avtar. Som ju kan vara fordi nåt är fel. I USA kan man köpa kit som heter ”count to save a life”. Det innehåller en påse och ett stoppur.

Men fungerar det? Hjälper det att räkna sparkar?

I Norge gjorde man en kjempestor studie och samlade in sparkräkningsdata från ett tusental gravida kvinnor. Data hamnade till slut hos mig, och man frågade: Vad säger alla dessa data för nåt?

När jag tog bort naturlig dag-till-dag-variaton, som att bort knastret i en gammal radio, upptäckte jag att det här också såg ut som korta ljudvågor. Fourier. Gånganalys. Men min statistiska analys avslöjade att metoden att räkna minuter och sekunder tills den tionde sparken hade inget prognosvärde. Men det fanns ett annat värde med att kvinnor räknar sparkar: de lär känna sin kropp bättre.

För mig var det succé men inte för de som trodde det här var ett bra sätt att monitorera foster. Några hävdade att vi hade räknat fel, men det hade vi självklart inte.

Så bytte jag uppdragsgivare, till Norska Luftambulansen. Det är ofta de som är först på plats vid en olycka. När ambulansen kommer fram till en trafikkolycka är en av de första sakerna att ta itu med att stoppa blödningar. Och för en dryg vecka sedan fick jag data på vad som händer med blodet när man har blödningar. Det finns olika mediciner som ska hjälpa till att stoppa blödningar. Och se, det är åter funktioner som kan approximeras med Fourierserier. De sa att problemet att förstå de här egenskaper med blödningarna varit olöst i flera årtionden. Jag tänkte lösa det i helgen.

Det är så, att vi som kan statistik, vi kan något annat än medicinarna. För de är gånganalys en specialdisciplin. Graviditet en annan box. Blodflöde en tredje box. Men jag är statistiker så jag ser på världen i en annan ledd. Jag ser på det från sidan. Jag ser den statistiska metoden. Så för mig är det samma problem, de bara kallas för olika saker.

Pingviner och andra djur ser väldigt verkliga ut i animerade filmer. Det beror på att man har filmat markörer på olika djur, analyserat data och sedan överfört statistiken i animeringen.

Folk frågar mig: vad ska vi egentligen med matematik och statistik till?

Då vet jag inte riktigt vad jag ska svara – för mig är det så uppenbart vad vi ska använda matematik och statistik till. Vi ska använda de till att ge människor ett bättre liv. Vi ska rädda liv. Vad är det som har revolutionerat medicin och hälsa de senaste 50 åren? Den medicinska revolutionen är inte medicinsk, det är en matematisk och statistisk revolution.

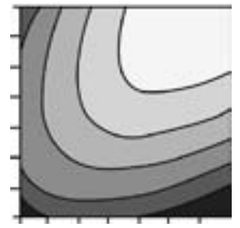
DAN HEDLIN

Knep som fångar publiken

■ Jo Røislien gav oss ett råd som han själv fått från de mediemänniskor han samarbetat med i flera år under inspelningen av tv-serien Siffror: tänk på att folk vill lära sig. De vill veta mer. Men de flesta är inte så bra på matte, de är inte vana att läsa böcker om matte och de gillar kanske inte att någon talar om för dem hur saker ting är. Men då faller ju hela vårt vanliga akademiska sätt att undervisa på.

Jo Røislien berättade hur han lagt upp det föredrag som vi just hade lyssnat på. När han berättar om sin egen forskning i populärvetenskapliga sammanhang, då visar han på stora förebilder (Fourier) och han väver in forskningen i sitt personliga liv, sina egna känslor. Så han gör allt fel, sett ur en traditionell akademisk synvinkel. Men det har visat sig att folk vill lyssna på det.

Ett annat knep fick vi veta i ett tidigare nummer av Qvintensen. Stort hår kan man åstadkomma med saltvattensprej.



Sex Sigmas historia

I förra numret av Qvintensen, nr 1/2014, redogjorde jag för statistikens användning inom Sex Sigma och diskuterade utbildningsaspekter av detta förbättringskoncept.

I litteraturen finns många uppgifter om detta framgångsrika koncept samtidigt som frågor ställs om orsakerna till framgången.

Jag vill här informera om några begrepp som Sex Sigma-konceptet bygger på, men som också kan ifrågasättas. Min hypotes är att det huvudsakligen är statistisk metodik som ger Sex Sigma framgångar, i kombination med ett projektorienterat arbetssätt för att uppnå förbättringar.

Sex Sigma är populärt och omtalat

Sedan introduktionen på nittiotalet är Sex Sigma fortfarande mycket omtalat. Tusentals böcker har skrivits i ämnet, och ca tio miljoner artiklar i olika tidskrifter. Termen ”Sex Sigma” ger idag 13 000 träffar på svenska Google och den engelska termen Six Sigma ger ca 16 miljoner träffar. Idag används Sex Sigma, samt Lean Sex Sigma, av bland andra kommuner, landsting, myndigheter, banker, sjukhus och industri. Konceptet utgår oftast från lönsamhetskalkyler och garanterar vinster, vilket tilltalar många företagsledare. Vi kan också se att Sex Sigma-konceptet följs med stort intresse i olika kretsar, inte bara av en mängd anhängare utan också av opponenter och kritiker.

Den vedertagna definitionen av Sex Sigma framhåller oftast att det är en metodik, vision, symbol, mål, ledningsfilosofi, vetenskaplig metodik, doktrin, affärsstrategi och arbetssätt. Andra grupper ser Sex Sigma som en attraktiv marknadsföringsstrategi, image, ett sätt att ratificera utbildning, rekryteringsstrategi och kompetensutveckling.

Sex Sigma-metodiken används främst för att effektivisera verksamhetsprocesser och är till skillnad från ISO-standarder och revisioner ett betydligt effektivare och mer jordnära sätt att uppnå företagets mål och åstadkomma förbättringar. Detta med flera andra egenskaper gör Sex Sigma-metodiken till ett populärt koncept.

Hur skapades Sex Sigma?

Men Sex Sigma tycks betyda olika saker för olika grupper av användare. Sex Sigmas tillblivelse är inte klarlagd men symbolen har registrerats som Motorolas varumärke. På åttiotalet använde Motorola egna diagram för processtyrning med färgmarkerade band. Ett grönt fält visade bandet med 50 % av specifikationsgränserna kring målet och gult band markerade området mellan det gröna bandet och specifikationsgränserna. Området utanför var rödmarkerat. Detta baserades på en processduglighet med kapabilitetsindexet Cpk=1,0, men medförde samtidigt alltför tillåtande toleranser, vilket ledde till defekter (Magnus Arnér och Kerstin Vännman skrev om detta index och andra kapabilitetsindex i Qvintensen nr 2/2011).

»Kanske denna spalt ger intryck av skepticism mot Sex Sigma-konceptet, men denna riktas dock främst mot de gamla, mindre nyttiga, begrepp som Sex Sigma ursprungligen introducerade.»

En av Motorolas ingenjörer, Bill Smith, införde då krav på variationen kring processens mål på högst ±1,5 sigma, där sigma är processens standardavvikelse. Emellertid är 1,5 ett värde med svag teoretisk grund. Man gjorde det felaktiga antagandet att data från alla processer var normalfördelade. Detta förekommer sällan och är inte den sanna bilden av industriella processer och transaktionsprocesser. Trots dessa felaktigheter utvecklades Sex Sigma vidare, framtvingat av Motorolas behov av kvalitetsförbättring.

Med facit i hand kan vi säga att det hade varit bättre om man tillämpat statistisk processtyrning från början för att påverka processkapabiliteten med hjälp av styrande processparametrar och genom eliminering av processfel, istället för att manipulera med processpecifikationer.

1995 blev Sex Sigma känt som begrepp när General Electrics ordförande, Jack Welch, gjorde det till GE:s framgångsrika affärsstrategi.

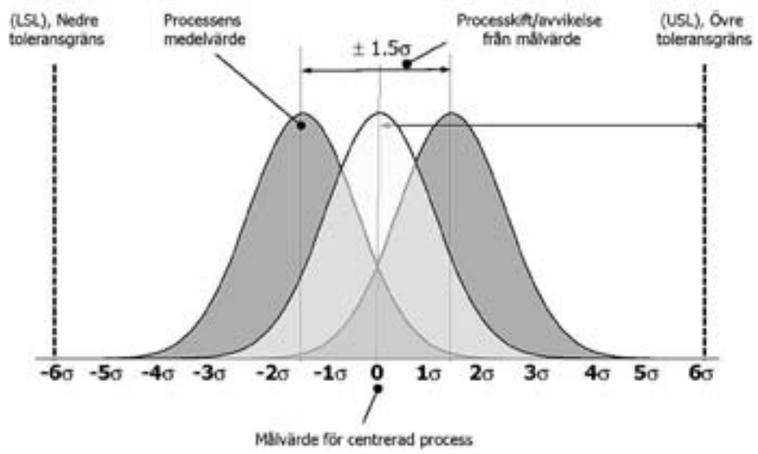
Namnet Sex Sigma

Själva namnet Sex Sigma härstammar från den praxis för toleranssättning som Motorolas ingenjörer använde. Där användes först gränserna ±1 sigma som processtyrgränser och de ändrades därefter till ±3 sigma. Eftersom detta fortfarande resulterade i felaktiga produkter ändrade man ±3 sigma till ±6 sigma för att höja processkapabiliteten från Cpk=1 till Cpk=2. Även om man menade väl så är det anmärkningsvärt och förvånande att man höjde processkapabiliteten utan att ändra på processkvaliteten. För en statistiker kan man beskriva det som att kraven på att ha liten variation i processen höjdes, men processens väntevärde lämnades därhän. Ännu märkligare är att just detta ledde till att Sex Sigma blev symbolen för kvalitet i världsklass.

Att räkna defekter och avkastning

Inom Sex Sigma används ytterligare några begrepp som till exempel defekt, DPU - defekt per enhet, defektmöjlighet, DPMO - defekter per miljon möjligheter (synonymt med PPM, parts per million), FTY – First Time Yield och RTY – Rolled Throughput Yield.

Skift – avvikelse från processens målvärde



Defekt är en produktegenskap som inte stämmer med specifikationen eller förväntningarna och som inte fungerar som den borde göra i rätt utförande. Här kan anmärkas att en produkt kan ha flera defekter och en defekt kan vara orsakad av en eller flera feltyper som har förekommit eller införts under arbetsgången i en process.

DPU är totalt antal defekter dividerat med totalt antal enheter i populationen. Man skattar DPU genom att observera det totala antalet defekter vid provtagning från en population. Det är ett föråldrat sätt att redovisa produktkvalitet och återspeglar snarare ett intensitetsmått. Rådata är mera användbart för att till exempel peka ut var i processen felen uppstår.

Defektmöjlighet är en möjlighet till fel på produkten. Denna beräkning av felmöjligheter är ett ytterst subjektiv sätt att använda proportionsräkning och kan vara mycket missvisande som kvalitetsmått. Istället kan vi använda begreppet sannolikheten att uppnå en viss kvalitetsnivå eller utbyte, till exempel max antal fel per antagen produktkomplexitet.

DPMO är antalet defekter per enhet som observerats på produkten under en produktionsomgång dividerat med antal felmöjligheter, normaliserad till en miljon felmöjligheter.

FTY anger processavkastning från ett processteg genom att dividera antalet felfria enheter med det totala antalet enheter som behandlas av processen. Vi antar att 100 enheter kommer in i processen A och bara 90 produceras felfria. 10 är felaktiga och kasseras. Avkastningen FTY_A blir 90%.

Processutbytet RTY anger total processavkastning som produkt av samtliga ingående processteg FTY. Om man antar att processerna är oberoende av varandra får vi som ett exempel med fyra processteg ett RTY lika med $FTY_A * FTY_B * FTY_C * FTY_D$, eller RTY=0,9 * 0,89 * 0,94 * 0,98 = 0,74.

RTY kan också beräknas genom att dividera antalet godkända enheter med det totala antalet som har bearbetats av hela processen. I detta fall 74/100 = 0,74 eller 74% utbyte.

Både FTY och RTY används fortfarande inom Sex Sigma för att beräkna avkastningen från en process. FTY-beräkningen påverkas mest av antalet enheter, vilket betyder att den inte tar hänsyn till om utgångsenheten är defekt. FTY fångar inte hur många defekter som hittas och korrigeras inom en och samma enhet eftersom man bara räknar antal defekta enheter och inte antal fel per enhet. Här missar man mycket viktig information genom att inte beräkna processkapabilitet ur processdata.

RTY, kapacitet av en s.k. ”genomlöpsavkastning”, återspeglar sannolikheten för att klara processens kriterier för varje steg i en process. Nackdelen med den beräkningen är att RTY inte fångar kostnader för till exempel omarbete, s.k. ”dolda fabriker”. Tillgång till den informationen är nödvändig för att kunna höja utbytet inom hela processen och därmed minska kostnaderna.

Bättre resultat kan åstadkommas genom användning av statistiska metoder som statistisk processtyrning, kapabilitetsanalys, försöksplanering och olika former av regressionsanalys.

Den största fördelen

Den största fördelen med Sex Sigma-konceptet är att det bygger på och populariserar användningen av mätningar och statistiska metoder. Den stora nackdelen ligger främst i ett föråldrat räknesätt som vilar på tämligen osäkra grunder. Hit hör Sex Sigmas ogrundade koppling till normalfördelningen och otydliga definitioner av processtyrning samt beräkningar av processutbyte. I den industriella verkligheten är normalfördelningen ytterst sällsynt förekommande och gamla beräkningar av processens utbyte grundar sig på overifierade antaganden om processförskjutning på ±1,5 sigma som helt saknar vetenskaplig grund. Men trots dessa brister så balanseras Sex Sigma-bilden av de positiva effekter som till exempel DMAIC-metodiken har fört med sig och mängden av nyttig statistisk metodik som används inom konceptet. (DMAIC-metodiken beskrevs i förra numret av Qvintensen, nr 1/2014).

Kanske denna spalt ger intryck av skepticism mot Sex Sigma-konceptet, men denna riktas dock främst mot de gamla, mindre nyttiga, begrepp som Sex Sigma ursprungligen introducerade.

Sex Sigma av idag representerar ett modernt förbättringskoncept med en mängd utprovade och ytterst effektiva statistiska metoder och verktyg. Enlig mångas erfarenhet är det statistisk metodik som bär Sex Sigma-konceptet till framgångar. Därför är det viktigt att slå vakt om en kvalitativt hög utbildning i Sex Sigma-konceptet så att goda kunskaper kan komma till nytta inom kommuner, landsting, myndigheter och industrin.

Romaren Cicero lär ha sagt att ”orsakerna till händelser är mera intressanta än händelserna själva”. Sex Sigma-konceptet levererar bevisligen goda resultat och det är glädjande att statistisk metodik möjliggör detta.

FOTO: LINA LANDE

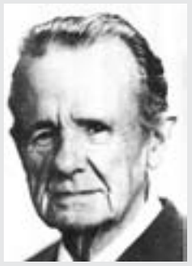
GÖRAN LANDE

KENNETH CARLING, NY ORDFÖRANDE I CRAMÉRSÄLLSKAPET

– Hedrande och spännande

– Ny ordförande i Cramérsällskapet – hedrande och spännande. Hedrande för att Harald Cramér var en lysande stjärna i den svenska statistikhistorien. Som doktorand i Uppsala i början av 90-talet hittade jag hans pedagogiska bok *Sannolikhetskalkylen och några av dess användningar* övergiven och dammig i en bokhylla i ett rum som stod outnyttjat. Jag satte genast i gång att läsa den trots att jag var under stor tidspress – vi hade tenta i en forskarutbildningskurs där examinationen bestod av att lösa ett antal problem inom 72 timmar. Hur tentan gick minns jag inte, men däremot minns jag hur nöjsamt det var att läsa Cramérs bok bland annat för att det var första gången jag läste en statistikbok på svenska. All kurslitteratur i statistik var redan på den tiden på engelska. Efter läsningen sökte jag mer information om mannen bakom boken, vilket var mödosamt i en tid då internets fria och snabba informationsflöde var i sin linda. När jag och min fru skulle ge namnet åt vår andra son de första dagarna av detta årtusende, så föll valet på Harald. Inte på grund av alla danska och norska kungar som burit det namnet, utan just på grund av Harald Cramér.

– Spännande för att Cramérsällskapets syfte är att främja forskning och utbildning i statistik. För mig är statistik (eller mer allmänt induktion) människans främsta redskap för att förstå och förutse sin omgivning. Och det är i princip lika viktigt för oss människor som att äta. Förståelse är centralt för vår sinnesro och tillfredsställelsen av vår intellektuella nyfikenhet; förutsägelsen en förutsättning för vår fortsatta överlevnad. Att genom forskning förfina detta redskap och genom utbildning lära yngre generationer detta redskap, ja det måste vara mycket viktigare än att följa Melodifestivalen. Tycker jag i alla fall! Och möjligheterna att förstå och förutse vår omgivning är fantastiska idag – massor med data, billiga mätare och enorm beräkningskraft. Än roligare blir uppgiften när man beaktar att omgivningen blir mer och mer komplex och svårförutsägbar. Jag tycker min vardag handlar om denna uppgift, men formellt heter det att jag är professor i statistik vid Högskolan Dalarna och forskningsledare i Komplexa system – Mikrodataanalys med forskningskollegor från ämnen som, utöver statistik, data- och vägteknik, informatik, kulturgeografi, företags- och nationalekonomi. I juni anordnar Cramérsällskapet sommarskola i min hemstad Falun, ni som ska dit får gärna hänga med mig ut och springa på Lugnet för där håller jag gärna till springandes, skidandes eller cyklandes tidiga morgnar eller sena kvällar.



Cramérsällskapet har fått sitt namn efter Harald Cramér. I Qvintensen nr 2/2010 berättar Tomas Cramér om sin far Harald. Vi tackar Ingram Olkin för bilden

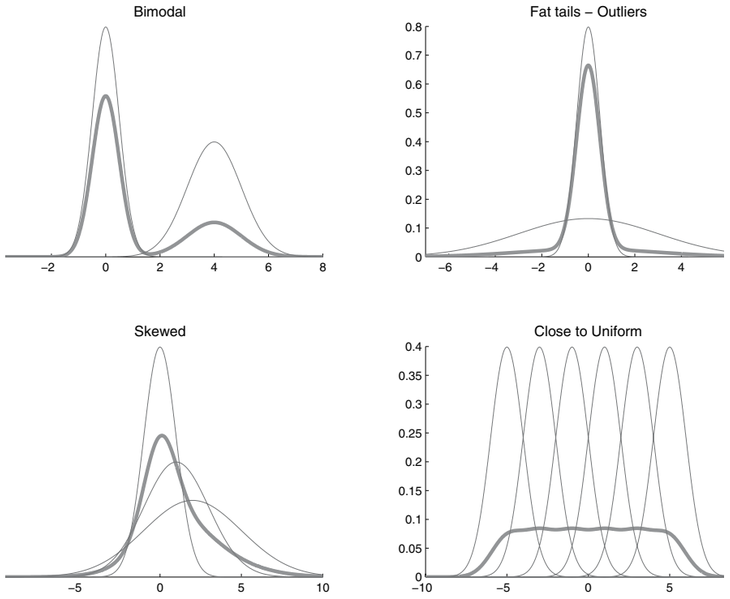
NY I STYRELSEN/ Martin Singull, ledamot

– Jag är lektor i matematisk statistik vid Linköpings universitet. Min grundutbildning, civilingenjör teknisk fysik, läste jag med inriktning mot matematik och statistik. Jag fastnade för statistiken och började mitt avhandlingsarbete vid Linköpings universitet som kom att handla om skattningar av kovariansmatriser med olika strukturer. Min forskning idag handlar fortfarande mycket om det men jag har också breddat mig och bihandleder fem doktorander inom olika områden, som till exempel tillväxtkurvor, random matrix theory, small area estimation och olika finanstillämpningar.

– Fram till nyligen var jag koordinator för forskarutbildningen vid matematiska institutionen. Jag följde bland annat upp forskarstudenternas årliga studieplaner och deras utveckling. Jag arbetade nära studierektorn och fick en bra inblick i hur forskarutbildningen fungerar. Nu fokuserar jag mer på min egen forskning, de fem doktorander som jag bihandleder, undervisning och LinStat2014, en konferens i linjär statistik som går av stapeln i Linköping i augusti och där jag är ordförande i organisationskommittén.

– Min tjänst erbjuder en hel del resor. Utöver konferenser åker jag regelbundet till Toronto där jag har ett forskningssamarbete med professor Muni S. Srivastava. Ett samarbete som påbörjades 2008 när jag var där i sex månader som doktorand. Jag försöker besöka University of Toronto åtminstone en gång per år. Ett annat resmål som jag återkommer till regelbundet är Rwanda där jag bidrar till ett SIDA-projekt som bland annat handlar om att utveckla ett masterprogram i statistik och vidareutveckla ett masterprogram i matematik samt forskarutbildningen vid University of Rwanda. Trots sin historia är Rwanda ett underbart land som ger många spännande möten med mycket trevliga och hjälpsamma människor.

– Min fritid tillbringar jag mycket tillsammans med min familj, fru och dotter som snart är två år, gärna vid en klätterklippa eller på cykel eller skidor. När barnpassning tillåter blir det gärna under lite mer äventyrliga former.



◀ **Figure 1. Using mixture of normal densities (thin lines) to approximate a density (bold line); four examples of rather difficult densities to model with other methods.**

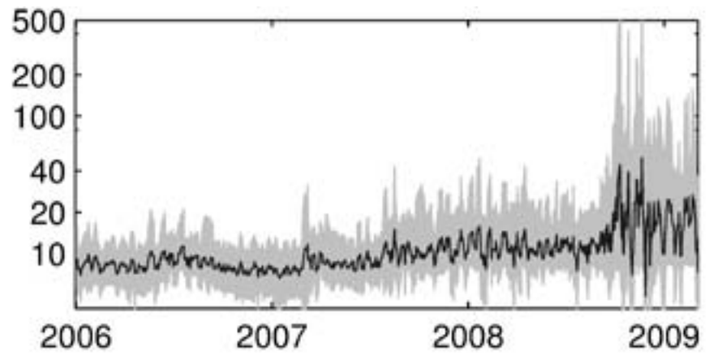


Figure 2. Time series plot of the posterior median (the dark line) and 95% probability intervals (the gray polygon) for kurtosis in terms of degrees of freedom of the return distribution for S&P 500 stock returns.

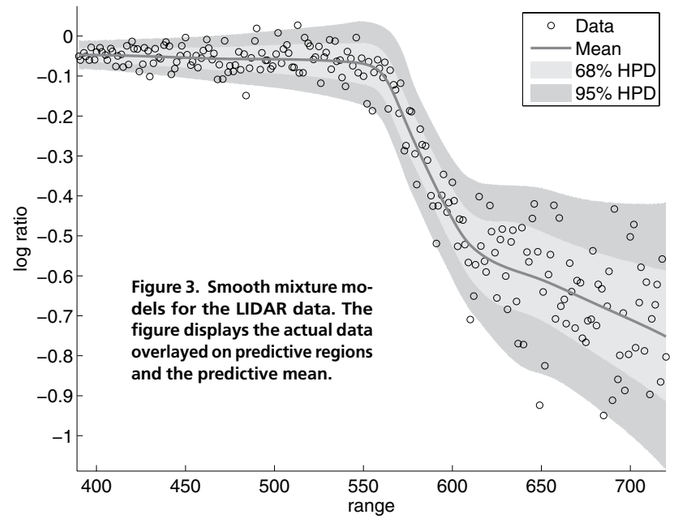


Figure 3. Smooth mixture models for the LIDAR data. The figure displays the actual data overlayed on predictive regions and the predictive mean.

Feng Li skrev sin doktorsavhandling på Statistiska institutionen vid Stockholms universitet. I år utnämndes han av Cramérsällskapet till årets Cramérpristagare.

Complex Models for Complex Data via the Bayesian Approach

Background

Statistical methods have developed rapidly in the past twenty years. One driving factor is that more and more complicated high-dimensional data require sophisticated data analysis methods. A noticeably successful case is the machine learning field which is now widely used in industry. Another reason is the dramatic advancements in the statistical computational environment. Computationally intensive methods that in the past could only be run on expensive super computers are now possible to run on a standard PC. This has created an enormous momentum for Bayesian analysis where complex models are typically analyzed with modern computer-intensive simulation methods.

Traditional linear models with Gaussian assumptions are challenged by the new large complicated datasets, which have in turn spurred interest in new approaches with flexible modeling with less restrictive assumptions. Moreover, research has shifted from merely modeling the mean and variance of the data to sophisticated modeling of skewness, tail-dependence, and outliers. However, such work demands efficient inference tools. The development of highly efficient Markov chain Monte Carlo (MCMC) methods has reduced the barrier. The Bayesian approach provides a natural way for prediction, model comparison and evaluation of complicated models, and has the additional advantage of being intimately connected with decision making.

The Bayesian density estimation

In statistics, density estimation is the procedure of estimating an unknown density $p(y)$ from observed data. Density estimation techniques trace back to the use of histograms, later followed by kernel density estimation in which the shape of the data is approximated through a kernel function with a smoothing parameter. However, kernel density estimation suffers from one obstacle, which is the necessary step of specifying the bandwidth.

Mixture models have recently become a popular alternative approach. A mixture density is a combination of different densities with different weights. Usually the mixture density is a weighted sum of densities. Mixture densities can be used to capture data characteristics such as multi-modality, fat tails, and skewness. Figure 1 shows four examples.

Conditional density estimation concentrates on the modeling of the relationship between a response y and set of covariates x through a conditional density function $p(y|x)$. In the simplest case, the homoscedastic Gaussian linear regression $y = x'\beta + \varepsilon$ is trivially equivalent to modeling $p(y|x)$ by a Gaussian density with mean function $\mu = x'\beta$ and constant variance.

In Bayesian statistics, inference of an unknown quantity θ, say $p(\theta|y)$, combines data information y, $p(y|\theta)$, with prior beliefs about θ, $p(\theta)$. In many simple statistical models with vague priors that play a minimal role in the posterior distribution, Bayesian inference draws similar conclusions to those obtained with a traditional frequentist approach. The Bayesian approach is however more easily extended to more complicated models using MCMC simulation techniques. In principle, MCMC can be applied to many hard-to-estimate models. However the efficiency depends heavily on how efficient the MCMC algorithm is. This is especially true in nonlinear models with many correlated parameters.

A key factor for evaluating a method's performance is to check how it balances the trade-off between goodness-of-fit and overfitting. It is common that if a model wins in goodness-of-fit, it will lose in prediction. Variable selection is a technique that is commonly used in such a context. Historically the purposes for using variable selection are to select meaningful covariates that contribute to the model, inhibit ill-behaved design matrices, and prevent model overfitting. Methods like backward and forward selection are standard routines in most statistical software packages. However, the drawbacks are obvious in those techniques; for example, the selection depends heavily on the starting points, which becomes more problematic with high dimensional data with many covariates. Most current methods rely on Bayesian variable selection via MCMC. A standard Bayesian variable selection approach is to augment the regression model with a variable selection indicator for each covariate. For the purpose of overcoming problems with overfitting, shrinkage estimation can also be used as an alternative, or even complementary, approach to variable selection. A shrinkage estimator shrinks the regression coefficients towards zero rather than eliminating the covariate completely. One

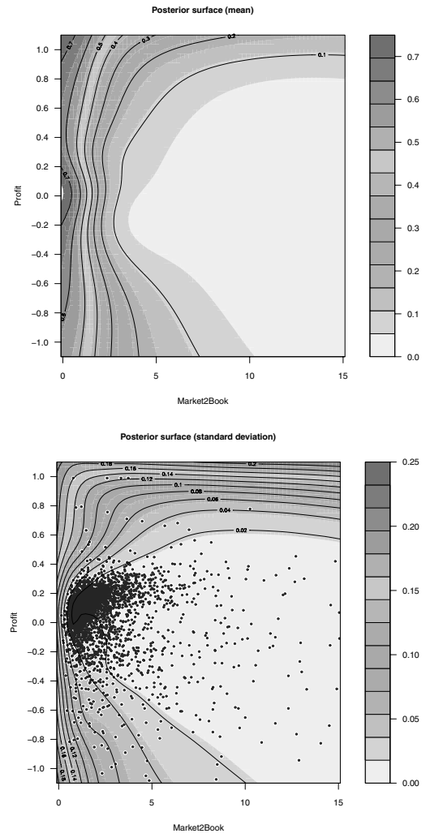


Figure 4. The posterior mean (upper panel) and standard deviation (lower panel) of the posterior surface for the model for firm leverage data. The shade indicates how the leverage varies with book values (x-axis) and profits (y-axis). The lower panel also shows an overlay of the covariate observations. Graferna finns i färg på tidningens baksida.

Patricia Geli Rolfhamre.

Anton Grafström.

Fredrik Jonsson.

Tidigare Cramér-pristagare

2013 David Bolin
2012 Fredrik Jonsson
2011 Anton Grafström
2010 Patricia Geli Rolfhamre

David Bolin.

way to select a proper value of the shrinkage is by cross-validation.

Bayesian models for complex data

Modeling the volatility and variability in financial data has been a highly active research area since the seminal paper by Engle introduced the ARCH model in 1982, and there are large financial markets for volatility-based instruments. Financial data, such as stock market returns, are typically heavy tailed and subject to volatility clustering, that is, time-varying variance. They also frequently show skewness and kurtosis that evolve in a very persistent fashion. Alternatively they may have been the result of a financial crisis with an unprecedented volatility. Figure 2 illustrates the modeling of the degree of freedom with S&P 500 returns. Modeling such data requires sophisticated MCMC treatment, but in return, we obtain better insights into a situation that other methods can hardly tackle.

LIDAR (Light Detection And Ranging) is a technique that uses laser-emitted light to detect chemical compounds in the atmosphere. The response variable (logratio) in a dataset we have analyzed consists of 221 observations on the log ratio of received light from two laser sources: one at the resonance frequency of the target compound, and the other from a frequency off this target frequency. The predictor is the distance traveled before the light is reflected back to its source (range). Our aim is to model the predictive density p(logratio | range). A smooth mixture of asymmetric densities is used to model such predictive density which involves a large number of parameters. See Figure 3 for the fitted curve with the confidence band. It is therefore easy to overfit the data unless model complexity is controlled effectively. Bayesian variable selection in all parameters can lead to important simplifications of the mixture components. Not only does this control complexity for a given number of components, but it also simplifies the existing components if an additional component is added to the model.

In finance applications, a firm's leverage (fraction of external financing) is usually modeled as a function of the proportion of fixed assets, the firm's market value in relation to its book value, firm sales, and profits. The relationships between leverage and the covariates are highly nonlinear. There are also outliers. Strong nonlinearities seem to be a quite general feature of balance sheet data, but only a handful articles have suggested using nonlinear/nonparametric models. One attempt is to extend the regression model by introducing a lot of auxiliary variables, also known as splines. A nonlinear curve/surface can then be constructed by choosing the correct number of splines and placing them in the right covariate space. See Figure 4 for the fitted mean curve and the standard deviation. Nonetheless, correctly allocating the splines in covariate space is not trivial. Bayesian methods treat the locations as unknown parameters that efficiently allocate the splines and therefore keep the number of splines to a minimum. Compared with the traditional deterministic spline approach, the Bayesian approach allows the splines to move freely in the covariate space and provides a dynamic surface with the measurement of surface uncertainty.

A model bigger than an elephant?

A linear regression model that was considered as very advanced in the 1950s is now standard course content for university students. The data are much more complicated nowadays, not only because the volume of data has increased but also the structure is now much more complicated. Very high-dimensional data that are a mix of numeric variables, character strings, images, or videos are not rare anymore. Sophisticated models are essential for such situations. In principle, a complicated model should be able to capture more complicated data features but estimating and interpreting such a model is not obvious. Personally speaking, there is a huge space to explore computationally and statistically. Statistical models that can adapt to modern computational architectures already flourish in industry. Techniques like high performance computing will be more widely used in statistics and will be made aware to statisticians eventually.

I would like to thank Professor Mattias Villani who introduced me to this exciting area.

FENG LI
CENTRAL UNIVERSITY OF FINANCE AND ECONOMICS, BEIJING, CHINA

Reference

Li, Feng (2013). Bayesian Modeling of Conditional Densities. Doctoral thesis. Department of Statistics, Stockholm University

ORDFÖRANDEN HAR ORDET

Vårkonferens i vår-vintrigt Stockholm

Det ett vårvintrigt Stockholm 20 mars på universitetets Frescatiområde hölls föreningens årsmöte med ett tjugotal deltagare. Mötet bestod som brukligt är av två delar: den informella med ett antal föredrag och den formella med årsmötespunkter såsom val av styrelse för det kommande verksamhetsåret (se nedan). Thomas Hvidtfeldt från Brottsförebyggande rådet inledde med att beskriva Nationella trygghetsundersökningen som är årlig och bygger på ett stort stickprov av personer (16-79 år) bosatta i Sverige. Undersökningens huvudsyfte är att beskriva befolkningens utsatthet för brott och dess upplevda trygghet. Peter Lundquist (SCB och Stockholms universitet) följde upp med att beskriva tekniken ”balanserad svarsmängd”. Det är ett sätt att justera för bortfall vid urvalsundersökningar i de fall där (bakgrunds-) hjälpinformation finns tillgänglig, genom att öka graden av representativitet med avseende på populationen.

”Kvalitet i webbpanelundersökningar” med underrubriken ”Metoder och mått” är namnet på den slutliga rapport som webbpanelkommittén arbetat fram och vars bakgrund och innehåll Gösta Forsman gjorde en genomgång av (se sidan 24 i detta nummer). I samband med detta kunde vi glädjande nog distribuera den direkt från tryckpressarna närmast rykande heta rapporten! Denna kommer också att delas ut i samband med kommande evenemang arrangerade av föreningen och även finnas tillgänglig elektroniskt. Den som redan nu vill ha pappersversionen kan kontakta Åke Wissing (ake@wissing.se).

Nästa talare blev årets uppsatsprisvinnare Andrea Gustafsen och Bergdís Sigurjónsdóttir, Stockholms universitet, med sin kandidatuppsats ”Coding Process and Quality Control”. De har gjort en studie för att kartlägga litteraturen angående svarskodning med tillhörande potentiella kodningsfel och hur olika statistikproducenter arbetar med kvalitetskontroll gällande kodning.

Carl-Erik Särndal avslutade föredragsdelen av mötet med det allvarligt syftande kåseriet ”Is that all there is” som en replik till en kritisk fundering av undertecknad i just den här ordförandespalten (Qvintensen nr 4/2013), angående vilka urvals- och skattningsmetoder som egentligen används vid till exempel opinionsundersökningar. Carl-Erik breddade också frågeställningen genom att skärskåda begreppen stickprov, slumpmässigt, representativt.

I vetenskaplig/akademisk mening är dessa väldefinierade men ute i ”verkligheten” blir de ibland urvattnade och oklara begrepp som ändå ofta nyttjas i till exempel marknadsföringssyfte.

Vid andra delen av årsmötet valdes inte mindre än sex nya styrelsemedlemmar och föreningens nya styrelse består nu av:

Per Gösta Andersson (omval, ordförande)
Anders Eklund (omval, sekreterare)
Marcus Berg (nyval, webbansvarig)
Anders Järleborg (nyval)
Jonas Ortman (nyval)
Karin Schneller (nyval)
Maria Varedian (nyval)
Åke Wissing (nyval)

De nyvalda presenterar sig själva i detta nummer. Den observante noterar att ingen är uppsatt som kassör. Denna fråga kommer dock att vara löst i samband med den nya styrelsens första möte. De avgående styrelsemedlemmarna Michael Carlson, Eva Lagerkrantz (kassör), Kadri Meister samt Maria Öman tackas för sin tid i styrelsen! Den nya hoppas kunna glädja föreningens medlemmar och övriga intressenter med lockande aktiviteter framöver!

PER GÖSTA ANDERSSON

NYA I STYRELSEN/

Maria Varedian, ledamot

– Jag jobbar på Trafikverket i Borlänge på enheten Statistikcenter. Vi tar själva fram en del statistik och ska även stödja verket med vår kompetens där det behövs. Det innebär att ganska skiftande frågor kan komma upp. Det mesta av min tid ägnar jag åt frågor som rör mätning av vägtrafikflöden och hastigheter. För det behövs urval av mätplatser och tidpunkter. Hur man bäst tar fram skattningar och hur man hanterar bortfall är också intressanta frågor. Dessutom behövs det beskrivningar av kvaliteten så att användarna förstår om data är lämpliga att använda för deras syften. Vi inventerar vilka krav användarna har för att kunna anpassa vad som produceras till behoven. Jag har länge arbetet med vägtrafikmätningar, först på Vägverket, då mer direkt med produktionen av data, och sedan under en period som konsult på Vectura. Men nu är jag alltså tillbaka på Trafikverket (som det heter efter sammanslagningen av Vägverket och Banverket). Man kan säga att jag bytt arbetsgivare oftare än arbetsuppgifter även om det blir lite olika infallsvinklar beroende på var man sitter.

Jonas Ortman, ledamot

– Jag läste vid Uppsala universitet och skrev magisteruppsats på ekonomprogrammet om tillvägagångssätt att utvärdera kvaliteten på tjänster företag emellan. Det är intressant med teorier om vad som formar individens upplevelse av kvalitet och olika metoder som utformats för att försöka mäta den upplevelsen - både kvantitativa och kvalitativa. Själv är jag framför allt intresserad av gränslandet däremellan, till exempel att man via webbenkäter låter respondenter svara på främst öppna frågor där de själva graderar och kategoriserar sina egna fritextsvar. Jag är också intresserad av (och har gjort några mindre undersökningar kring) respondenters syn på enkäter och vad de har för positiva och negativa associationer kring enkäter.

– Jag har arbetat med enkätundersökningar på Action Dialog Partner sedan 2008 där jag nu är vd. Vi jobbar framför allt med undersökningsföretag, konsulter och myndigheter där vi hjälper till med tekniska lösningar för webbenkäter och graf- och tabellframställning.

– Jag är uppvuxen i Vällingby. Nu bor jag i Sickla med frun, en femårig dotter och en treårig son. På den lilla tid som blir över spelar jag gärna brädspel med släkt och vänner, racketsport eller golf.

RAPPORT FRÅN

Webbpanelkommittén

Surveyföreningens webbpanelkommitté har under våren färdigställt en rapport, Kvalitet i webbpanelundersökningar. Metoder och mått. Den finns nu tillgänglig som pdf-fil på Surveyföreningens webbplats. Gå till http://statistikframjandet.se och välj Surveyföreningen.

Här följer en kort sammanfattning av arbetet.

Bakgrund till rapporten

Marknads- och opinionsundersökningar genomförs sedan sekelskiftet i ökande grad via webbpaneler. Då webbpanelundersökningar kännetecknas av andra möjligheter och problem än traditionella urvalsundersökningar har producenter och användare ställts inför svårigheter att bedöma kvaliteten i undersökningsresultaten. Traditionella mått, som konfidensintervall, är inte relevanta för att bedöma den totala osäkerheten i undersökningsresultat och det har saknats andra etablerade redskap för sådana bedömningar.

Surveyföreningens styrelse tillsatte därför 2009 en kommitté med uppgift att, via litteraturstudier och egna bedömningar, ta fram en skrift som belyser kvalitetsfrågor och presenterar mått som på olika sätt kan beskriva kvalitet i webbpanelundersökningar. Skriften vänder sig till aktörer som kommer i kontakt med webbpaneler: leverantörer, undersökningsköpare och andra användare av undersökningsresultat.

Rapportens syfte är att ge riktlinjer för kvalitetsredovisning av webbpanelundersökningar. I praktiken handlar det huvudsakligen om marknads- och opinionsundersökningar, som utförs av kommersiella undersökningsföretag. Inom detta område finns redan vissa internationella standarder, som är allmänt accepterade inom undersökningsbranschen. Två viktiga dokument på detta område är ISO-publikationen SS-ISO 26362:2009 Accesspaneler inom marknads-, opinions- och samhällsundersökningar – Vokabulär och servicekrav samt den internationella organisationen ESOMAR:s (European Society for Opinion and Market Research) skrift 28 questions to help buyers of online samples (ESOMAR 2012). Den rapport som nu föreligger avser att komplettera de nämnda dokumenten med riktlinjer speciellt inriktade mot kvalitetsredovisning (och speciellt tänkta för svenska förhållanden).

Webbpanelundersökningar

Med en webbpanel menas en databas som innehåller uppgifter om personer, ofta kallade panellister eller panelmedlemmar, som är villiga att medverka som respondenter i undersökningar. Den viktigaste uppgiften i databasen är personens e-postadress som inhämtas i samband med att denne rekryteras. Övrig information samlas in från panellisterna via en så kallad profilundersökning. Olika undersökningar, så kallade webbpanelundersökningar, genomförs i ett senare skede på hela panelen eller på ett urval av panellister där alltså panelen används som urvalsram.

Panelmedlemmar kan rekryteras på olika sätt:

• Självrekrytering. Allmänheten bjuds in via annonser i tidningar, pop-up-rutor på hemsidor, mm. Alla som ser annonsen är inbjudna att vara med i panelen.

• Rekrytering i samband med annan undersökning med sannolikhetsurval. En fråga om intresse att vara med i panelen ställs i samband med den ordinarie intervjun. Endast de som deltar i den första undersökningen inbjuds att vara med i panelen. Ofta är den första undersökningen en telefonbuss[2] eller en engångsundersökning med telefonintervju.

• Direkt rekrytering med sannolikhetsurval. En separat undersökning genomförs med enda syfte att rekrytera panellister.

• Blandat rekryteringsförfarande. Ofta sätts paneler samman med panellister som rekryterats med olika metoder. Det är dock vanligt att paneler rekryterade via sannolikhetsurval hålls separerade från självrekryterade paneler.

Vi konstaterar i rapporten att rekryteringsmetoden är mycket viktig för undersökningsresultatens kvalitet. Ju starkare inslag av sannolikhetsurval desto bättre utgångsläge. Direkt rekrytering med sannolikhetsurval är den bästa metoden medan självrekrytering leder till stora svårigheter att bedöma kvaliteten i undersökningsresultaten. I praktiken har man dock problem även vid rekrytering med sannolikhetsurval eftersom bortfallsproblemen vanligen är mycket stora.

Gösta Forsman.

»Transparens är ett nyckelord när det gäller redovisning av undersökningsresultat.»

Man kan därför konstatera att det vid båda rekryteringsformerna görs undersökningar av skiftande kvalitet. Rapporten ska inte ses som ett inlägg i en debatt för eller emot användande av självrekryterade paneler. I stället koncentreras framställningen på vilka olika typer av information som bör redovisas.

Vi konstaterar att det ofta saknas en vetenskaplig grund för att generalisera resultat från webbpanelundersökningar av allmänheten till hela populationen. Webbpaneler är därför i de flesta fall mindre lämpliga att använda när man vill skatta nivåer i populationen med god precision. Webbpaneler kan i stället övervägas i följande situationer:

• För sådana engångsundersökningar där man i första hand vill undersöka förekomsten av olika egenskaper hos populationen eller om undersökningen syftar till att generera idéer rent allmänt och inte fokuserar på exakta nivåer. Någon teoretisk grund för att beräkna osäkerhetstal finns inte. Under denna punkt faller olika typer av kvalitativa undersökningar liksom undersökningar som annars skulle ha genomförts med ”på-stan-intervjuer”, det vill säga där datainsamlingen skulle ha skett genom intervjuer av förbipasserande vid ett fåtal mätpunkter.

• Om kraven på noggrannhet i nivåskattningar är högre kan panelen ändå vara användbar ifall man kan verifiera empiriskt att urval från panelen ger tillförlitliga skattningar. Detta ställer krav på att man samlar på sig kunskap om hur urval från panelen fungerar vid mätning av olika variabler. Möjligheter att jämföra resultaten med pålitlig statistik (som en traditionell sannolikhetsbaserad undersökning eller en totalräkning) bör då finnas. Även här gäller att osäkerhetstal inte kan beräknas.

• För återkommande undersökningar där man

■ Webbpanelkommittén har bestått av följande personer:
Gösta Forsman, adj. professor em, Linköpings universitet (ordförande).
Mats Nyfjäll, fil.lic., seniorkonsult, Statisticon AB.
Åke Wissing, seniorkonsult, Åke Wissing & Co.
Jan Wretman, professor em, Stockholms universitet.
Karin Dahmström[1], universitetslektor, Stockholms universitet.

Dessutom har Henrik Kronberg, Norstat, deltagit som rådgivare. Bengt Larsson, Svenska ISO-kommittén, har fungerat som samordnare i definitionsfrågor med ISO-standarderna 26362 och 20252.

»»» RAPPORT FRÅN WEBBPANELKOMMITTÉN

varje gång använder exakt samma frågeformulär och procedurer i övrigt och därför (optimistiskt) kan anta att ett eventuellt systematiskt fel (på grund av bortfall) är lika stort och har samma riktning varje omgång. Intresset bör då vara riktat mot att upptäcka trender och inte mot nivåer. Antagandet om ”lika fel varje gång” bör man försöka styrka på något sätt.

• Webbpaneler kan också användas för experiment där panelen, eller en delgrupp av panelen, slumpmässigt delas upp (randomiseras) i olika grupper som alla utsätts för olika ”behandling”. Man kan till exempel vara intresserad av att jämföra hur olika varianter av ett frågeformulär påverkar svarsfördelningar. Randomiseringen gör att grupperna blir jämförbara även om panelen i sig är snedfördelad i förhållande till populationen.

Ovanstående förutsätter att rutiner för panelunderhåll och undersökningsprocedurer är ambitiösa och väl genomförda.

Vad bör ingå i en kvalitetsredovisning?

Transparens är ett nyckelord när det gäller redovisning av undersökningsresultat. Vi försöker dock inte ge några enkla tumregler för vilka slutsatser som ska dras av informationen. Vår förhoppning är snarare att informationen ska kunna ge upphov till diskussion och ökad medvetenhet om den osäkerhet, som alltid finns hos resultat från sådana här undersökningar.

I en kvalitetsredovisning ska ingå dels viss verbal information (se också ovan nämnda ESOMAR- och ISO-dokument), dels vissa numeriska beskrivningsmått.

De verbala beskrivningarna är av stor vikt, men vi anger ingen exakt mall för hur den verbala informationen ska utformas. Det viktigaste kan dock sammanfattas i följande tre punkter:

• Hur rekryteringen till panelen har gått till.

• Hur urvalet från panelen till den specifika undersökningen har gjorts.

• Hur eventuell vägning har gjorts vid beräkning av skattningar.

De numeriska beskrivningsmåtten kan sammanfattas i följande nio punkter.

A. Mått som beskriver rekrytering och deltagande:

• Kumulativ deltagarandel. Detta mått beräknas som produkten av rekryteringsandel och profilandel[3] samt deltagarandel i den specifika undersökningen. Det är jämförbart med svarsandelen i en traditionell undersökning.

B. Mått som beskriver själva panelen:

• Panelstorlek vid visst datum. Panelstorleken räknas i antal ”aktiva panelmedlemmar”, vilket bland annat innebär att rekryteringen och samtycke skett enligt vissa regler och att medlemmen deltagit i minst en undersökning senaste året (om medlemmen varit med hela året).

• Andel nyrekryterade. Andel aktiva panelmedlemmar som vid en viss tidpunkt varit med ett år eller kortare i panelen.

• Uppgiftslämnarbörda. Här föreslås två mått: genomsnittligt antal inbjudningar och genomsnittligt antal besvarande enkäter per aktiv panelmedlem under ett givet år.

• Utträdesandel (engelska attrition rate). Andel aktiva medlemmar som utträtt ur panelen under ett referensår. Man kan här dela upp måttet efter utträdesorsak, till exempel frivilligt eller ofrivilligt utträde.

• Dominans. I många paneler står en relativt liten del av medlemmarna för en stor andel av alla inkomna svar. Här föreslås ett mått som anger hur stor del av samtliga svar, under ett år, som de 20 procent mest aktiva panelmedlemmarna har bidragit med.

C. Mått som beskriver den specifika undersökningen:

• Anpassat svarsbeteende (engelska conditioning). Här föreslås mått som bygger på jämförelser mellan medlemmar som varit med länge och nyare medlemmar vad beträffar till exempel mediantid för undersökningen eller angivna svar.

• Nåbarhetsandel. Ett mått på kvaliteten i e-postadresserna fås genom att beräkna andel inbjudningar som inte når medlemmarna.

• Avbrottsandel. Andel enkäter som öppnats men avbrutits och klassats som ”ej svar”.

Observera att deltagandemåtten (under A ovan) bara kan beräknas vid rekrytering med sannolikhetsurval. De övriga måtten kan däremot beräknas både vid självrekryterade paneler och vid paneler rekryterade genom sannolikhetsurval.

Avslutande kommentarer

Rapporten finns som nämnts i en pdf-version på Surveyföreningens webbplats och föreligger också i en tryckt version som med fördel kan delas ut vid seminarier och kurser där webbpanelundersökningar behandlas. Den som vill hjälpa till att på detta sätt sprida rapporten är välkommen att kontakta någon i styrelsen eller i kommittén[4]. Rapporten är avsedd för en svensk läsekrets och därför skriven på svenska. Surveyföreningens styrelse har dock annonserat intresse för en engelsk version.

Planer finns att upplåta plats på Surveyföreningens webbplats för panelleverantörer att beräkna och redovisa värden på måtten för sina egna paneler och panelundersökningar. Tanken är naturligtvis att underlätta jämförelser mellan olika paneler för den som står i begrepp att beställa en webbpanelundersökning, men också att bygga upp erfarenheter av vilka värden på beskrivningsmåtten som är möjliga och lämpliga att uppnå.

GÖSTA FORSMAN

Referenser

ESOMAR (2012). 28 questions to help buyers of online samples. www.esomar.org.

SS-ISO 26362:2009. Accesspaneler inom marknads-, opinions- och samhällsundersökningar - Vokabulär och servicekrav (ISO 26362:2009, IDT).

SS-ISO 20252:2007. Marknads-, opinions- och samhällsundersökningar - Vokabulär och servicekrav (ISO 20252:2006, IDT).

Fotnoter

1. Karin Dahmström avled i juni 2012. En minnesartikel infördes i Qvintensen nr 3/2012.

2. Här används ibland metoder som skulle kunna betecknas som ”approximativa sannolikhetsurval”.

3. Profilandel är andelen som svarar i profilundersökningen.

4. Detta gäller även Surveyföreningens tidigare skrift Standard för bortfallsberäkning som också finns som pdf-fil på Surveyföreningens webbplats.

Bästa uppsats

■ Surveyföreningens pris för bästa uppsats på surveyområdet vanns i år av Andrea Gustafsen och Bergdís Sigurjónsdóttir som skrev kandidatuppsats på Statistiska institutionen vid Stockholms universitet. Uppsatsen har titeln Coding Process and Quality Control: Literature overview and the Case of Statistics Sweden. De skriver om sitt arbete i nästa nummer av Qvintensen.

NYA I STYRELSEN/

Karin Schneller, ledamot

– Jag jobbar sedan 1996 som statistiker och projektledare på TNS Sifo. Det är mycket roligt att arbeta med marknadsundersökningar, mina arbetsuppgifter är både omväxlande och spännande. Exempel på arbetsuppgifter är undersöknings- och urvalsdesign, konstruktion av frågeformulär, databearbetning, statistiska analyser (ofta multivariata analyser, även tidsserieanalys) och att dra slutsatser från alla data. Arbetet med kunder från olika branscher och olika ”business solutions” (t ex varumärkes- och kundundersökningar, produkt- och koncepttester, kartläggning av attityder och köpbeteenden) är stimulerande och lärorikt.

Ursprungligen kommer jag från Tyskland. Jag är civilekonom och statistiker och tog examen vidUppsala universitet 1994.

Åke Wissing

– Jag är statistiker och företagsekonom. Sedan mer än 40 år tillbaka arbetar jag som oberoende, fristående och renodlad konsult i marknadsundersökningar. Jag har ansvarat för undervisningen i marknadsundersökningar på Marknadsinstitutet och på Stockholms universitet. Jag har några specialintressen bland metodfrågorna som jag tycker särskilit mycket om att arbeta med; dit hör kontaktmetod, bortfall och frågekonstruktion.

– Jag ser det som hedersuppdrag åt Svenska statistikfrämjandet att ha fått vara med i Bortfallskommittén (Standard för bortfallsberäkning) och Webbpanelkommittén (Kvalitet i webbpanelundersökningar).

Anders Järleborg, ledamot

– Jag kom ut som nybakad statistiker år 1996 och har därefter jobbat inom olika statistiska metodområden som till exempel tidsserier och prognoser samt sambandsanalyser av skiftande slag. Under de senaste sju åren har jag jobbat som internkonsult inom surveyområdet på Socialstyrelsen där det genomförs en mängd olika typer av surveyundersökningar. Jag har således brottats med urvalsansatser, bortfallsproblem och mätproblem med mera under den senaste tiden. Privat är jag en gift 55-åring som trivs i norra Roslagen och då helst i en kajak.

Marcus Berg, webbansvarig

– Jag arbetar som adjunkt på Statistiska institutionen på Stockholms universitet där jag undervisar i statistik på grundnivå. Jag har tidigare arbetat som webbanalytiker, en analysgren där man tittar på människors beteendemönster online. Jag har även arbetat en del med grafisk design och bloggat tidigare. Jag blev specifikt nominerad till Surveyföreningens styrelse för att ta ansvar för, uppdatera, och förvalta föreningens hemsida, så det kommer att vara mitt huvudsakliga uppdrag i styrelsen under det kommande året.

– Jag har ett stort spelintresse – tv-spel, kortspel och brädspel. Jag spelar dock knappt alls längre. Jag tycker mycket mer om att läsa om nyheter och trender inom dessa områden och sedan diskutera detta med vänner och bekanta.

– Forskningsmässigt intresserar jag mig för alternativa datainsamlingsformer, framför allt i samband med internet. Mycket av detta faller under paraplybegreppet Big Data, och jag sitter för tillfället med i en task force som är initierad av AAPOR (American Association for Public Opinion Research) vars uppgift är att ta fram en rapport kring hur Big Data kan användas i samband med surveyer.

FMS

ORDFÖRANDEN HAR ORDET

»Nu är det dags för dig att ta över stafettpinnen!»

Jag har av och till haft tanken på att engagera mig i FMS utan att det lett fram till annat än just en tanke, men när vår utvecklingsansvarige kom med så tydliga instruktioner behövde jag inte fundera länge. Min förhoppning var att få komma in i styrelsen och få ett läroår innan eventuella tyngre uppdrag kunde vara aktuella. När sedan valberedningen kom med en förfrågan om jag kunde tänka mig ordförandeposten var det förstås oväntat, men efter att jag fått smälta frågan ett tag kände jag att det kunde vara en utvecklande och spännande utmaning. Här är jag nu, och ser fram emot ett spännande år med FMS!

Vårmötet inklusive årsmötet hölls i år på Stockholms universitet, och jag vill börja med att tacka programkommittén för att de lyckades få ihop ett program som uppvisade en härlig mix av olika tillämpningsområden av statistik inom det medicinska området. Sedan vill jag naturligtvis tacka talarna som fyllde programmet med innehåll. Jag blev både imponerad och inspirerad av vad som kan uppnås med en så hög ambitionsnivå, och med engagemang att verkligen fördjupa sig inom ett område.

Under dagen avhandlades även årsmötet utan större dramatik. Styrelsen för FMS för 2014 består av mig som ordförande (nyval), Nasser Nuru Mahmud som kassör (omval), Caroline Weibull som sekreterare (nyval), Karin Jensevik som arkiv- och hemsidesansvarig (omval), Frank Miller som FMS representant i Statistikfrämjandets styrelse (omval) samt Holger Kraiczi och Anna Berglind som ledamöter (omval). Som ny representant i European Federation of Statisticians in the Pharmaceutical Industry (EFSPI) valdes Magnus Kjaer. Till styrelsen finns även Marie Linder knuten som kontakt med Qvintensen.

»Jag blev både imponerad och inspirerad av vad som kan uppnås med en så hög ambitionsnivå»

Så vad är nu nästa steg? Enligt arbetsbeskrivningen skall ordförande ansvara för att lägga upp en långsiktig planering. Än så länge har jag haft fullt upp med att sätta mig in i FMS verksamhet vilket i huvudsak inneburit att titta bakåt i tiden snarare än framåt. Ett område som jag dock gärna vill få ett grepp om är medlemssituationen i FMS. Var finns våra medlemmar, hur ska vi kunna rekrytera fler och vad har vi att erbjuda våra medlemmar, både nya och befintliga? Traditionellt har de flesta av FMS medlemmar arbetat inom läkemedelsindustrin och mycket av aktiviteterna har varit riktade åt det hållet. Dock är det rimligt att anta att andelen medlemmar som arbetar inom andra områden ökar, och att aktiviteterna breddas för att återspegla denna fördelning. Jag kommer sannolikt återkomma i detta ämne i senare spalter.

FMS har sedan någon månad tillbaka haft kontakt med ordföranden för PSI (Statisticians in the Pharmaceutical Industry), och fört diskussioner om och i så fall hur de två föreningarna skall kunna ha ett samarbete. En konkret åtgärd redan nu är att information delats avseende presentatörer på respektive förenings konferenser, vilket kan ge många nya uppslag. För övrigt får vi se vad denna kontakt kan leda till. Då PSI har ett omfattande utbud av kurser/workshopar/konferenser och publikationer så finns det mycket i ett sådant samarbete som våra medlemmar skulle kunna ha nytta av.

Det är nu beslutat att årets höstmöte kommer att hållas på Hilton på Köpenhamns flygplats den 2 oktober, och det arrangeras i år i samarbete med DSBS (Danish Society for Biopharmaceutical Statistics). Även om temat och alla talare i skrivande stund inte är klara så ser jag personligen mycket fram emot att få träffa och lyssna på Stephen Senn, vars bok ”Statistical issues in drug development” varit mig till mycken nytta. Mer detaljer kring detta evenemang har redan skickats ut till FMS medlemmar och jag hoppas jag inte utlovar för mycket om jag säger att mer information om evenemanget kommer i nästa nummer av Qvintensen.

Nu ser vi fram emot en lång och skön sommar där vi hoppas på tur med den nordatlantiska oscillationen!

MARCUS THURESSON

NYA I FMS:S STYRELSE

Caroline Weibull, ny sekreterare

– Jag är från Stockholm och tog min magisterexamen i matematisk statistik på Stockholms universitet år 2009. I slutet av min utbildning fick jag möjligheten att skriva mitt examensarbete på institutionen för medicinsk epidemiologi och biostatistik (MEB) på Karolinska Institutet, och där har jag blivit kvar! Jag arbetar främst med registerbaserad forskning, med fokus på reproduktiv epidemiologi och canceröverlevnad. Det bästa med mitt jobb är balansen mellan forskning, undervisning och min egen utveckling som statistiker – och möjligheten att arbeta i en stor grupp av fantastiska biostatistiker från hela världen.

Sedan mars i år är jag doktorand på halvtid, vilket är spännande och utmanande. På min fritid njuter jag av tiden med min ettåriga son samt semestrar med familjen, ofta med inslag av vandring och bergsbestigning.

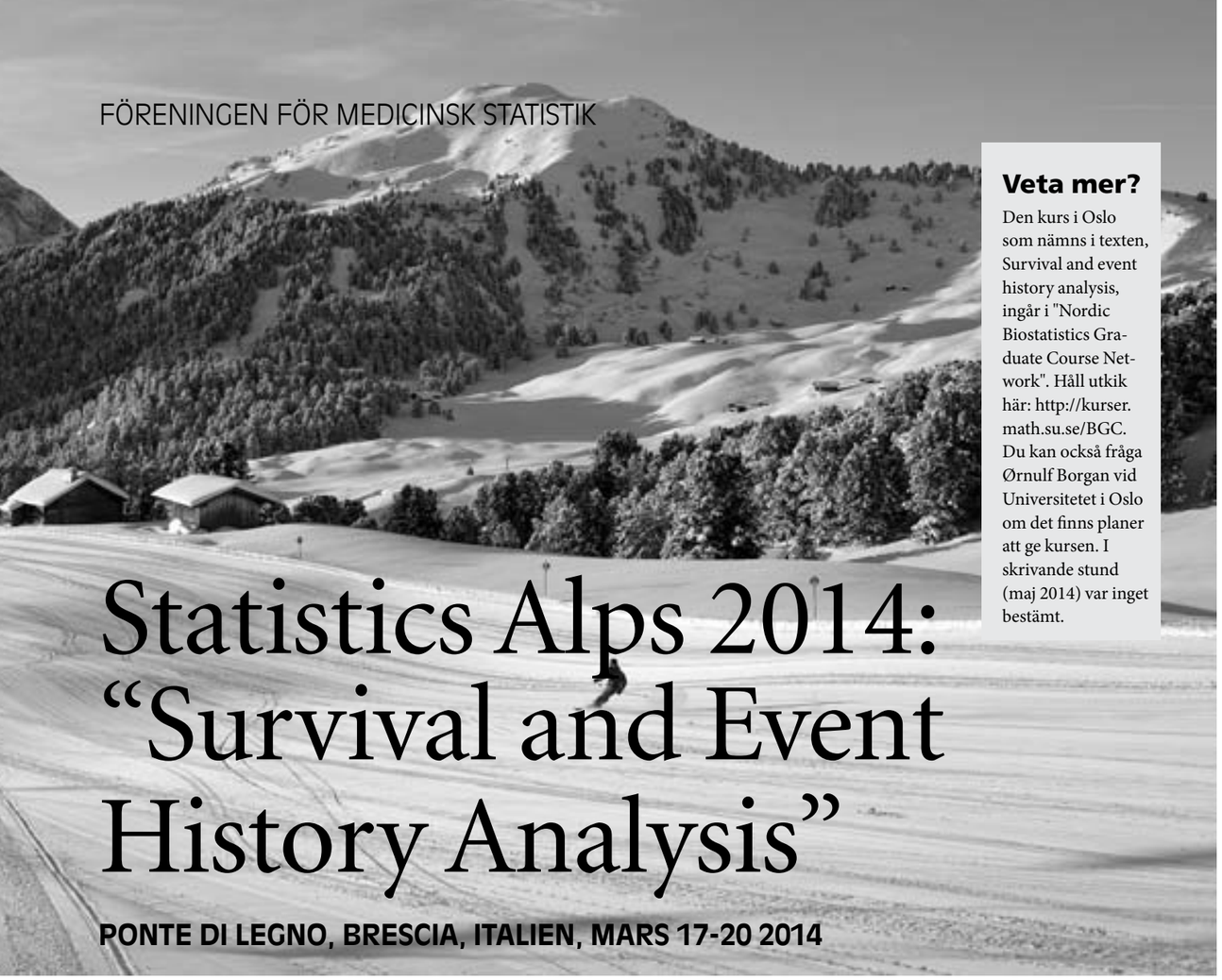
Och eftersom jag ofta får frågan om mitt efternamn; min mans farfarsfar var kusin med Waloddi Weibull – som var professor vid KTH och definierade den statistiska Weibullfördelningen.

Jag ser fram emot ett intressant år som sekreterare i FMS!

Marcus Thuresson, ny ordförande

– Den yrkesmässiga vägen fram till den plats där jag är idag har varit långt ifrån spikrak, och ett antal mer eller mindre medvetna vägval har haft avgörande betydelse. Jag är född och uppvuxen i den lilla byn Pilgrimstad i Jämtland där det inte fanns några vidare traditioner med högre utbildning, vilket sannolikt hade stort inflytande över mitt val att utbilda mig till byggnadsarbetare. Även om yrket som sådant har många positiva sidor var det de negativa som övervägde (läs: kalla vintrar). Efter kompletterande studier på Komvux bestämde jag mig för att flytta till Stockholm och studera till sjukgymnast, och om jag har ett beslut i mitt liv som jag inte kan komma ihåg hur jag motiverade det, så är det just detta. Efter sjukgymnastexamen fick jag förfrågan om att arbeta vidare med och publicera min C-uppsats, vilket så småningom ledde till både magister- och doktorsexamen i sjukgymnastik. Under forskarutbildningstiden insåg jag att statistik var det som jag verkligen brann för och parallellt med forskningen på Karolinska Institutet cyklade/gick jag vägen runt Brunnsviken till universitetet och läste kurser i statistik. Sålunda kunde jag samma år, 2005, plocka ut en kandidatexamen i statistik och en doktorsexamen i sjukgymnastik.

Sedan 2006 arbetar jag som biostatistiker på Statisticon i Uppsala med väldigt varierande typer av uppgifter och kunder. Dock kan mina uppdrag i stort sett delas in i tre huvudkategorier: registerstudier, uppdrag åt läkemedels- och biotechföretag, samt konsultationer med forskare inom akademin. Även om uppdragen är väldigt olika har de gemensamt att allesammans oftast är väldigt inspirerande och ständigt kommer med nya spännande utmaningar.



Statistics Alps 2014: "Survival and Event History Analysis"

PONTE DI LEGNO, BRESCIA, ITALIEN, MARS 17-20 2014

FOTO: WOLFGANG MORODER

Veta mer?

Den kurs i Oslo som nämns i texten, Survival and event history analysis, ingår i "Nordic Biostatistics Graduate Course Network". Håll utkik här: http://kurser.math.su.se/BGC. Du kan också fråga Ørnulf Borgan vid Universitetet i Oslo om det finns planer att ge kursen. I skrivande stund (maj 2014) var inget bestämt.

Sedan några år tillbaka anordnar centrum för biostatistik på Bicocca-universitet i Milano en statistikkurs i Alperna en gång per år. Varför i Alperna kan man undra – men när vi anlände till denna lilla alpby i vårvärmen (18+ grader!) med dramatisk inramning av bergen i bakgrunden var vi rätt nöjda över detta faktum.

Årets kurs gavs av några välkända norska namn: Odd O. Aalen, Ørnulf Borgan samt Håkon K. Gjessing. Dessa tre gav 2008 ut boken "Survival and event history analysis – a process point of view". Kursen baserades helt på denna bok och författarna gav smakprov på varje kapitel genom att ge översiktliga föreläsningar. I Oslo brukar författarna ge motsvarande kurs under två veckor, men i Ponte di Legno gav de en nerbantad kurs på fyra dagar. Det var en utmaning både för dem och för oss kursdeltagare att navigera genom det digra materialet.

På Karolinska Institutet är vi involverade i doktorandkursen i överlevnadsanalys för epidemiologer. Vi undervisar i detta från ett tolkningsfokuserat perspektiv, vilket medför att man som lärare ibland kommer bort från att tänka på överlevnadsanalys som stokastiska processer med allt vad det innebär. Därför var det med skräckblandad förtjusning vi satte oss i undervisningssalen den första dagen.

I kort behandlades följande: räkneprocesser, martingaler och stokastiska integraler, Nelson-Aalen-estimatorn av den kumulativa intensitetsprocessen, regressionsmodellering inom överlevnadsanalys ur ett processperspektiv, frailtymodeller, modellering av återkommande händelser, samt kausalinferens inom överlevnadsanalys.

För vår egen del, eftersom detta inte var vårt första möte med överlevnadsanalys, var undervisningen inom processteori det mest intressanta. Det var spännande att se hur denna teori med hjälp av martingaler så elegant ligger till grund för de analyser som vi genomför dagligen i vårt arbete.

Vid sidan av de intensiva kursdagarna fick vi möjlighet till några timmars skidåkning eller annan fri aktivitet. Dessutom anordnades det utflykt till 2000 meters höjd för lunch och middag på annan ort. Det var, som alltid, väldigt intressant att träffa statistiker från andra länder i Europa. Majoriteten var från Italien, men andra deltagare kom bland annat från Tyskland, Nederländerna, Storbritannien, Frankrike och Danmark.

Den längre Oslo-kursen kan vi varmt rekommendera om den utannonseras igen, för såväl doktorander som statistiker i arbetslivet. Förutom de faktiska kunskaperna vi tog med oss hem påmindes vi om hur viktigt det är att då och då utmana sig själv. Det är också viktigt att gå tillbaka och läsa teoretiska kurser, även några (eller många!) år in i ens yrkesverksamma liv som statistiker.

Statistic Alps återkommer igen i september 2015. För mer information: www.statmed.medicina.unimib.it/statisticalps2014/statisticalps.htm

CAROLINE WEIBULL OCH ANNA JOHANSSON
KAROLINSKA INSTITUTET

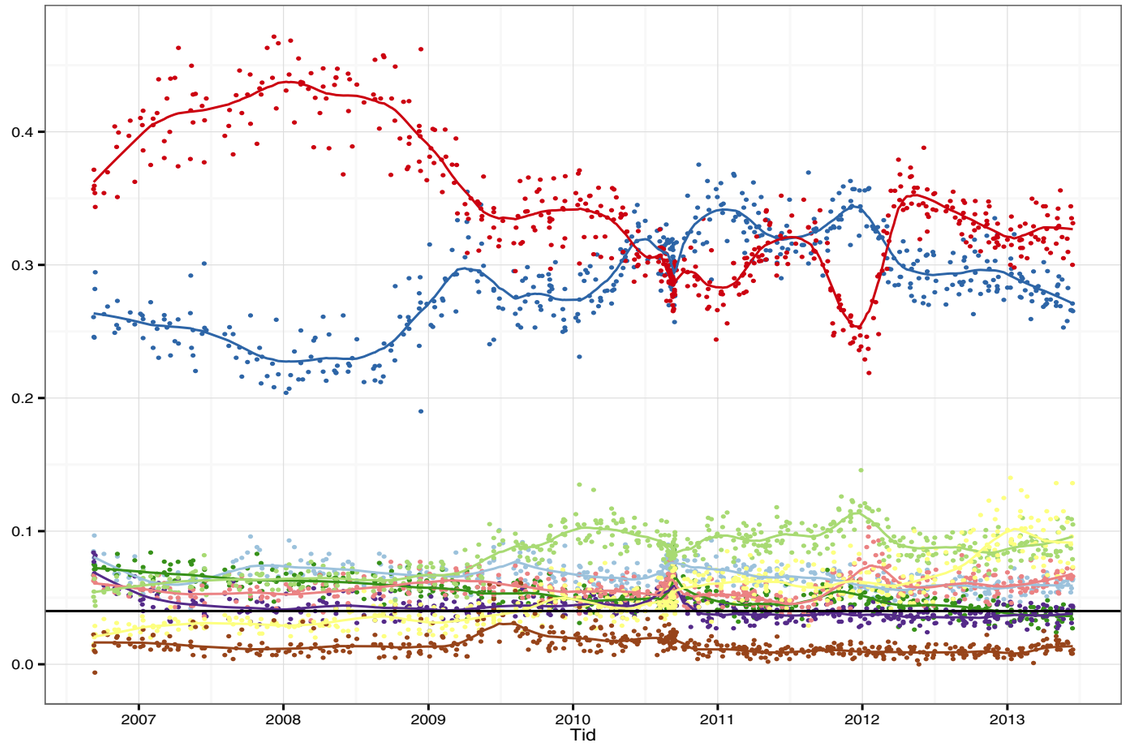


ILLUSTRATION TILL ARTIKELN: ATT VÄGA IHOP OLIKA OPINIONSUNDERSÖKNINGAR (SIDAN 10)

Figur 1. Resultatet av 509 partisympatiundersökningar mellan september 2006 och juni 2013, samt för varje parti en sammanvägd skattning (heldragen kurva) av alla skattade partisympatier för det partiet. Data från Novus.

Har du flyttat?

■ Meddela din nya adress genom att skicka ett e-brev till Thorbjörn Gudmundsson på sekrfram@gmail.com.

Så blir du medlem

Svenska statistikfrämjandet

■ 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 Thorbjörn Gudmundsson på sekrfram@gmail.com. Du får Qvintensen i brevlådan och platsannonser via e-post.

Det ställs inga krav för att bli medlem; alla som är intresserade av statistik och vill stödja statistikens roll i samhället är välkomna.

Qvintensen

POSTTIDNING B

SVERIGE
PORTO BETALT
PORT PAYÉ

Avsändare/Sender: Svenska statistikfrämjandet
c/o: T Gudmundsson, Institutionen för matematik,
KTH, 100 44 Stockholm, Sweden

Illustrationer till artikel på sidorna 20-21.

COMPLEX MODELS FOR COMPLEX DATA VIA THE BAYESIAN APPROACH

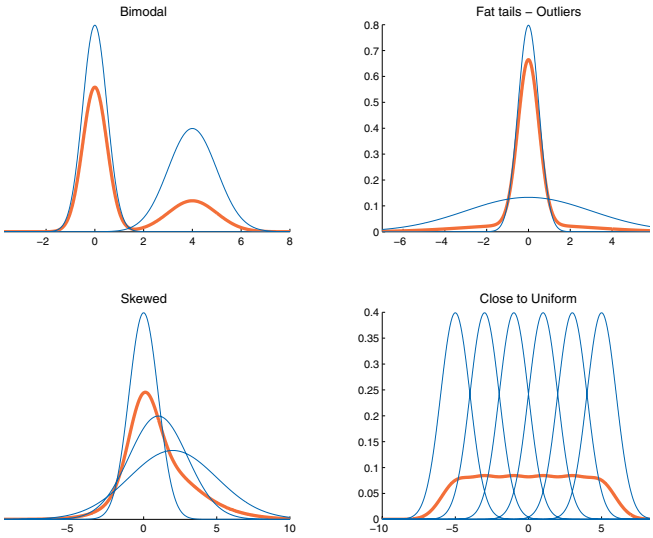


Figure 1. Using mixture of normal densities (blue lines) to approximate a density (red line); four examples of rather difficult densities to model with other methods.

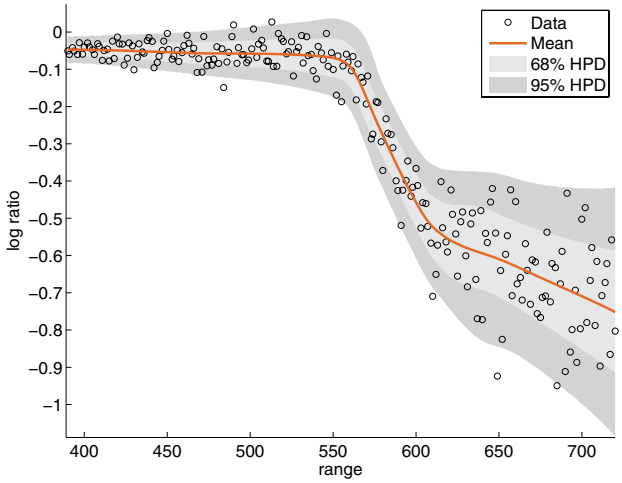


Figure 3. Smooth mixture models for the LIDAR data. The figure displays the actual data overlayed on predictive regions and the predictive mean.

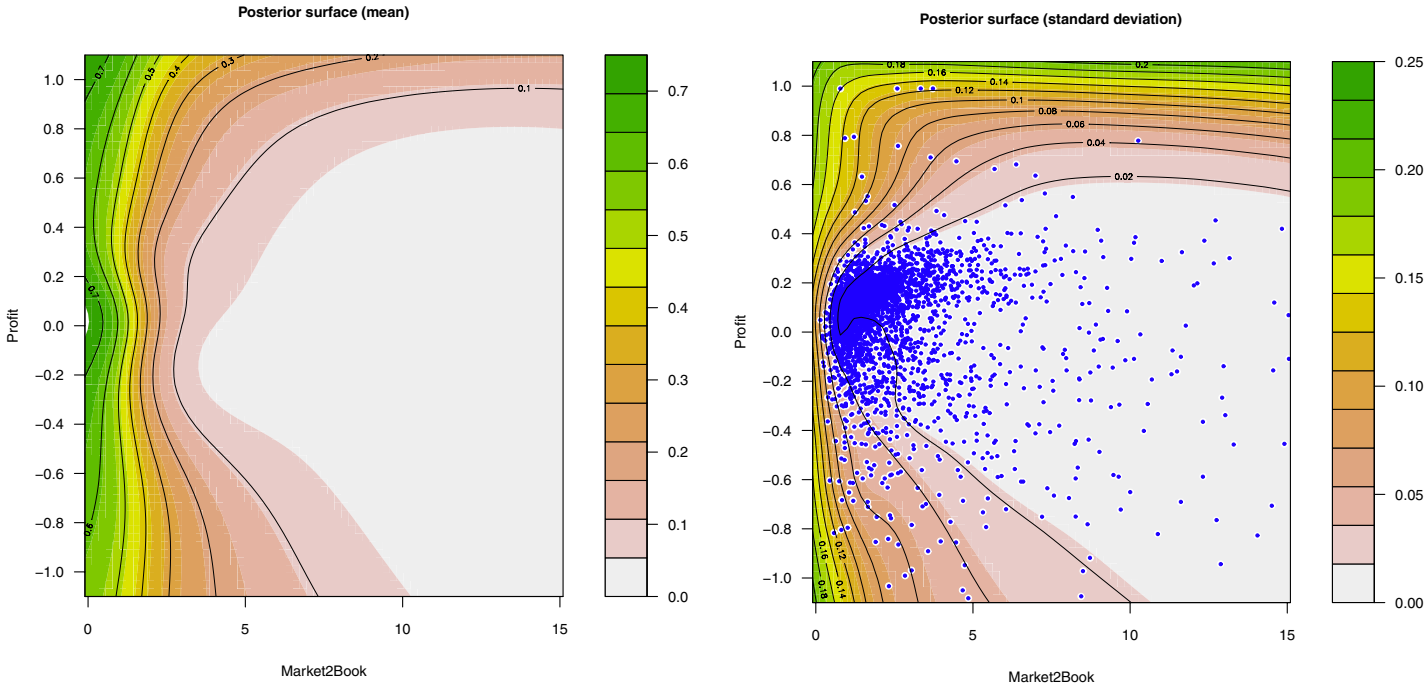


Figure 4. The posterior mean (left) and standard deviation (right) of the posterior surface for the model for firm leverage data. The depth of the colour indicates how the leverage varies with book values and profits. The subplot to the right also shows an overlay of the covariate observations (blue).